1. Introduction

"*Goldman Sachs is using ChatGPT-style AI in-house to assist developers with writing code.*"
(CNBC, 2023)

"ChatGPT Stock Predictions: 10 Stocks That Have 10X Potential"
(Yahoo Finance, 2024)

"Morgan Stanley CEO says AI could save financial advisers 10-15 hours a week"
(Reuters, 2024)

The rapid advancement of AI is transforming the finance and accounting industry, prompting the reevaluation of roles traditionally held by humans (Acemoglu et al., 2010; Mcelheran et al., 2024). While AI offers significant benefits for financial analysts, auditors, or portfolio managers (Asparouhova et al., 2024; Aubry et al., 2023; Estep et al., 2024; Fedyk et al., 2022; Hodge et al., 2021; Lopez-Lira & Tang, 2023), it also raises some concerns about the (potential) displacement of skilled workers, its broader economic effect, and reliability issues (Acemoglu, 2024; L. Chen et al., 2024; Mollick & Mollick, 2022). Recent advances in large language models (LLMs), such as ChatGPT, have accelerated AI capabilities, intensifying these debates. However, the critical question emerges: To what extent do investors utilize AI in financial decision-making?

On the one hand, the recent empirical literature finds that investors could deem AI technology credible. For instance, Cao et al. (2024) show that a well-trained AI model beats most humans.[1] However, humans maintain an edge when institutional knowledge is important, showing that a combination of *Man + Machine* has potential.[2] Glikson and Williams Woolley (2020) highlight

[1] There is a larger literature discussing analyst forecast quality (e.g., De Silva & Thesmar, 2023; Gloria et al., 2012; Kadous & Thayer, 2009; Zhang, 2006).
[2] This conclusion aligns with other papers showing the promise of artificial intelligence (Cao et al., 2024; Chak et al., 2022; J. Chen et al., 2024; Hodge et al., 2021; Jha et al., 2024; Lopez-Lira & Tang, 2023).

that people are more likely to (1) follow highly intelligent machines, even when they are faulty, and (2) trust them generally more on technical or data-driven tasks. This implies that receiving AI-generated information corresponds to increased trust and utilization.

On the other hand, there are sizeable differences in the performance between machine learning tools (Gu et al., 2020) and concerns about ChatGPT's performance decreasing over time (Chen et al., 2024). Eslami et al. (2015) highlight negative repercussions when people are unaware of AI usage. Moreover, Longoni et al. (2019) find evidence of "algorithm aversion" in the medical sector (e.g., Germann & Merkle, 2023; Önkal et al., 2009). This implies that investors could have less trust in AI, particularly when the provider is revealed afterward. In this paper, we conduct three incentivized experiments to examine these competing hypotheses, done through the lens of analyst forecasts.

Taken together, our findings highlight the key role of the forecast provider in belief formation. We document that, although investors update their beliefs about expected returns toward the forecast, they are *less* responsive when analysts incorporate AI. In other words, investors seem to trust human analysts relatively more. This stems from the perception that AI-incorporated forecasts lack credibility. We also reveal some nuances: women, Democrats, and investors with high AI literacy are *more* responsive to AI forecasts, while perceived model complexity reduces the likelihood of updating return expectations. This emphasizes the complex relation between conventional financial advice and emerging technologies.

In Experiment 1, we randomly assign 1,800 US participants to one of three groups: *Man*, *Man + Machine*, or *Machine,* and study the impact of forecast provider on trust. To make this setting realistic and avoid deceiving participants, we use the wording of Goldman Sachs' stock market outlook of November 2023 and only change the source to (1) "The analysts of Goldman Sachs" (*Man*), (2) "Analysts of Goldman Sachs incorporating an advanced AI model" (*Man + Machine*), (3) or "An advanced AI model (*Machine*)". More specifically, we keep the forecast estimate and text constant, changing only the source to isolate its effects. This setup ensures that differences in the posterior variables are due solely to the forecast source. Moreover, it mirrors the current landscape of financial advising, where there is an interest in robo-advisory services (e.g., Chak et al., 2022; Hodge et al., 2021; Rossi & Utkus, 2020) and the integration of LLMs by investment banks.

Our findings paint a nuanced picture. While investors generally update their beliefs positively in the direction of new information (i.e., the 12-month ahead return forecast of 5% for the S&P 500), the forecast provider plays a key role. For instance, we show that the average investor in our sample is less likely to update his return beliefs toward the signal when it comes from the combination of *Man + Machine*. If this treatment is effective, participants will align their return beliefs with the provided forecast. However, the results imply that investors react less toward AI-generated forecasts.

Our results also reveal heterogeneities in this pattern. First, we document that AI forecasts are perceived as less credible than those from humans. Perceived credibility plays a crucial role in shaping investor's beliefs. We conclude that these insights have implications for designing and implementing AI in financial decision-making, highlighting that technical excellence by itself cannot generate trust. It also shows the key role of AI model disclosure in analyst forecasts by making such reports more trustworthy for the average investor.

Second, on average, women exhibit a greater propensity to positively update their return belief to AI-generated forecasts, particularly when they have larger initial misperceptions. Indeed, while the average investor moves away from the signal from AI sources, female investors seem to update their beliefs *toward* the signal. This conclusion contributes to the growing discussion about gender differences in financial decision-making and technology adoption (Bhattacharya et al., 2024; Bucher-Koenen et al., 2023).

Third, we show differences across political affiliations. More specifically, Democrats are more likely to update their return beliefs in line with AI forecasts, particularly when there is a larger gap between their prior beliefs and the forecast. In other words, Democrats are more receptive to AI-generated forecasts. This finding is consistent with the literature that political affiliation shapes expectations, such as stock market beliefs (Barrios & Hochberg, 2021; Cookson et al., 2020; Gerber & Huber, 2009; Leblang & Mukherjee, 2005; Meeuwis et al., 2022) and technology adoption (Blank & Shaw, 2015).

Finally, higher AI literacy corresponds with a significantly larger return belief update toward the signal when receiving a *Man + Machine* forecast. Previous studies highlight the importance of different types of literacy (Filippini et al., 2024; Lusardi & Tufano, 2015), in making financial decisions. Given the rising importance of AI in the finance and accounting industry, our paper underscores the relevance of improving AI education, particularly among investors.

Building on the baseline evidence, we examine whether using a more familiar AI tool, such as ChatGPT, mitigates the observed trust deficit. Importantly, ChatGPT can also be used for stock return forecasts, and it has been shown to have potential (e.g., Lopez-Lira & Tang, 2023). For Experiment 2, we recruited 600 new participants to study this issue. Contrary to prior research suggesting familiarity enhances trust in technology (Lee & See, 2004), our results indicate that replacing AI with ChatGPT does not improve investor trust. In fact, we document evidence of negative reactions to ChatGPT-generated forecasts. This indicates investors distrust ChatGPT-generated advice, perhaps even more than "the generic AI model." This finding aligns with the literature showing the decreasing accuracy of ChatGPT (Lingjiao Chen et al., 2024), implying that the average investor is potentially aware of these issues.

In Experiment 3, we recruit 600 participants to highlight that our observed credibility findings are robust by making the forecast source more salient. In this setup, we first show the forecast without a source. After retrieving the respondents' posterior return expectations, we randomly assign them to one of the treatment groups (*Man*, *Man+Machine*, or *Machine)* and tell them who made the forecast. We find that the credibility of AI-generated forecasts decreases significantly once the forecast source is revealed. This result aligns with Eslami et al. (2015), who argue that this evokes anger in people. Overall, the experiment confirms our main findings and suggests a causal link between the forecast provider and perceived credibility. The increased skepticism of investors underscores the challenges of integrating (new) technology into financial decision-making (Davis, 1989; Dietvorst et al., 2014; Lee & See, 2004; Longoni et al., 2019).

In our baseline experiment, we introduce three manipulations to further examine the nuances of investor reactions in various scenarios. First, we vary the information richness by providing participants with a “basic” forecast, one with an earnings estimate, and a more comprehensive package, including visual aids (see Bazley et al., 2021; Bradshaw, 2004). Second, we manipulate analyst forecast dispersions, presenting either high or low analyst consensus (e.g., Palley et al., 2024). Finally, we introduce an analyst downgrade with and without accompanying earnings estimates (e.g., Kecskés et al., 2017). Overall, the sign of return revisions to these manipulations aligns with the established results in the prior literature, lending credence to our experimental setting. More importantly, we show that the identity of the forecast sources (i.e., *Man*, *Machine*, or *Man + Machine*) does not interact with manipulations that impact trust perceptions.

Experiment 4 provides suggestive evidence that investors will update their return expectations differently depending on the disclosed statistical model, highlighting a preference for ordinary least squares (OLS) forecasts over more complex-sounding methods (i.e., Best Linear Unbiased Estimator or deep learning). This may be attributed to complexity aversion (Oberholzer et al., 2024; Umar, 2022). It underscores the multifaceted nature of investor responses to AI forecasts.

Our paper makes several key contributions. First, we contribute to the literature on the role of AI in analyst forecasts. There is a (rapidly) growing stream of the literature that focuses on the ability of AI to help financial analysts better dissect balance sheet information (Cao et al., 2024; Jha et al., 2024; Kim et al., 2024) and provide investment advice (Chak et al., 2022; Coleman et al., 2022; Hodge et al., 2021). We are the first to study investors' reactions to AI forecasts from a credibility and return beliefs perspective. This revolves around the question of trust in equity markets (Bhagwat & Liu, 2020; Guan & Tsang, 2020; Pevzner et al., 2015; Wei & Zhang, 2023) and in new technology (e.g., Davis, 1989; Glikson & Williams Woolley, 2020; Lee & See, 2004).

Second, a large emerging body of literature examines the impact of AI on asset predictability, including the effect of ChatGPT on financial markets (J. Chen et al., 2024; Jha et al., 2024; Lopez-Lira & Tang, 2023). These papers document that ChatGPT dissects economic news better than traditional methods and forecasts future returns. Our paper focuses on the human component by studying how much investors trust ChatGPT or other AI tools. Hence, it addresses a rapidly evolving technology and its impact on the finance and accounting industry (Acemoglu, 2021;

Acemoglu et al., 2010; Acemoglu & Estrepo, 2022). We show that investors perceive LLMs, like ChatGPT, as less credible. This has large implications for the implementation of these tools in financial decision-making.

Third, we advance the literature on algorithm aversion (e.g., Dietvorst et al., 2014; Germann & Merkle, 2023; Longoni et al., 2019; Önkal et al., 2009). Our paper makes several contributions. On the one hand, we offer a more realistic setting, distinguishing between three groups: *Man, Man + Machine,* and *Machine.* Indeed, the recent advances in LLMs, AI, and robo-advising have made investors more aware of such topics, reflecting the current landscape better. As such, we replicate Önkal et al. (2009) for a period with arguably higher AI awareness. On the other hand, by examining gender, political affiliations, and AI literacy variation, we add a novel dimension to the concept of "trust" in AI. Indeed, we highlight that credibility is a main driver of reduced trust in the forecasts. Building on these results, we find that complexity also plays an important role. Unlike previous studies focusing on algorithm aversion, after witnessing errors, we show that the forecast provider does not amplify the impact of its content on investor reactions. This indicates a more complex relation between content and provider than previously understood.

Finally, we contribute to the household finance literature on technology adoption and financial advice (Chak et al., 2022; Jha et al., 2024; Rossi & Utkus, 2020; Yang & Zhang, 2022). We showed that, on average, investors deem AI less credible than human forecasts despite prior evidence showing AI is better at data-driven and technical tasks and serving as a (potentially) powerful tool to improve financial decision-making. Furthermore, we highlight that AI literacy plays an essential role, implying that, as the participants become more (AI) educated, they become more responsive to its signals, contributing to the literature on the financial education of households.

2. Study design

As motivated by the introduction, we experimentally examine the impact of forecast providers on investor beliefs. We conduct four incentivized experiments to examine this.

[INSERT FIGURE 1]

Figure 1 plots the flow of the main experiment, which consists of three parts, discussed below.

2.1. Survey flow

Baseline, prior beliefs. In Part 1 of the survey, participants were asked about their priors. First, we measure risk preferences similar to Dohmen et al. (2021):

> *"Generally speaking, are you the kind of person who is willing to take risks or who prefers to avoid risks?"*

Responses are recorded from 1 (very risk-averse) to 7 (very risk-taking) (see, e.g., Borsboom & Zeisberger, 2020; Holzmeister et al., 2024).

To enhance the realism of our experiment, we incorporate questions about Goldman Sachs, a renowned investment bank. This allows us to test for prior beliefs on Goldman Sachs to ensure the integrity of the randomization in Part 2.

> *"How capable do you think Goldman Sachs' analysts are of making financial forecasts?"*

> *"In general, do you think analyses done by Goldman Sachs analysts can be trusted?"*

The questions are on a scale from 1 to 7, similar to above, and are inspired by YouGov surveys, while the latter is inspired by papers examining trust in financial institutions (Okat et al., 2024). We refer to these measures, respectively, as *Goldman Sach's capabilities* and *Goldman Sachs' trust.*

We include several questions related to AI. These are important to validate the integrity of the randomization of respondents in Part 2. More importantly, we add the following measures as control variables to ensure that the effects we document are not driven by heterogeneity in AI usage, trust, and relative capability beliefs:

> *"How capable do you think artificial intelligence (AI) is of making financial forecasts compared to Goldman Sachs' analysts?"*
>
> *"In general, do you think analyses done by artificial intelligence (AI) are more trustworthy compared to those done by Goldman Sachs?"*
>
> *"In general, how useful has ChatGPT (or similar artificial intelligence tools) been to you?"*

These questions are recorded on a scale from 1 to 7 and were inspired by YouGov surveys. We label the measures, respectively, as *AI relative capabilities*, *AI trust*, and *ChatGPT usefulness*.

To poll the expected return priors, we borrow the question of Giglio et al. (2024). Figure 2 plots the screen the participants see. Moreover, we ask them for their confidence in this estimate on a scale from one to seven, as in Ungeheuer and Weber (2021).

[INSERT FIGURE 2]

Our study focuses on the S&P 500 index, chosen for its status as a well-known and extensively covered US equity market index, enhancing our findings' generalizability. We implement an incentive structure for prior and posterior expected return estimates to encourage participants' best efforts. After one year, the computer randomly selects either prior or posterior responses. The five participants whose estimates most closely match the realized S&P 500 returns receive a monetary bonus of €50. This incentive method, similar to those employed in previous studies (e.g., Bauer et al., 2024), is designed to elicit accurate expected returns while avoiding potential hedging motives in participants' responses.

2.2. Investment information

Treatment stage. In Part 2 of the experiments, all participants are randomly assigned to one of three groups:

- ***Man***: *Analysts from Goldman Sachs forecast a 12-month return of 5% for the S&P 500 index. The baseline assumption during the next year is that the US economy will continue to expand at a modest pace and avoid a recession.*
- ***Man + Machine***: *Analysts from Goldman Sachs, incorporating advanced AI tools to enhance their analyses, forecast a 12-month return of 5% for the S&P 500 index. The baseline assumption during the next year is that the US economy will continue to expand at a modest pace and avoid a recession.*
- ***Machine*****:** *An advanced AI model, trained on financial data and market trends, forecasts a 12-month return of 5% for the S&P 500 index. The baseline assumption during the next year is that the US economy will continue to expand at a modest pace and avoid a recession.*

As can be observed, we keep the information in the text constant except for the source to isolate its effect. Since individuals are randomly assigned to one of the three messages, any differences in the posterior variables retrieved after the random assignments to the information group will be attributable to the source of the forecast.

In designing our setup, we made specific choices to enhance realism and credibility. First, we have adopted the wording from Goldman Sachs' November 2023 market outlook introduction. Goldman Sachs validates our setup as a renowned investment bank that issues annual equity predictions and already incorporates AI.[3][4] Second, we verify that AI tools (i.e., ChatGPT) could produce similar forecasts independently, confirming the plausibility of the information given to the participants (cfr. Figure A.1). This approach grounds our setting in an authentic financial industry practice while accurately representing current artificial intelligence capabilities.

Posterior beliefs. In all experiments, respondents are again asked to provide a one-year ahead forecast of the S&P 500 index after receiving one of the forecasts. Furthermore, we ask them to elicit the credibility of the forecast provider and risk perception beliefs (on a seven-point scale), which aligns with Kadous and Thayer (2009). Finally, we provide them with five questions on their AI and financial literacy (see Appendix II).

2.3. Participants

In Experiment 1, we recruited 1,800 participants via the online platform Prolific on August 27th and 28th, 2024. We focus on U.S. participants with investment experience in exchange-traded funds or stocks who are at least 18 years of age. We paid participants £9 per hour – as pounds are the standard currency on Prolific. The median time to complete the survey was 9 minutes.[5]

[INSERT TABLE 1]

Sample statistics. Panel A of Table 1 presents the summary statistics.[6] Participants expect the 12-month ahead return of 8.29% for the S&P 500, with a confidence of 4.92 out of 7. They report

[3] Appendix I highlights the introduction of Goldman Sachs' equity market outlook (Kostin et al., 2023).
[4] https://www.wsj.com/articles/goldman-sachs-deploys-its-first-generative-ai-tool-across-the-firm-cd94369b
[5] In total, we reject 44 participants who exhibit S&P 500 return expectations above 1,000% over the next 12-months or who fail an attention check. We collect additional participants to reach our pre-determined sample size of 1,800.
[6] We drop all participants with ignorant return expectations, below -30% or above 30% (Merkoulova & Veld, 2022). The results do not change when we did include return ignorant participants.

a risk preference of 4.07 on a seven-point scale and demonstrate a high level of AI[7] (4.06 on a 5-point scale) and financial[8] literacy (4.22 out of 5). Furthermore, they exhibit lower perceived trust in AI than Goldman Sachs (3.81 out of 7), while they deem it equally capable of Goldman Sachs (4.08 on a seven-point scale).

Integrity of the randomization. Columns 6 to 10 highlight the randomization of respondents. When randomly assigning participants to one of the groups, we find no significant differences across the average investor characteristics. We include these control variables in our regression specifications to eliminate concerns. Doing so ensures that other characteristics, such as prior beliefs in AI capabilities, gender, or age, do not drive changes in investor beliefs across forecast providers.

3. Updating across forecast providers

In Experiment 1, we study the effect of forecast providers on investor beliefs. More specifically, we calculate the differences in return updating across the different groups. Following Haaland et al. (2023), we define *Update* as the point estimate difference between the posterior and prior return expectations. In other words, it captures how much respondents changed their expected returns following the information treatment (i.e., the analyst forecast).

We calculate the participant's perception gap (*Perception gap*), which is the difference between "the provided signal" (i.e., the 5% forecast) and prior return beliefs. In other words, it quantifies how far away they are from the forecast signal. Larger perception gaps should lead to strong

[7] AI literacy is measured as the number of correct answers on five questions (see Appendix II).
[8] Financial literacy is measured as the number of correct answers on five questions (see Appendix II).

return updating if the participants believe this forecast signal is credible (Beutel & Weber, 2024; Haaland et al., 2023; Laudenbach, 2024). To study this, we use the following specifications:

$$Update_i = \beta_0 + \beta_1 Treatment_{t,i}\, x\ Perception\ gap_i + \beta_2 Treatment_{t,i} + \beta_3 Perception\ gap_i + \gamma X_i + \varepsilon_i \tag{1}$$

where $Update_i$ is the difference between posterior and prior return expectations for individual i. Since we are interested in the impact of receiving a forecast from AI versus (human) financial analysts, we define individuals receiving the *Machine* or *Man + Machine* treatment, while those receiving the *Man* message are in the control group. Hence, we choose an active control group, as recommended for this type of experiment (e.g., Haaland et al., 2023). The variable of interest is β_1, which captures the extent to which an investor updates their priors toward the provided forecast in the AI treatments relative to the control group, which only receives Goldman Sachs' forecast. This is referred to as "the learning rate" (Haaland et al., 2023; Laudenbach, 2024). The coefficient on β_2 quantifies the average amount of updating across different treatments which is not related to the participants' priors. In our case, this could be related to prior beliefs on AI. Therefore, we include several control variables, such as AI trust or capabilities. The coefficient on β_3 captures the average "learning rate" that is unrelated to the treatments. Finally, X_i is the vector of control variables, such as gender, risk preference, and age.

Expected Sharpe Ratio. Alternatively, we employ an alternative specification:

$$ESR_i = \beta_0 + \beta_1 Treatment_{t,i}\, x\ Perception\ gap_i + \beta_2 Treatment_{t,i} + \beta_3 Perception\ gap_i + \gamma X_i + \varepsilon_i \tag{2}$$

where ESR_i is a ratio between the posterior return point estimate and risk perception measure for individual i. This risk perception measure is retrieved through the following question:

"How risky do you perceive an investment in the S&P500 to be?"

This indicates that investors have to make an investment choice based on the information they receive. Risk perception is measured on a scale from 1 (Risk-free) to 7 (Very risky). We use this measure as the denominator to quantify the *Expected Sharpe ratio* (ESR) in Equation 2.

3.1. Main findings

Revisiting our key research question of whether investors trust AI-generated or AI-integrated analyst forecasts, Table 2 presents our results. It reveals a nuanced picture: We find substantial heterogeneity in how investors update their beliefs about future returns in response to the new information. On average, the learning rates range from 0.514 to 0.524 (all significant at the 1% level). This suggests that investors update significantly toward the provided signal, indicating that the average investor responds to the forecast. The learning rate is comparable to previous studies (e.g., Beutel & Weber, 2024).

[INSERT TABLE 2]

To address our hypothesis about investor responsiveness to AI-generated forecasts, we study how participants updated their beliefs based on different forecast sources: The combination of *Man* and *Machine* recommendations significantly diminishes this learning rate. The coefficient on the interaction of *Perception Gap* and *Man + Machine* is negative (significant at the 5% level), implying that individuals are less responsive to new information. In other words, the investors trust the *Man + Machine* forecast less than financial analysts. Overall, this adds to the previous literature on belief updating that generally finds a positive impact on belief updating through (new) information.

These results challenge the prevailing conclusions in the literature, which show *Man + Machine* combinations outperforming other configurations (Cao et al., 2024; Kim et al., 2024; Lopez-Lira & Tang, 2023). Instead, it aligns more closely with the conclusions on algorithm aversion (e.g., Dietvorst et al., 2014; Germann & Merkle, 2023; Önkal et al., 2009), suggesting concerns related to the perceived credibility of AI-incorporated financial decision-making (Davis, 1989; Glikson & Williams Woolley, 2020; Lee & See, 2004).

Our results are robust when considering risk-adjusted return expectations, defined as the ratio of posterior return expectation and perceived riskiness. The interaction between *Perception Gap* and *Machine* and *Man + Machine* treatments is negative and statistically significant (at the 5% and the 1% levels). This indicates that AI-based recommendations lead to lower risk-adjusted return expectations, with the effect being stronger as the gap between the provided signal and the investor's prior expectation increases. This result confirms our qualitative conclusions from Columns 1 and 4.

Among the controls, we find that investors with high risk preferences have lower risk-adjusted return expectations (significant at the 1% level). This aligns with the standard portfolio theory. Moreover, we observe that women exhibit lower risk-adjusted return expectations, in line with the growing literature on the gender differences in investor beliefs (Almenberg & Dreber, 2015; Barber & Odean, 2011; Bucher-Koenen et al., 2023; Lawrence et al., 2024).

Result 1. Although investors generally update their beliefs about future returns in response to forecast information, the average investor is less responsive when analysts incorporate AI.

3.2. Exploring reasons why investors do not update their beliefs

Although extensive literature shows AI's benefits in financial forecasting, we have found that investors seem reluctant to update return beliefs *toward* AI-generated forecasts. From previous papers, we identify four potential explanations. First, investors may find the forecast providers less credible. Second, the previous literature has pointed to differences in AI adoption between men and women. Third, individuals with different political affiliations can respond differently to new information. Finally, investors may lack the knowledge to evaluate this information in an investment context.

3.2.1. The role of credibility

Low perceived credibility can explain the observed differences in learning rates. To study this, we use one of the posterior perceptions questions we asked, *Credibility*:

> *"How credible do you think the information containing the 5% forecast you previously received is?"*

Responses are recorded on a seven-point scale from 1 (Not credible) to 7 (Very credible).

[INSERT FIGURE 3]

Figure 3 illustrates the respondents' average beliefs on *Credibility* across the three groups. We highlight significant differences between treatments, with *Man* being the most credible relative to *Machine* being the least credible forecast provider. This can explain the observed differences in return updating.

We use the following specification with *Credibility* as a dependent variable to formally examine whether forecasts involving artificial intelligence are perceived as less credible:

$$Credibility_i = \beta_0 + \beta_1 Treatment_{t,i} + \beta_2 Perception\ gap_i + \gamma X_i + \varepsilon_i \tag{3}$$

Where β_1 captures the extent to which the investors perceived the two forecasts involving AI (i.e., *Machine* and *Man+Machine*) as having different credibility than those made solely by *Man*, and the control variables are similar to Equation 1. Table 3 reports the regression coefficients.

[INSERT TABLE 3]

The coefficients reveal that *Machine* and *Man + Machine* recommendations are perceived as less credible than human recommendations (significant at the 1% and 10% level, respectively). This explains why investors are less likely to update their beliefs according to the forecast provider (cfr. Table 2). Overall, the result that people perceive the AI forecast as less credible contributes to the larger literature on trust by focusing on AI-generated financial advice (e.g., Hodge et al., 2021; Rossi & Utkus, 2020) and signals a challenge to its adoption in financial decision-making.

In other words, the forecast source is the key driver of our results, arguably mediating through its credibility. To support the view that the signal's credibility is important in updating beliefs, Table A.1 reports that the perception gap does not have explanatory power in belief updating once credibility is considered. Moreover, the higher the signal credibility, the greater the belief adjustment toward the provided forecast. Therefore, belief updating is significantly driven by the credibility of the signal.

Result 2. AI forecasts are perceived as less credible compared to human analyst forecasts.

3.2.2. The role of gender

Table 2 documented gender differences in return updating across forecast providers. Indeed, a growing body of literature finds a relation between gender and financial advice. For instance, Bhattacharya et al. (2024) argued that women get worse financial advice, mainly if they exhibit high risk tolerance and higher confidence. Bucher-Koenen et al. (2023) find evidence of gender discrimination among financial advisors. Our setup directly adds to the literature by analyzing the effect of gender on financial advice when they receive the same treatment. [9]

Building on our previous results, We estimate the specifications of the following form:

$$\begin{aligned} Update_i = \beta_0 + \beta_1 Treatment_{t,i}\ x\ Perception\ gap_i\ x\ Female_i \\ + \beta_2 Treatment_{t,i}\ x\ Female_i + \beta_3 Perception\ gap_i\ x\ Female_i \\ + \beta_4 Treatment_{t,i} + \beta_5 Perception\ gap_i + \beta_6 Female_i + \gamma X_i + \varepsilon_i \end{aligned} \tag{4}$$

where $Update_i$ is the difference between posterior and prior return expectations for individual i. Our variable of interest is β_1, which captures the rate at which women update beliefs toward the provided forecast. $Female_i$ is a dummy variable that yields one if respondent *i* is a woman.

[INSERT TABLE 5]

Table 5 presents these findings. Our analysis reveals stark gender differences in how investors respond to forecast providers, particularly from *Machine*. Women, on average, are more likely to positively update their return beliefs to the information treatment. The triple interaction of *Female*, *Perception Gap*, and *AI* is positive (significant at the 5% level), highlighting that women seem to trust AI-generated forecasts. Moreover, the interaction coefficients between *Female* and the *Perception Gap* are positive (significant at the 1% level).

The results suggest that women are more responsive to the provided information, particularly from AI sources. This is interesting from two perspectives. First, this goes against the findings

[9] Table A.2. reports statistical differences across genders. Women in our sample are older, less financial and AI literate, find ChatGPT less useful, more risk averse, less confident in their return estimate while having a higher expected return, and have more trust in Goldman Sachs. These results align with the tendencies across genders (Almenberg & Dreber, 2015; Barber & Odean, 2011; Carvajal et al., 2024).

from Experiment 1, where we found that AI-generated forecasts lead to *a decrease* in updating, arguably showing a distrust in artificial intelligence. Second, Second, Brenner and Meyll (2020) showed that women prefer human advice to robo-advisors. Instead, we document that women are *more* likely to follow an AI-generated forecast when presented with them. This conclusion aligns with D'Acunto et al. (2020); women update their inflation expectations more strongly in response to information provision.

Turning to risk-adjusted return expectations, we find further evidence of gender differences. Indeed, women show *higher* risk-adjusted returns when they have a larger perception gap and receive AI forecasts (significant at the 5% level). The interaction between the *perception gap* and *women* is also positive (significant at the 1% level), confirming that women with a larger initial misperception tend to form a more optimistic (risk-adjusted) return expectation after receiving the information treatment.

Result 3. Women are more likely to update their beliefs in line with AI-generated forecasts.

This seemingly contradictory conclusion can be explained by considering different aspects of trust and technology adoption. The general trend of lower trust in AI forecasts could reflect a broader skepticism towards (new) technologies in financial decision-making (Lee & See, 2004; Longoni et al., 2019). However, the gender differences could be rooted in differing approaches to financial advice and technology adoption (e.g., Brenner & Meyll, 2020).

Prior research has shown that women are often open to seeking and following financial advice (e.g., Bucher-Koenen et al., 2023). Therefore, this could extend to novel sources of advice, such as AI. Additionally, studies have found that women may be more likely to admit uncertainty in financial matters (Barber & Odean, 2001), which could make them more receptive to external inputs, including AI forecasts.

Furthermore, the gender difference might reflect a 'catch-up' effect. If women have historically been less confident in their financial decision-making (as some literature suggests), they might view AI as an objective, non-discriminatory source of advice, free from potential gender biases (Bhattacharya et al., 2024; Bucher-Koenen et al., 2023), which they may have experienced with human advisors.

Our findings present another intriguing paradox. That is, while women report an average trust of 3.8 out of 7 for AI (which is below the scale's midpoint), they simultaneously demonstrated a tendency to update their beliefs toward the AI forecasts. This disconnect between stated trust and observed behavior highlights the complex nature of trust in financial decision-making. It underscores the importance of examining self-reported attitudes and (actual) decision-making patterns when studying investor behavior.

3.2.3. The role of political affiliation

There is a growing body of literature revealing that political affiliation shapes expectations, such as stock market expectations (e.g., Barrios & Hochberg, 2021; Cookson et al., 2020; Gerber & Huber, 2009; Leblang & Mukherjee, 2005; Meeuwis et al., 2022) or technology adoption (e.g., Blank & Shaw, 2015). This is particularly important in this framework since we conducted the experiment less than two months before the U.S. general election.[10]

[10] Table A.3. reports statistical differences across political affiliation. Democrats in our sample are older, more AI and financial literate, more confident in their (low) return expectations, more risk averse, have more trust in and find Goldman Sachs more capable. This is in line with the literature.

To examine the effect of political affiliation on investor beliefs, we include an indicator variable *Democrat* to the following regression specification:

$$\begin{aligned} Update_i = \beta_0 + \beta_1 Treatment_{t,i}\ x\ Perception\ gap_i\ x\ Democrat_i \\ + \beta_2 Treatment_{t,i}\ x\ Democrat_i + \beta_3 Perception\ gap_i\ x\ Democrat_i \\ + \beta_4 Treatment_{t,i} + \beta_5 Perception\ gap_i + \beta_6 Democrat_i + \gamma X_i + \varepsilon_i \end{aligned} \tag{5}$$

where $Update_i$ is the difference between posterior and prior return expectations for individual i. Our variable of interest is β_1, which captures the rate at which Democrats update their beliefs toward the provided forecast. $Democrat_i$ is an indicator variable that yields one if respondent i is affiliated with the Democrat party, according to Prolific.

[INSERT TABLE 6]

Table 6 reveals the results. We document that Democrats exhibit a stronger tendency to update their return beliefs in line with AI forecasts, especially when there is a gap between their prior beliefs and the provided forecast signal. Columns 1 and 3 show that for both AI-generated and – integrated forecasts, the triple interactions are positive (significant at the 1% and 5% levels). The findings on risk-adjusted returns further reinforce this conclusion. Democrats exhibit high risk-adjusted return expectations when presented with AI forecasts and when they have larger perception gaps, as shown by Columns 4 and 6. These results suggest that Democrats are more receptive to AI-generated financial advice, particularly when it diverges from initial beliefs.

Result 4. Democrats are more likely to update their beliefs in line with AI-generated forecasts.

The findings contribute to the growing literature on the intersection of political affiliation and financial decision-making (e.g., Coibion et al., 2020; Cookson et al., 2020; Leblang & Mukherjee, 2005). They show that political ideology plays a role in shaping stock market expectations and determining receptiveness to novel sources of financial information, such as AI forecasts. This is particularly intriguing given the broader context of our study, where we observed a general trend of low trust in AI forecasts. Indeed, we add another layer to the heterogeneity in investor responses to AI-generated forecasts we have observed throughout this study, namely, political affiliation.

3.2.4. The role of AI literacy

Our findings may vary depending on respondent's level of AI knowledge. To address the lack of an established measure for AI knowledge in the finance literature, we develop an AI literacy variable. This measure is based on the number of correct responses on a Pew Research survey, which assesses AI awareness in everyday activities.[11] This approach is inspired by prior studies emphasizing financial literacy's importance (see, e.g., Filippini et al., 2024; Lusardi & Mitchell, 2014; Lusardi & Tufano, 2015).

Given the increasing use of AI in forecasting, we study how this literacy affects investors' belief updating processes. Understanding AI literacy's capabilities, limitations, and applications can impact trust in AI-generated forecasts, mitigate algorithmic aversion, and enable more critical evaluation of AI outputs (e.g., Davis, 1989; Glikson & Williams Woolley, 2020; Lee & See, 2004). In this section, we explore the effects by analyzing how AI literacy moderates the relationship between forecast sources and return belief updating. In particular, we estimate the following regression:

---

[11] https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/

$$Update_i = \beta_0 + \beta_1 Treatment_{t,i}\ x\ Perception\ gap_i\ x\ AI\ literacy_i \qquad (6)$$
$$+ \beta_2 Treatment_{t,i}\ x\ AI\ literacy_i + \ \beta_3 Perception\ gap_i\ x\ AI\ literacy_i$$
$$+ \beta_4 Treatment_{t,i} + \ \beta_5 Perception\ gap_i + \beta_6 AI\ literacy_i + \gamma X_i + \varepsilon_i$$

where *AI literacy* is measured as the number of correct answers to five questions (see Appendix II), and all other variables are defined in the same way as in Equation 1.

[INSERT TABLE 6]

Table 6 presents the results. We show that investors with higher AI literacy are more likely to adjust their return beliefs toward AI-generated forecasts, and this relation is amplified when the perception gap increases. Columns 1 to 3 report positive coefficients ranging from 0.071 to 0.082 (significant at the 1% level) for the triple interaction between *Man + Machine*, perception gap, and AI literacy. This implies that AI-literate investors favor human-AI combinations over purely human ones.

Our results on risk-adjusted returns further support this conclusion, with AI literacy combined with AI-generated forecasts leading to high risk-adjusted return expectations. The coefficients of *Man* (significant at the 10% level) and *Man + Machine* (significant at the 1% level) are positive. In other words, the findings of this section indicate that AI literacy plays an important role in whether or not investors update their return beliefs in response to a forecast that involves AI as the source.

Result 5. Investors are more likely to update their beliefs toward AI-generated forecasts if they exhibit higher levels of AI literacy.

3.2.5. Horse race

Our comprehensive analysis reveals a nuanced picture of investor behavior in response to AI-generated forecasts. The findings show that three factors (e.g., gender, political affiliation, and AI literacy) shape investors' reactions but with varying degrees across different providers. One question, however, is whether the results hold once we control for the factors simultaneously.

$$\begin{aligned} Update_i = \beta_0 &+ \beta_1 Treatment_{t,i}\ x\ Perception\ gap_i\ x\ Female_i \\ &+ \beta_2 Treatment_{t,i}\ x\ Perception\ gap_i\ x\ AI\ literacy_i \\ &+ \beta_3 Treatment_{t,i}\ x\ Perception\ gap_i\ x\ Democrat_i\ + \delta Y_i + \gamma X_i + \varepsilon_i \end{aligned} \tag{7}$$

where Y_i is a vector of interaction variables, such as $Perception\ gap\ x\ Female$, and X_i are other control variables similar to Equation 1.

[INSERT TABLE 7]

Table 7 reports the horse race results. The results show that all three factors significantly shape investors' reactions but with varying degrees of influence between different forecast providers. For instance, women tend to update their beliefs in line with AI-generated (*Machine*) forecasts. Similarly, the results hold when we look at the risk-adjusted returns. This result reinforces our earlier observations about gender differences in receptiveness to AI in analyst forecasting.

Columns 1 to 3 document that AI literacy is a robust predictor of belief updating for forecasts combining human and AI input (*Man* + *Machine*). Similarly, Political affiliation also plays a key role, with Democrats showing a high propensity to update their return beliefs in line with *Man* + *Machine* forecasts, albeit with a smaller effect size than AI literacy. Furthermore, in Columns 4 to 6, the findings for expected Sharpe ratios largely mirror those for return expectations, with the additional finding that both AI literacy and Democratic affiliation are associated with more positive responses to purely AI-generated forecasts.

The horse race tells us that investor responses to AI in analyst forecasting are multifaceted and complex. While gender, AI literacy, and political affiliation all influence belief return updating, their relative importance varies depending on the forecast source. These findings highlight the need for a nuanced approach in developing and deploying AI-based financial advisory tools and the importance of AI education in fostering trust and the usage of AI in financial decision-making.[12]

3.3. Do our results replicate?

We conduct two additional experiments further to validate our results on investor trust in AI forecasts and explore the role of provider familiarity. In our first study, we replace the "generic AI" provider with ChatGPT, reflecting the growing prominence of LLMs in various industries (J. Chen et al., 2024; Jha et al., 2024; Lopez-Lira & Tang, 2023). We keep the flow of Experiment 1, however, we remove Stage 3 from the survey due to the fewer participants.

Our second robustness test addressed potential framing effects by providing a generic forecast, asking about the posterior return expectation, and then revealing the source in a second stage. This experiment aims to better understand the effect of forecast sources on investor credibility, in line with Eslami et al. (2015). Once the forecast source is revealed, we ask investors whether they want to change their return forecast or opinion on the source's credibility. This allows us to examine the impact of the provider better by making this information more salient.[13]

[12] Figure A.2. documents the differences in credibility across political affiliation and gender.
[13] Similar to Experiment 2, we remove Stage 3 from the survey due to a lower number of participants.

3.3.1. Participants

We recruited 600 participants who had not participated in Experiment 1 for both experiments on September 2nd, 2024. This corresponds to 1,200 new individuals. As in Experiment 1, we use Prolific to recruit participants with investment experience in exchange-traded funds or stocks at least 18 years of age.[14]

3.3.2. Updating

We are interested in the robustness of two results in particular. First, are investors more likely to update their return beliefs when they come from human analysts? To answer this question, we use the regression specification of Equation 1. Table 8 shows the results for Experiment 2.[15]

[INSERT TABLE 8]

The evidence in Table 8 broadly aligns with our conclusions. In Columns 1 to 4, the perception gap remains positive and statistically significant (at the 1% level), implying that investors with larger return misperceptions tend to update their beliefs more toward the signal. Furthermore, the *Perception Gap* and ChatGPT interaction is negative (significant at the 10% level). In other words, the average investor is less likely to update beliefs toward ChatGPT positively than a financial analyst, especially when their initial misperceptions are larger.

3.3.3. Credibility

---

[14] Table A.4. reports the summary statistics. We paid participants £9 per hour, next to the additional €50 for five participants. The median time to complete the survey was six minutes, as we do not go through additional manipulations. We rejected six individuals with return expectations above 1,000% and failed the attention check. We collect additional participants to reach the pre-determined sample size of 600.

[15] We cannot include Experiment 3 in this specification since the analyst source is revealed after eliciting the respondent's posterior return belief. As such, they cannot adjust their beliefs to the analyst provider.

Second, does credibility play a main role in understanding why investors change their beliefs? To answer this question, we first combine all experiments and study the changes in Credibility.

[INSERT FIGURE 4]

Figure 4 summarizes the evidence on *Credibility* across our experiments. A clear trend emerges: investors consistently find human analysts (*Man*) most credible. The difference relative to *Man + Machine* is statistically significant across the experiments. Even in an (increasingly) AI-driven financial landscape, this persistent preference suggests that the human component in financial advice remains essential. We use the regression specification of Equation 1, with *Credibility* as the dependent variable, to explore this more.

[INSERT TABLE 9]

Table 9 reports that, once the AI source is revealed, there is a significant negative relationship on *Credibility*. For instance, Columns 5 to 8 document that the coefficients range between -0.26 and -0.29 for the *Machine*-treatment (significant at the 5% level) or between -0.18 and -0.22 for *Man + Machine* (significant at the 10% level). In other words, ChatGPT exhibits a stronger effect when it is being used by itself (i.e., AI-generated) than when used in combination with humans (i.e., AI-incorporated). This aligns with recent literature on the perceived accuracy of ChatGPT. Overall, the evidence suggests a causal relationship between the forecast source and perceived investor credibility, confirming our conclusion: Investors are hesitant to trust AI forecasts.

3.4. Discussion

The evidence reveals a nuanced picture of investor behavior to analyst forecasts, highlighting the general responsiveness to new information and a reluctance toward AI-integrated analyst recommendations. This complexity in reactions underscores the interplay between traditional financial advice and the emerging AI technology in influencing financial decision-making and return belief-updating processes.

Our study also revealed notable gender differences in the perception and trust of AI-generated financial advice. These differences have vital implications for the design of financial education programs and the presentation of financial advice. Indeed, we argue that tailoring information provision strategies by gender could potentially enhance the effectiveness of efforts to improve financial literacy and decision-making.

The stronger response of women to the AI-generated forecasts is a novel finding that warrants further investigation. It may reflect differences in trust and perceived credibility of AI systems across genders or imply that women are more open to incorporating AI-generated information into their financial decision-making processes. This, however, goes against most of the current research, highlighting the increase in gender inequality due to AI (e.g., Brenner & Meyll, 2020; Carvajal et al., 2024; Humlum & Vestergaard, 2024).

Our results on AI literacy underscore the importance of (AI) education in fostering trust in AI-generated financial advice. As more findings emerge in favor of AI's capabilities in forecasting, financial institutions can benefit greatly from implementing strategies to enhance their clients' AI literacy. This, in turn, can lead to more informed decision-making and a greater willingness to incorporate AI-generated insights into investment strategies.

4. Updating across forecast content

In Stage 3 of Experiment 1, we randomly assign the respondents to one of three manipulations: information richness, analyst dispersion, or analyst downgrades. In other words, respondents remain in the same group (*Man*, *Man + Machine* or *Machine*) from Stage 1. This part then studies whether investors revise their return estimates after they receive one of the manipulations. As in Weber et al. (2024), we ask:

"Considering this new information, relative to your previous estimate for the return of the S&P 500 over the next year, will your estimate change?"

- *Improve substantially*
- *Improve slightly*
- *Remain the same*
- *Worsen slightly*
- *Worsen substantially*
- *Don't know*

The responses are coded on a scale from 1 (Worsen substantially) to 5 (Improve substantially). We label this as *Return revision* and use it as a dependent variable in the following specification:

$$\begin{aligned} Return\ Revision_i &= \beta_0 + \beta_1 Treatment_{t,i} + \beta_2 Manipulation_{t,i} \\ &+ \beta_3 Treatment_{t,i}\ x\ Manipulation_{t,i} + \beta_4 Perception\ gap_i + \gamma X_i + \varepsilon_i \end{aligned} \quad (8)$$

where $Manipulation_{t,i}$ is referred to as *Manipulation* where t = 1, 2, or 3, as we will define below.

4.1. Manipulation 1: information richness

In general, there are two ways in which an analyst could increase forecast credibility. The most standard method is including earnings estimates (Bradshaw, 2004; Kecskés et al., 2017; Keung & Keung, 2010). Additionally, they could include visuals, which are shown to impact financial decision-making (Bazley et al., 2021).

We build on Experiment 1 to test the effects of earning estimates and visuals. More specifically, after participants make an investment decision and provide beliefs, 600 of them are randomly assigned to one of three groups:

- **Control group**: *As a reminder, the 12-month return forecast for the S&P 500 index is 5%.*
- ***Earnings estimate***: *As a reminder, the 12-month return forecast for the S&P 500 index is 5%. The report also noted that firms' earnings are expected to rise by 5% over the next 12 months, and the equity market valuation will be 30x, close to the current P/E level.*
- ***Earnings estimate and visual***: *As a reminder, the 12-month return forecast for the S&P 500 index is 5%. The report also noted that firms' earnings are expected to rise by 5% over the next 12 months, and the equity market valuation will be 30x, close to the current P/E level. The following visual was added.*

[INSERT FIGURE 5]

Figure 5 plots the graph that we add for the third group. The data comes from S&P Global. We lagged the time series by two quarters for visual purposes and to avoid deceiving respondents. Moreover, we choose the equity market valuation to accurately reflect the price/earnings (P/E) level of August 2024.

[INSERT TABLE 10]

Columns 1 and 2 of Table 10 reveal two results. First, we find that including earnings estimates significantly increases the likelihood of the investors revising their return expectations upward (significant at the 1% level). This effect is amplified when visuals are added (significant at the 1% level). Overall, the results align with prior literature, validating the design and conclusions of our experiment.

Second, we observe no significant differences across different forecast providers for these two manipulations. This suggests that once investors engage with the forecast content, they do not appear to adjust their responses more based on the forecast provider. While the positive effects of an earnings estimate and visuals persist, they lose their statistical significance. This indicates that forecast providers do not amplify the impact of forecast content.

4.2. Manipulation 2: analyst dispersion

There is extensive literature on the impact of analyst dispersion, both from an accounting (e.g., Gloria et al., 2012; Kadous & Thayer, 2009) and a financial perspective (e.g., Barron et al., 2009; Johnson, 2004; Palley et al., 2024). The takeaway from this literature is that analyst dispersion corresponds to low future returns. Given the observed lack of trust in AI, we hypothesize that analyst dispersion negatively impacts investment beliefs for AI-generated forecasts.

Similar to Manipulation 1, 600 individuals are randomly assigned to one of three groups:

- ***Control group*****:** *After polling 14 leading investment banks, the average analyst forecast for the S&P500 index equals 5%.*
- ***Low dispersion***: *After polling 14 leading investment banks, the average analyst forecast for the S&P500 index equals 5% and ranges from -3.1% to 8.7%.*
- ***High dispersion***: *After polling 14 leading investment banks, the average analyst forecast for the S&P500 index equals 5% and ranges from -6.1% to 17.4%.*

The range of the High Dispersion condition is factual as it represents the analyst forecast range for 14 investment banks (at the end of 2023). The average forecast, however, is chosen to match the initial forecast. The Low Dispersion condition cuts the forecast range by half while keeping the average analyst forecast at 5%. This last adjustment ensures that the change in our variables of interest is driven by the analyst dispersion rather than the place of Goldman Sachs Research within this dispersion. Finally, the No Dispersion condition serves as our control group.

After receiving the second statement, we ask all participants whether they want to change their investment decision, how confident they are in it, and how credible and informative it is. This enables us to study the impact of uncertainty on an investor's asset allocation in light of analyst dispersion.

[INSERT TABLE 10]

Columns 3 and 4 of Table 10 report the coefficients. The results reveal that lower dispersion is associated with a higher likelihood of investors lowering their return revisions. This finding is statistically significant at the 10% level in our more comprehensive model (Column 4). These outcomes align with previous literature on analyst dispersion, again lending credibility to our experimental design and conclusions.

Again, we observe no significant differences in the effects between different forecast providers (*Man*, *Machine*, or *Man + Machine*). This indicates that investors' responses are primarily driven by the content of the forecast rather than its source. In other words, the source of the forecast (i.e., *Man*, *Machine*, or *Man + Machine*) does not magnify the impact of the forecast's content on investor behavior.

4.3. Manipulation 3: forecast revision

There is an extensive literature that documents the impact of analyst revisions on stock prices (Hsu & Wang, 2021; Kecskés et al., 2017; Zhang, 2006). The main finding from this literature is that prices tend to move *toward* the revised forecasts. A recent stream of the literature focuses on revisions either without revising earnings forecasts (Berger et al., 2019) or without earnings estimates (Kecskés et al., 2017). Moreover, recent AI papers highlight a decrease in trust when errors are involved (Dietvorst et al., 2014). In our setting, we consider "an analyst revision" to be such an error.

We follow the same structure as Manipulations 1 and 2 and randomly assign participants to:

- ***Control group***: *Suppose that, after six months, the return forecast for the S&P 500 index has not been revised.*
- ***Downgrades with no earnings estimate***: *Suppose that, after six months, the return forecast for the S&P 500 index is revised from 5% to 3%.*
- ***Downgrades with an earnings estimate***: *Suppose that, after six months, the return forecast for the S&P 500 index is revised from 5% to 3%. The equity market valuation will be 30x, close to the current P/E level.*

The magnitude of the revision accurately reflects Goldman Sachs' forecast error of the S&P 500 index over the last 20 years. This adds realism to the setting and avoids deceiving participants.

[INSERT TABLE 10]

Columns 5 and 6 of Table 10 present the results. Investors are significantly more likely to revise their return expectations downward when presented with an analyst downgrade. This effect is robust across specifications and statistically significant at the 1% level. It really underscores the substantial influence that a revision wields over investor perceptions. It is noteworthy that our results reveal that the magnitude of this effect is more pronounced when downgrades are presented *without* accompanying earnings estimates. This contrasts prevailing wisdom, which typically posits that detailed information should elicit a stronger investor reaction (Kecskés et al., 2017).

Notably, the analysis reinforces our earlier conclusions: the absence of significant variation in investor responses across different forecast providers. This finding has important implications for understanding information processing in financial markets. That is, there is a dual impact of forecast source and content. Furthermore, the financial and accounting industry's efforts to integrate AI into forecasting may need to focus more on improving its credibility.

Result 6. The forecast provider does not amplify investor reactions from forecast content.

4.4. Does AI model disclosure matter?

There is large interest among academics and practitioners in finding the best machine-learning model, as these models are used for a number of applications, such as financial forecasting and economic predictions (e.g., Aubry et al., 2023; Cao et al., 2024; Luyang Chen et al., 2024; Gu et al., 2020). The main finding of this literature is that OLS models yield the worst (out-of-sample) performance, while deep learning models do significantly better (Luyang Chen et al., 2024).

This leads to the question of whether investors find information related to the chosen AI model useful. In other words, does AI model disclosure matter? We recruit 600 individuals (who have not yet participated in the previous experiments) to examine this question on September 11th, 2024.[16] We follow the same procedure as Experiment 1, in which we use Prolific and only focus on U.S. participants with investment experience and at least 18 years old.

Following the flow of Experiment 1, they are randomly assigned to one of four groups:

[16] Table A.5. reports the summary statistics. We paid participants £9 per hour, next to the additional €50 for five participants. The median time to complete the survey was six minutes. Overall, we rejected one participant with a return expectation above 1,000% and collected one participant to reach sample size.

- ***Control***: *An AI model, trained on financial data and market trends, forecasts a 12-month return of 5% for the S&P 500 index. The baseline assumption during the next year is that the US economy will continue to expand at a modest pace and avoid a recession.*
- ***OLS*****:** *An AI model, trained on financial data and market trends, forecasts a 12-month return of 5% for the S&P 500 index. The baseline assumption during the next year is that the US economy will continue to expand at a modest pace and avoid a recession. This AI system applies an Ordinary Least Squares regression model.*
- ***BLUE***: *An AI model, trained on financial data and market trends, forecasts a 12-month return of 5% for the S&P 500 index. The baseline assumption during the next year is that the US economy will continue to expand at a modest pace and avoid a recession. This AI system applies a method that satisfies the Best Linear Unbiased Estimator properties.*
- ***Deep learning:*** *An AI model, trained on financial data and market trends, forecasts a 12-month return of 5% for the S&P 500 index. The baseline assumption during the next year is that the US economy will continue to expand at a modest pace and avoid a recession. This AI system applies a deep neural network learning model.*

Similar to Experiment 1, we keep the information in the text constant except for the method to isolate its effect. We made specific choices for the treatments: OLS is arguably the most familiar method with the worst out-of-sample performance in a financial setting (Gu et al., 2020). It also sounds more straightforward relative to the other two. In turn, the deep learning method has the highest out-of-sample performance while being rather technical (Luyang Chen et al., 2024) and also sounding more complex. We also include the Best Linear Unbiased Estimator (BLUE) treatment, which shares properties with OLS while also sounding sophisticated. Prior research has shown that complexity can be off-putting to investors (Oberholzer et al., 2024; Umar, 2022).

We use the regression specification in Equation 1 to examine the impact of AI model disclosure on the return update learning rates.

[INSERT TABLE 11]

Table 11 reports the results. Columns 1 to 4 find that the interaction between the *OLS-treatment* and *Perception Gap* is positive (significant at the 1% level). Additionally, the *Perception Gap* itself is positive (significant at the 1% level). These results indicate that investors are more likely to use an OLS method to update their beliefs in the forecast's direction. Notably, we do not find significant effects for the BLUE or deep learning treatments or interactions with the perception gap. This implies that, from a return updating viewpoint, investors trust OLS the most, as they are more likely to adjust their expectations in line with this signal.

Columns 5 to 8 show that investors are not more likely to adjust their return beliefs according to OLS methods on a risk-adjusted basis. Nevertheless, we highlight that using more complex-sounding methods leads to a lower risk-adjusted return. For instance, the interaction between the *BLUE*-treatment and *Perception gap* is negative (significant at the 10% level). The interaction with deep learning, while being statistically insignificant, is also negative. This aligns with the evidence in Columns 1 to 4, highlighting that individuals are not more likely to follow complex methods.

There are two potential explanations for these results. First, it can imply that people only react to what they are familiar with, as Davis (1989) or Lee and See (2004) suggested. Our data does, unfortunately, not allow us to test this hypothesis formally. Second, if BLUE and deep learning treatments are perceived as more complex, individuals are *more* likely to shy away from them (e.g., Oberholzer et al., 2024). For instance, Umar (2022) highlights that complexity negatively affects investors. In our framework, this is shown by the positive interaction of return updating for the simple (arguably more familiar) model and simultaneously the negative interactions of risk-adjusted returns for the "complex-sounding" methods. In other words, there is suggestive evidence of complexity aversion (Oberholzer et al., 2024; Umar, 2022) and familiarity bias (e.g., Davis, 1989; Lee & See, 2004). We invite future research to investigate the mechanisms further.

Result 7. Model complexity negatively affects the likelihood of return updating.

Overall, the findings highlight the potential dangers of AI forecasting: Adding a more familiar statistical model to an analyst forecast increases its effect on return beliefs. Despite the worst-performing method (Gu et al., 2020), the significant coefficient of the OLS treatment highlights the potential misalignment between the perceived and actual reliability of forecasting models.

5. Conclusion

Our study provides novel insights into the complex relationship between artificial intelligence, financial analyst forecasts, and investor trust. Through four incentivized experiments, we find that investors generally update their return beliefs toward (new) analyst forecast information. However, compared to expectations, they are less responsive when analysts incorporate AI in their forecasts. This highlights that the average investor does not entirely trust AI-incorporated financial forecasts.

We document five potential explanations for this result. First, we argue that there are sizeable differences across genders. On average, women are more likely to follow AI-generated advice. Second, we highlight that AI literacy plays a key role. Investors are more likely to update their beliefs in line with new information if they have higher AI literacy levels. Third, we document that Democrats are *more* likely to update their return beliefs in line with the AI forecast. Fourth, we report a strong relation between perceived provider credibility and learning rates. Indeed, investors are more likely to update their beliefs if they perceive the forecast as more credible. Finally, we document complexity's role in AI forecasts, with OLS being trusted more than deep learning models. This finding hints at familiarity bias and complexity aversion.

Our manipulations of information richness, analyst dispersion, and revisions demonstrate that investors are more likely to respond to the informational content of the forecast rather than its source. This indicates that while the source of the forecasts (*Man* vs. *Machine*) influences initial credibility perceptions, its content also shapes investor beliefs. However, we highlight that the forecast sources do not amplify the reaction to variation in forecast content. Overall, the results challenge prevailing notions about AI integration in financial decision-making.

References

Acemoglu, D. (2021). *Harms of AI*.

Acemoglu, D. (2024). The Simple Macroeconomics of AI. *Economic Policy*.

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2010). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, *40*.

Acemoglu, D., & Estrepo, P. (2022). Tasks, Automation, and the Rise in the U.S. Wage Inequality. *Econometrica*, *90*(5), 1973–2016.

Almenberg, J., & Dreber, A. (2015). Gender, stock market participation and financial literacy. *Economics Letters*, *137*, 140–142.

Asparouhova, E., Bossaerts, P., Cai, X., Rotaru, K., Yadav, N., & Yang, W. (2024). Humans in charge of trading robots: the first experiment. *Review of Finance*, *28*, 1215–1244.

Aubry, M., Kräussl, R., Manso, G., & Spaenjers, C. (2023). Biased Auctioneers. *The Journal of Finance*, *78*(2), 795–833.

Barber, B. M., & Odean, T. (2011). Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment. *Quarterly Journal of Economics*, *116*(1), 261–292.

Barrios, J. M., & Hochberg, Y. V. (2021). Risk perceptions and politics: Evidence from the COVID-19. *Journal of Financial Economics*, *142*(2), 862–879.

Barron, O. E., Stanford, M. H., & Yu, Y. (2009). Further Evidence on the Relation between Analysts' Forecast Dispersion and Stock Returns. *Contemporary Accounting Research*, *26*(2), 329–357.

Bauer, R., Gödker, K., Smeets, P., & Zimmermann, F. (2024). *Mental Models in Financial Markets: How Do Experts Reason about the Pricing of Climate Risk?* (Issue 17030).

Bazley, W. J., Cronqvist, H., & Mormann, M. (2021). Visual Finance: The Pervasive Effects of Red on Investor Behavior. *Management Science*, *67*(9), 5616–5641.

Berger, P. G., Ham, C. G., & Kaplan, Z. R. (2019). Do Analysts Say Anything About Earnings Without Revising Their Earnings Forecasts? *The Accounting Review*, *94*(2).

Beutel, J., & Weber, M. (2024). *Beliefs and Portfolios: Causal Evidence*.

Bhagwat, V., & Liu, X. (2020). The Role of Trust in Information Processing : Evidence from Security Analysts. *The Accounting Review*, *95*(3), 59–83.

Bhattacharya, U., Kumar, A., Visaria, S., & Zhao, J. (2024). Do Women Receive Worse Financial Advice? *The Journal of Finance*, *0*, 1–48.

Blank, J. M., & Shaw, D. (2015). Does Partisanship Shape Attitudes toward Science and Public Policy? The Case for Ideology and Religion. *The Annals of the American Academy of Political and Social Science*, *658*(March), 18–35.

Borsboom, C., & Zeisberger, S. (2020). What makes an investment risky? An analysis of price path characteristics. *Journal of Economic Behavior and Organization*, *169*, 92–125.

Bradshaw, M. T. (2004). How Do Analysts Use Their Earnings Forecasts in Generating Stock Recommendations? *The Accounting Review*, *79*(1), 25–50.

Brenner, L., & Meyll, T. (2020). Robo-advisors: A substitute for human financial advice? *Journal of Behavioral and Experimental Finance*, *25*, 100275.

Bucher-Koenen, T., Hackethal, A., Koenen, J., & Laudenbach, C. (2023). *Gender Differences in Financial Advice*.

Cao, S., Jiang, W., Wang, J., & Yang, B. (2024). From Man vs. Machine to Man + Machine: The art and AI of stock analyses. *Journal of Financial Economics*, *160*(103910).

Carvajal, D., Franco, C., & Isaksson, S. (2024). *Will Artificial Intelligence Get in the Way of Achieving Gender Equality?*

Chak, I., Croxson, K., D'Acunto, F., Reuter, J., Rossi, A. G., & Shaw, J. M. (2022). *Improving Household Debt Management with Robo-Advice*.

Chen, J., Tang, G., Zhou, G., & Zhu, W. (2024). *ChatGPT, Stock Market Predictability and Links to the Macroeconomy*.

Chen, Lingjiao, Zaharia, M., & Zou, J. (2024). *How Is ChatGPT's Behavior Changing over Time ?*

Chen, Luyang, Pelger, M., & Zhu, J. (2024). Deep Learning in Asset Pricing. *Management Science*, *70*(2), 714–750.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2020). *Political Polarization and Expected Economic Outcomes* (No. 28044; NBER Working Paper Series).

Coleman, B., Merkley, K., & Pacelli, J. (2022). Human Versus Machine: A Comparison of Robo-Analyst and Traditional Research Analyst Investment Recommendations. *The Accounting Review*, *97*(5), 221–244.

Cookson, J. A., Engelberg, J. E., & Mullins, W. (2020). Does Partisanship Shape Investor Beliefs? Evidence from the COVID-19 Pandemic. *Review of Asset Pricing Studies*, *10*(September), 863–893.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Technology Information. *MIS Quarterly*, *13*(3), 319–340.

De Silva, T., & Thesmar, D. (2023). Noise in Expectations: Evidence from Analyst Forecast. *Review of Financial Studies*, *37*(5), 1494–1537.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2014). Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err. *Journal of Experimental Psychology: General*, *143*(6), 1–13.

Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., & Sandvig, C. (2015). "I always assumed that I wasn't really that close to [her]": Reasoning about Invisible Algorithms in News Feeds. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, *April*, 153–162.

Estep, C., Griffith, E. E., & Mackenzie, N. L. (2024). How do financial executives respond to the use of artificial intelligence in financial reporting and auditing ? *Review of Accounting Studies*, *29*(3), 2798–2831.

Fedyk, A., Khimich, N., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies*, *27*, 938–985.

Filippini, M., Leippold, M., & Wekhof, T. (2024). Sustainable finance literacy and the determinants of sustainable investing. *Journal of Banking and Finance*, *163*(March), 107167.

Gerber, A. S., & Huber, G. A. (2009). Partisanship and Economic Behavior: Do Partisan Differences in Economic Forecasts Predict Real Economic Behavior? *American Political Science Review*, *103*(3).

Germann, M., & Merkle, C. (2023). Algorithm aversion in delegated investing. In *Journal of Business Economics* (Vol. 93, Issue 9). Springer Berlin Heidelberg.

Glikson, E., & Williams Woolley, A. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management*, *14*(2), 627–660.

Gloria, X., Natarajan, R., & Liu, X. G. (2012). The Effect of Financial Analysts' Strategic Behavior on Analysts' Forecast Dispersion. *The Accounting Review*, *87*(6), 2123–2149.

Gu, S., Kelly, B. T., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *Review of Financial Studies*, *33*(5), 2223–2273.

Guan, Y., & Tsang, A. (2020). Societal Trust and Management Earnings Forecasts. *The Accounting Review*, *95*(4), 149–184.

Haaland, I., Roth, C., & Wohlfart, J. (2023). *Designing Information Provision Experiments*. *61*, 3–40.

Hodge, F. D., Mendoza, K. I., & Sinha, R. K. (2021). The Effect of Humanizing Robo-Advisors on Investor Judgments. *Contemporary Accounting Research*, *E1*(E8), 770–792.

Holzmeister, F., Huber, J., Kirchler, M., Lindner, F., Weitzel, U., & Zeisberger, S. (2024). laypeople What Drives Risk Perception? A Global Survey with Financial Professionals and Laypeople. *Management Science*, *66*(9), 3977–4002.

Hsu, C., & Wang, R. (2021). Bundled Earnings Guidance and Analysts' Forecast Revisions. *Contemporary Accounting Research*, *38*(4), 3146–3181.

Humlum, A., & Vestergaard, E. (2024). *The Adoption of ChatGPT* (Issue 16992).

Jha, M., Qian, J., Weber, M., & Yang, B. (2024). *ChatGPT and Corporate Policies* (No. 23–15; Chicago Booth Paper).

Johnson, T. C. (2004). Forecast dispersion and the cross-section of expected returns. *The Journal of Finance*, *59*(5), 1957–1978.

Kadous, K., & Thayer, J. (2009). Is There Safety in Numbers? The Effects of Forecast Accuracy and Forecast Boldness on Financial Analysts' Credibility with Investors. *Contemporary Accounting Research*, *26*(3), 933–968.

Kecskés, A., Michaely, R., & Womack, L. (2017). Do Earnings Estimates Add Value to Sell-Side

Analysts' Investment Recommendations? *Management Science*, *63*(6), 1855–1871.

Keung, E. C., & Keung, E. C. (2010). Do Supplementary Sales Forecasts Increase the Credibility of Financial Analysts? *The Accounting Review*, *85*(6), 2047–2074.

Kim, A. G., Muhn, M., & Nikolaev, V. V. (2024). *Financial Statement Analysis with Large Language Models Large Language Models* (No. 2024–65; Becker Friedman Institute For Economics).

Kostin, D. J., Snider, B., Hammond, R., Conners, C., Calcagnini, L., Ma, J., & Chavez, D. (2023). *2024 US Equity Outlook: "All You Had To Do Was Stay"* (Issue November 2023).

Laudenbach, C. (2024). Beliefs About the Stock Market and Investment Choices: Evidence from a Survey and a Field Experiment. *Review of Financial Studies*.

Lawrence, E. R., Nguyen, T. D., & Wick, B. (2024). Gender difference in overconfidence and household financial literacy. *Journal of Banking and Finance*, *166*(May), 107237.

Leblang, D., & Mukherjee, B. (2005). Government partisanship, elections, and the stock market: Examining American and British stock returns, 1930-2000. *American Journal of Political Science*, *49*(4), 780–802.

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, *46*(1), 50–80.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650.

Lopez-Lira, A., & Tang, Y. (2023). *Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models*.

Lusardi, A., & Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature*, *52*(1), 5–44.

Lusardi, A., & Tufano, P. (2015). Debt litarecy, financial experiences, and overindebtedness. *Journal of Pension Economics & Finance*, *14*(4), 332–368.

Mcelheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. (2024). AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy*, *June 2022*, 375–415.

Meeuwis, M., Parker, J. A., Schoar, A., & Simester, D. (2022). Belief Disagreement and Portfolio Choice. *The Journal of Finance*, *LXXVII*(6), 3191–3247.

Merkoulova, Y., & Veld, C. (2022). Stock return ignorance. *Journal of Financial Economics*, *144*, 864–884.

Mollick, E., & Mollick, L. (2022). *New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments*.

Oberholzer, Y., Olschewski, S., & Scheibehenne, B. (2024). Complexity aversion in risky choices and valuations: Moderators and possible causes. *Journal of Economic Psychology*, *100*(November 2023), 102681.

Okat, D., Paaso, M., & Pursiainen, V. (2024). Trust in Finance and Consumer Fintech Adoption. *The Review of Corporate Finance Studies*, *November*.

Önkal, D., Goodwin, P., Thomson, M., Gönul, S., & Pollock, A. (2009). The Relative Influence of Advice From Human Experts and Statistical Methods on Forecast Adjustments. *Journal of Behavioral Decision Making*, *409*(February), 390–409. https://doi.org/10.1002/bdm

Palley, A. B., Steffen, T. D., & Zhang, F. X. (2024). The Effect of Dispersion on the Informativeness of Consensus Analyst Target Prices. *Management Science*, *Forthcomin*.

Pevzner, M., Xie, F., & Xin, X. (2015). When firms talk, do investors listen? The role of trust in stock market reactions to corporate earnings announcements. *Journal of Financial Economics*, *117*(1), 190–223.

Rossi, A. G., & Utkus, S. (2020). *The Needs and Wants in Financial Advice: Human versus Robo-advising*.

Umar, T. (2022). Complexity aversion when Seeking Alpha. *Journal of Accounting and Economics*, *73*(2–3), 101477.

Ungeheuer, M., & Weber, M. (2021). The Perception of Dependence, Investment Decisions, and Stock Prices. *The Journal of Finance*, *76*(2), 797–844.

van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), 449–472.

Weber, M., Candia, B., Coibion, O., & Gorodnichenko, Y. (2024). *Do You Even Crypto, Bro? Cryptocurrency in Household Finance*.

Wei, C., & Zhang, L. (2023). Trust in Financial Markets : Evidence from Reactions to Earnings News. *Management Science*, *69*(10), 6393–6415.

Yang, T., & Zhang, X. (2022). FinTech adoption and financial inclusion: Evidence from household consumption in China. *Journal of Banking and Finance*, *145*, 106668.

Zhang, F. X. (2006). Information Uncertainty and Analyst Forecast Behavior. *Contemporary Accounting Research*, *23*(2).

Figures

Figure 1. Experiment flow

This figure illustrates the flow of the experimental survey.

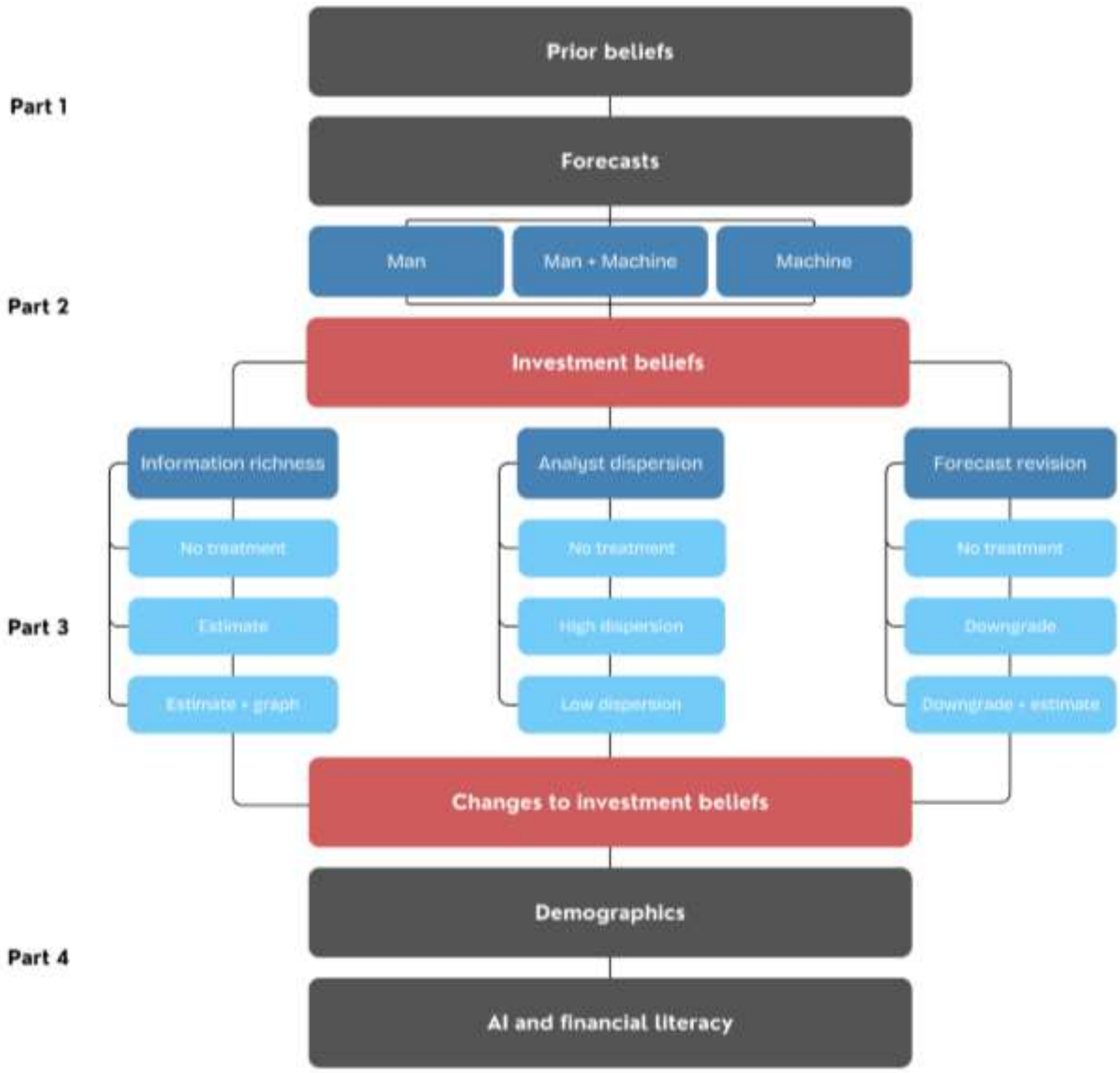

Figure 2. Incentivized question

This figure plots the screen participants will see to ask about their return expectations priors.

What do you expect the return of the S&P 500 to be over the next 12 months?

Note: This expected return is the change in value, in percentage, that you expect to receive over the next year from investing in an index fund replicating the S&P 500. It includes both dividends and capital gains/losses.

(Please answer only with a positive or negative numeric value with at most 1 decimal.)

% per year over the next year

Bonus Condition: If our computer randomly selects this question, the five respondents who provide the most accurate answers (i.e., those with expected returns closest to the realized ones) will receive a bonus payment of 50 euros one year from now.

Figure 3. Average credibility

This figure plots the average credibility of the analyst forecast across treatments in Experiment 1. The bars represent the mean values by treatment. The error bars indicate 95% confidence intervals. The red reference line is the average belief across all treatments.

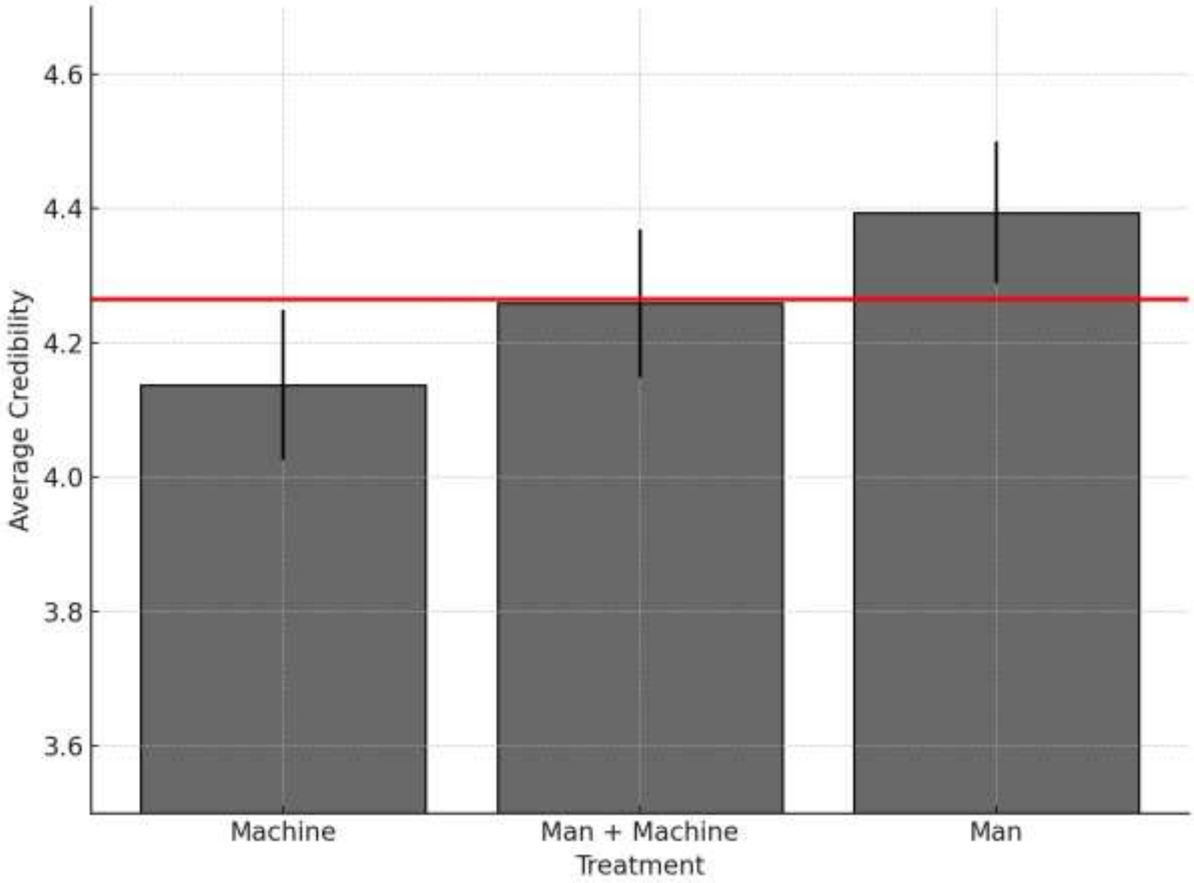

Figure 4. Summary of the results: Credibility

This figure plots the average credibility of the analyst forecast across treatments in the three experiments.

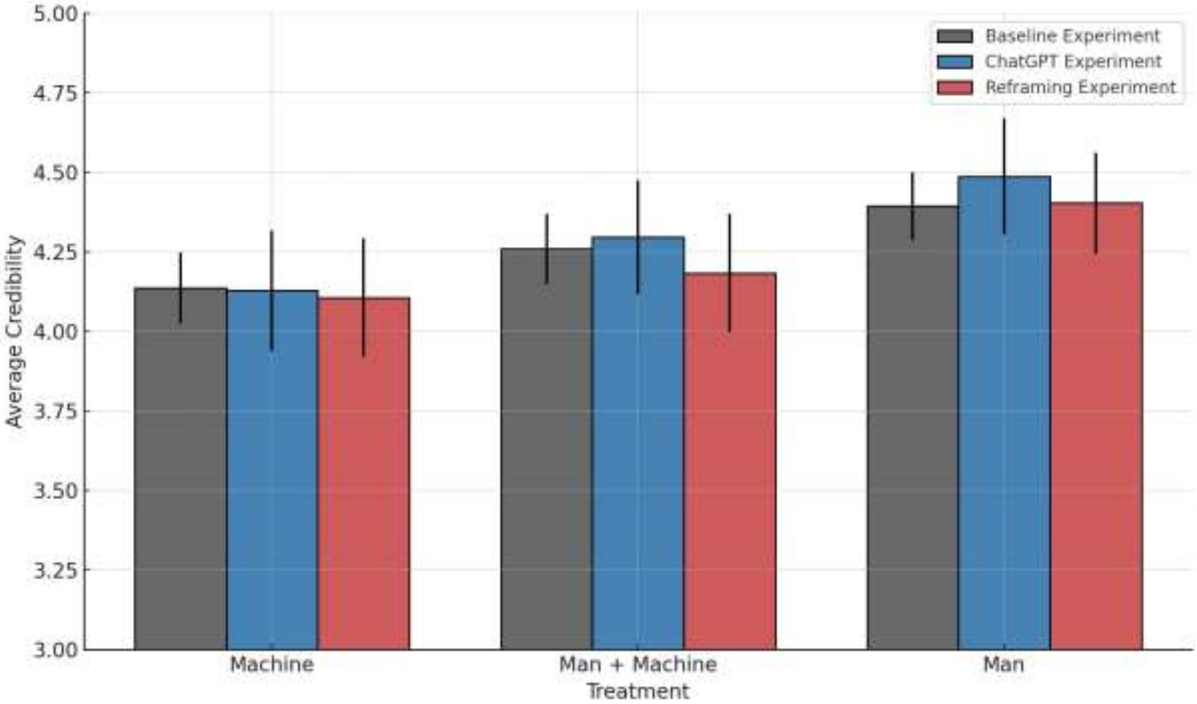

Figure 5. S&P 500 earnings per share

This figure plots the (lagged) S&P500 earnings per share from January 2012 to June 2024 (blue) and the 5% forecast (red) for the next four quarters.

Tables

Table 1. Summary statistics

This table reports summary statistics. Numbers in parentheses imply the range of possible values. AI and Financial literacy are the number of correctly answered questions in the tests (out of 5). Columns 6-10 check the balance of means across the control and two treatment groups, and *p*-values between control and treatment groups (Columns 9 and 10). Belief variables refer to prior elicited before the information manipulation. The return expectation prior is incentivized. We drop participants who exhibit return expectations below -30% and above 30%.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
	Mean	SD	p10	p50	p90	Man Mean	Man + Machine Mean	Machine Mean	p-value (7 = 6)	p-value (8 = 6)
Age	42.413	12.643	21	41	70	41.053	41.926	41.381	0.251	0.669
Female	0.466	0.499	0	0	1	0.465	0.426	0.441	0.196	0.436
University	0.701	0.458	0	1	1	0.687	0.707	0.711	0.477	0.397
AI literacy (1-5)	4.058	1.086	1	4	5	4.078	4.031	4.103	0.462	0.697
Confidence (1-7)	4.920	1.396	1	5	7	4.961	4.826	4.894	0.109	0.433
Financial literacy (1-5)	4.222	0.895	1	4	5	4.239	4.248	4.207	0.848	0.548
Return expectation	8.287	7.546	-15	8	17	8.469	8.338	8.610	0.768	0.741
Risk preference (1-7)	4.065	1.424	1	4	6	4.046	4.059	4.141	0.874	0.266
AI relative capabilities (1-7)	4.085	1.314	1	4	6	4.073	4.022	4.079	0.520	0.938
AI trust (1-7)	3.808	1.413	1	4	6	3.819	3.741	3.751	0.346	0.408
ChatGPT usefulness (1-7)	4.356	1.832	3	5	7	4.459	4.276	4.329	0.098	0.234
Goldman Sachs capability (1-7)	4.582	1.199	3	5	6	4.582	4.611	4.551	0.689	0.667
Goldman Sachs trust (1-7)	4.431	1.276	3	5	6	4.406	4.433	4.429	0.724	0.525

Table 2. Baseline result

This table shows the coefficients of *Update* (the difference between posterior and prior expectation for the 12-month-ahead S&P500 index return) and *Expected Sharpe Ratio* (the ratio of posterior return expectation and risk perception) on a set of independent variables. The perception gap (*Perc. gap*) is the difference between the analyst forecast and the prior expectation about the 12-month-ahead S&P500 returns. The treatments are *Man + Machine* (I_C) and *Machine* (I_{AI}). The control variables are defined in Table 1. Demographics fixed effects include education and income levels. P-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	**Update**				**Expected Sharpe ratio**			
	1.	2.	3	4.	5.	6.	7.	8.
$Perc. gap\ x\ I_{AI}$	-0.027	-0.027	-0.026	-0.027	-0.048***	-0.050**	-0.049***	-0.050***
	(0.335)	(0.338)	(0.344)	(0.341)	(0.002)	(0.002)	(0.002)	(0.002)
$Perc. gap\ x\ I_C$	-0.056**	-0.051*	-0.048*	-0.050*	-0.033**	-0.033**	-0.031**	-0.032**
	(0.037)	(0.055)	(0.070)	(0.064)	(0.033)	(0.031)	(0.042)	(0.030)
$Perc. gap$	0.524***	0.517***	0.516***	0.514***	-0.144***	-0.147***	-0.148***	-0.149***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
I_{AI}	0.166	0.167	0.169	0.174	-0.206	-0.199	-0.204	-0.196
	(0.456)	(0.448)	(0.446)	(0.433)	(0.105)	(0.155)	(0.105)	(0.122)
I_C	0.155	0.173	0.186	0.177	0.129	0.133	0.132	0.112
	(0.476)	(0.425)	(0.393)	(0.418)	(0.300)	(0.285)	(0.288)	(0.369)
Risk preference		0.095*	0.092	0.089		-0.085***	-0.093***	-0.089***
		(0.100)	(0.120)	(0.136)		(0.010)	(0.006)	(0.008)
AI literacy		0.130*	0.131*	0.119		0.091**	0.085*	0.084*
		(0.096)	(0.094)	(0.130)		(0.042)	(0.056)	(0.061)
Financial literacy		0.425***	0.434***	0.447***		0.197***	0.202***	0.208***
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Female			-0.068	-0.048			-0.179*	-0.172*
			(0.691)	(0.780)			(0.067)	(0.079)
Age			-0.006	-0.005			-0.009**	-0.008**
			(0.363)	(0.439)			(0.020)	(0.033)
AI trust			0.053	0.073			-0.036	-0.027
			(0.549)	(0.414)			(0.475)	(0.598)
AI capability			-0.098	-0.118			-0.030	-0.041
			(0.296)	(0.211)			(0.573)	(0.447)
Constant	-0.340**	-3.077***	-2.638***	-3.299***	1.488***	0.617**	1.355***	1.231**
	(0.028)	(0.000)	(0.000)	(0.000)	(0.000)	(0.045)	(0.000)	(0.020)
R-squared	0.542	0.550	0.550	0.550	0.303	0.313	0.317	0.319
Demographics FE	No	No	No	Yes	No	No	No	Yes
Observations	1,668	1,668	1,664	1,664	1,668	1,668	1,664	1,664

Table 3. Credibility

This table shows the coefficients of *Credibility*, measured as posterior perceptions, on a set of independent variables. The perception gap (*Perc. gap*) is the difference between the analyst forecast and the prior expectation about the 12-month-ahead S&P500 index returns. Treatments are *Man + Machine* (I_C) and *Machine* (I_{AI}). All other control variables are defined in Table 1. Demographics fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	**Credibility**				
	1.	2.	3	4.	5.
I_{AI}	-0.257***	-0.256***	-0.261***	-0.256***	-0.253***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
I_C	-0.135*	-0.136*	-0.140*	-0.123	-0.113
	(0.085)	(0.083)	(0.072)	(0.106)	(0.138)
$Perc. gap$		0.006	0.007	0.009**	0.009**
		(0.182)	(0.114)	(0.031)	(0.037)
Risk preference			0.075***	0.043*	0.048**
			(0.001)	(0.059)	(0.037)
AI literacy			-0.067**	-0.066**	-0.064**
			(0.030)	(0.028)	(0.035)
Financial literacy			0.013	-0.005	0.005
			(0.733)	(0.902)	(0.901)
Female			0.004	-0.007	0.004
			(0.957)	(0.917)	(0.957)
Age			0.002	0.001	0.002
			(0.482)	(0.621)	(0.482)
AI trust			0.016	0.015	0.016
			(0.644)	(0.661)	(0.644)
AI capability			0.232***	0.233***	0.232***
			(0.000)	(0.000)	(0.000)
Constant	4.394***	4.414***	4.336***	3.458***	2.905***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.005	0.005	0.013	0.072	0.070
Demographics FE	No	No	No	No	Yes
Observations	1,668	1,668	1,668	1,664	1,664

Table 4. Gender heterogeneity

This table shows the coefficients of *Update* (the difference between posterior and prior expectation for the 12-month-ahead S&P500 index return) and *Expected Sharpe Ratio* (the ratio of posterior return expectation and risk perception) on a set of independent variables. The perception gap (*Perc. gap*) is the difference between the analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. Treatments are *Man + Machine* (I_C) and *Machine* (I_{AI}). All other control variables are defined in Table 1. Demographics fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	**Update**			**Expected Sharpe ratio**		
	1.	2.	3	4.	5.	6.
$Perc.gap\ x\ I_{AI}\ x\ Female$	0.135**	0.123**	0.121**	0.078**	0.073**	0.075**
	(0.014)	(0.025)	(0.028)	(0.013)	(0.019)	(0.017)
$Perc.gap\ x\ I_C\ x\ Female$	-0.040	-0.043	-0.049	-0.036	-0.038	-0.040
	(0.450)	(0.418)	(0.355)	(0.241)	(0.209)	(0.185)
$Perc.gap\ x\ Female$	0.163***	0.169***	0.171***	0.081***	0.081***	0.079***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$I_{AI}\ x\ Female$	0.060	0.121	0.056	0.336	0.394	0.395
	(0.892)	(0.782)	(0.899)	(0.182)	(0.116)	(0.120)
$I_C\ x\ Female$	-0.687	-0.618	-0.651	-0.540**	-0.493**	-0.515**
	(0.110)	(0.150)	(0.132)	(0.029)	(0.046)	(0.038)
$Perc.gap\ x\ I_{AI}$	-0.080**	-0.073**	-0.073**	-0.078***	-0.077***	-0.079***
	(0.031)	(0.048)	(0.049)	(0.000)	(0.000)	(0.000)
$Perc.gap\ x\ I_C$	-0.028	-0.020	-0.018	-0.011	-0.010	-0.011
	(0.430)	(0.577)	(0.608)	(0.574)	(0.629)	(0.587)
$Perc.gap$	0.445***	0.435***	0.433***	-0.184***	-0.187***	-0.187***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
I_{AI}	0.178	0.158	0.189	-0.331**	-0.351**	-0.341**
	(0.534)	(0.580)	(0.509)	(0.044)	(0.032)	(0.038)
I_C	0.433	0.442	0.449	0.350**	0.336**	0.327**
	(0.124)	(0.116)	(0.122)	(0.031)	(0.037)	(0.044)
Female	0.614**	0.794***	0.856***	0.179	0.189	0.2000
	(0.044)	(0.010)	(0.007)	(0.305)	(0.285)	(0.262)
Constant	-0.584***	-2.951***	-3.717***	1.423***	1.226***	1.082**
	(0.004)	(0.000)	(0.000)	(0.000)	(0.001)	(0.038)
R-squared	0.565	0.571	0.572	0.332	0.343	0.344
Controls	No	Yes	Yes	No	Yes	Yes
Demographics FE	No	No	Yes	No	No	Yes
Observations	1,668	1,664	1,664	1,668	1,664	1,664

Table 5. AI literacy

This table shows the coefficients of *Update* (the difference between posterior and prior expectation for the 12-month-ahead S&P500 index return) and *Expected Sharpe Ratio* (the ratio of posterior return expectation and risk perception) on a set of independent variables. The perception gap (*Perc. gap*) is the difference between the analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. Treatments are *Man + Machine* (I_C) and *Machine* (I_{AI}). All other control variables are defined in Table 1. Demographics fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	**Update**			**Expected Sharpe ratio**		
	1.	2.	3	4.	5.	6.
$Perc. gap\ x\ I_{AI}\ x$ AI literacy	0.030	0.032	0.034	0.027*	0.027*	0.028*
	(0.234)	(0.204)	(0.184)	(0.068)	(0.061)	(0.052)
$Perc. gap\ x\ I_C\ x$ AI literacy	0.071***	0.078***	0.082***	0.038***	0.041***	0.041***
	(0.004)	(0.001)	(0.001)	(0.007)	(0.004)	(0.003)
$Perc. gap\ x$ AI literacy	-0.088***	-0.090***	-0.091***	-0.047***	-0.047***	-0.048***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$I_{AI}\ x$ AI literacy	0.384*	0.369*	0.374*	0.189	0.183	0.180
	(0.074)	(0.084)	(0.082)	(0.126)	(0.135)	(0.143)
$I_C\ x$ AI literacy	0.513**	0.535***	0.547***	0.155	0.163	0.150
	(0.013)	(0.009)	(0.008)	(0.189)	(0.166)	(0.203)
$Perc. gap\ x\ I_C$	-0.345***	-0.367***	-0.386***	-0.185***	-0.196***	-0.201***
	(0.001)	(0.000)	(0.000)	(0.002)	(0.001)	(0.001)
$Perc. gap\ x\ I_{AI}$	-0.161	-0.168	-0.175*	-0.160***	-0.163***	-0.168***
	(0.127)	(0.110)	(0.098)	(0.008)	(0.007)	(0.005)
$Perc. gap$	0.881***	0.882***	0.885***	0.045	0.044	0.046
	(0.000)	(0.000)	(0.000)	(0.286)	(0.292)	(0.281)
I_{AI}	-1.481	-1.413	-1.426	-1.018*	-0.990*	-0.968*
	(0.108)	(0.123)	(0.121)	(0.054)	(0.059)	(0.066)
I_C	-1.994**	-2.073**	-2.132**	-0.530	-0.570	-0.540
	(0.023)	(0.018)	(0.015)	(0.293)	(0.255)	(0.283)
AI literacy	-0.327**	-0.382***	-0.396***	-0.083	-0.128	-0.124
	(0.027)	(0.010)	(0.008)	(0.325)	(0.131)	(0.146)
Constant	1.063*	-0.406	-1.225	1.861***	2.285***	2.068***
	(0.095)	(0.623)	(0.248)	(0.000)	(0.000)	(0.001)
R-squared	0.553	0.559	0.559	0.316	0.326	0.328
Controls	No	Yes	Yes	No	Yes	Yes
Demographics FE	No	No	Yes	No	No	Yes
Observations	1,668	1,664	1,664	1,668	1,664	1,664

Table 6. Political affiliation

This table shows the coefficients of *Update* (the difference between posterior and prior expectation for the 12-month-ahead S&P500 index return) and *Expected Sharpe Ratio* (the ratio of posterior return expectation and risk perception) on a set of independent variables. The perception gap (*Perc. gap*) is the difference between the analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. Treatments are *Man + Machine* (I_C) and *Machine* (I_{AI}). All other control variables are defined in Table 1. Demographics fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	**Update**			**Expected Sharpe ratio**		
	1.	2.	3	4.	5.	6.
$Perc.gap\ x\ I_{AI}\ x\ Democrat$	0.163***	0.176***	0.165***	0.111***	0.115***	0.109***
	(0.004)	(0.002)	(0.004)	(0.001)	(0.000)	(0.001)
$Perc.gap\ x\ I_C\ x\ Democrat$	0.125**	0.139**	0.127**	0.107***	0.108***	0.105***
	(0.037)	(0.020)	(0.034)	(0.002)	(0.002)	(0.002)
$Perc.gap\ x\ Democrat$	0.027	0.020	0.024	-0.032	-0.032	-0.029
	(0.502)	(0.618)	(0.549)	(0.170)	(0.162)	(0.210)
$I_{AI}\ x\ Democrat$	0.087	0.209	0.176	-0.099	0.337	0.312
	(0.848)	(0.643)	(0.699)	(0.706)	(0.194)	(0.230)
$I_C\ x\ Democrat$	-0.249	-0.204	-0.208	0.328	-0.118	-0.129
	(0.586)	(0.654)	(0.650)	(0.208)	(0.652)	(0.622)
$Perc.gap\ x\ I_{AI}$	-0.103***	-0.107***	-0.103***	-0.094***	-0.097***	-0.096***
	(0.004)	(0.002)	(0.004)	(0.000)	(0.000)	(0.000)
$Perc.gap\ x\ I_C$	-0.085***	-0.083***	-0.081**	-0.063***	-0.064***	-0.065***
	(0.008)	(0.008)	(0.011)	(0.001)	(0.000)	(0.000)
$Perc.gap$	0.517***	0.514***	0.510***	-0.132***	-0.133***	-0.135***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
I_{AI}	0.153	0.114	0.131	-0.305*	-0.301*	-0.283*
	(0.586)	(0.685)	(0.643)	(0.059)	(0.061)	(0.080)
I_C	0.374	0.382	0.364	0.248	0.256	0.239
	(0.171)	(0.161)	(0.184)	(0.113)	(0.101)	(0.128)
Democrat	0.544*	0.490	0.492	0.184	0.157	0.169
	(0.086)	(0.121)	(0.122)	(0.311)	(0.386)	(0.352)
Constant	-0.552***	-2.491***	-2.944***	1.396***	1.353***	1.330***
	(0.005)	(0.000)	(0.001)	(0.000)	(0.000)	(0.008)
R-squared	0.551	0.559	0.558	0.312	0.325	0.327
Controls	No	Yes	Yes	No	Yes	Yes
Demographics FE	No	No	Yes	No	No	Yes
Observations	1,668	1,664	1,664	1,668	1,664	1,664

Table 7. Horse race

This table shows the coefficients of *Update* (the difference between posterior and prior expectation for the 12-month-ahead S&P500 index return) and *Expected Sharpe Ratio* (the ratio of posterior return expectation and risk perception) on a set of independent variables. The perception gap (*Perc. gap*) is the difference between the analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. Treatments are *Man + Machine* (I_C) and *Machine* (I_{AI}). All other control variables are defined in Table 1. Demographics fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	**Update**			**Expected Sharpe ratio**		
	1.	2.	3	4.	5.	6.
$Perc. gap\ x\ I_{AI}\ x$ Female	0.119**	0.108**	0.107*	0.073**	0.069**	0.072**
	(0.031)	(0.050)	(0.055)	(0.022)	(0.028)	(0.025)
$Perc. gap\ x\ I_{C}\ x$ Female	-0.014	-0.011	-0.016	-0.026	-0.026	-0.028
	(0.796)	(0.830)	(0.767)	(0.404)	(0.411)	(0.371)
$Perc. gap\ x\ I_{AI}\ x$ AI literacy	0.034	0.034	0.036	0.026*	0.026*	0.027*
	(0.178)	(0.171)	(0.157)	(0.073)	(0.073)	(0.061)
$Perc. gap\ x\ I_{C}\ x$ AI literacy	0.079***	0.086***	0.089***	0.037***	0.039***	0.039***
	(0.001)	(0.000)	(0.000)	(0.010)	(0.006)	(0.006)
$Perc. gap\ x\ I_{AI}\ x$ Democrat	0.086	0.098*	0.089	0.066**	0.069**	0.065**
	(0.131)	(0.083)	(0.120)	(0.044)	(0.033)	(0.048)
$Perc. gap\ x\ I_{C}\ x$ Democrat	0.100*	0.114*	0.103*	0.093***	0.094***	0.092***
	(0.087)	(0.052)	(0.078)	(0.006)	(0.005)	(0.007)
$Perc. gap$	0.769***	0.768***	0.768***	-0.011	-0.011	-0.009
	(0.000)	(0.000)	(0.000)	(0.799)	(0.804)	(0.841)
Constant	0.515	-1.107	-1.818*	1.664***	1.956***	1.828***
	(0.429)	(0.179)	(0.083)	(0.000)	(0.000)	(0.003)
R-squared	0.578	0.584	0.583	0.346	0.354	0.356
Interactions	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Demographics FE	No	No	Yes	No	No	Yes
Observations	1,668	1,664	1,664	1,668	1,664	1,664

Table 8. Robustness: Updating

This table shows the coefficients of *Update* (the difference between posterior and prior expectation for the 12-month-ahead S&P500 index return) and *Expected Sharpe Ratio* (the ratio of posterior return expectation and risk perception) on a set of independent variables for Experiment 2 (i.e., ChatGPT). The perception gap (*Perc. gap*) is the difference between the analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. The treatments are *Man+Machine* (I_C) and *Machine* (I_{AI}). All other control variables are defined in Table 1. Demographics fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	**Update**				**Expected Sharpe ratio**			
	1.	2.	3	4.	5.	6.	7.	8.
$Perc. gap\ x\ I_{AI}$	-0.096*	-0.095*	-0.096*	-0.101*	-0.060**	-0.056**	-0.047*	-0.049**
	(0.092)	(0.094)	(0.098)	(0.090)	(0.013)	(0.021)	(0.051)	(0.047)
$Perc. gap\ x\ I_C$	-0.059	-0.061	-0.057	-0.061	-0.025	-0.024	-0.024	-0.025
	(0.271)	(0.253)	(0.289)	(0.269)	(0.279)	(0.295)	(0.279)	(0.269)
I_{AI}	-0.164	-0.115	-0.104	-0.017	0.015	0.030	0.087	0.121
	(0.688)	(0.779)	(0.801)	(0.968)	(0.929)	(0.863)	(0.611)	(0.485)
I_C	-0.216	-0.212	-0.193	-0.082	-0.082	-0.074	-0.074	-0.030
	(0.600)	(0.608)	(0.641)	(0.846)	(0.638)	(0.671)	(0.699)	(0.865)
$Perc. gap$	0.549***	0.552***	0.550***	0.552***	-0.115***	-0.116***	-0.114***	-0.112***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Risk preference		0.249**	0.198*	0.210*		0.009	-0.020	-0.002
		(0.012)	(0.060)	(0.053)		(0.835)	(0.653)	(0.957)
AI literacy		0.007	-0.007	0.023		0.061	0.049	0.052
		(0.959)	(0.954)	(0.864)		(0.250)	(0.354)	(0.344)
Financial literacy		0.037	0.037	0.042		0.083	0.083	0.066
		(0.819)	(0.823)	(0.803)		(0.218)	(0.221)	(0.346)
Female			-0.227	-0.235			-0.282**	-0.301**
			(0.457)	(0.455)			(0.027)	(0.021)
Age			0.002	0.008			-0.012**	-0.010**
			(0.851)	(0.502)			(0.015)	(0.045)
AI trust			0.027	0.010			0.001	-0.005
			(0.874)	(0.954)			(0.987)	(0.941)
AI capability			0.123	0.148			-0.037	-0.026
			(0.482)	(0.414)			(0.616)	(0.725)
Constant	0.209	-1.056	-1.408	-1.202	1.643***	1.022***	1.942***	1.534**
	(0.469)	(0.241)	(0.214)	(0.441)	(0.000)	(0.007)	(0.000)	(0.018)
R-squared	0.466	0.470	0.468	0.461	0.284	0.285	0.297	0.290
Demographics FE	No	No	No	Yes	No	No	No	Yes
Observations	558	558	555	555	558	558	555	555

Table 9. Robustness: Credibility

This table shows the coefficients of *Credibility*, measured as posterior perceptions, on a set of independent variables for Experiments 2 (ChatGPT) and 3 (Reframing). The perception gap (*Perc. gap*) is the difference between the analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. Treatments are *Man+Machine* (I_C) and *Machine* (I_{AI}). Other control variables are defined in Table 1. Demographics fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Experiment:	**ChatGPT**				**Reframing**			
	1.	2.	3	4.	5.	6.	7.	8.
I_{AI}	-0.363***	-0.335**	-0.384***	-0.372***	-0.294**	-0.272**	-0.268**	-0.281**
	(0.007)	(0.013)	(0.003)	(0.006)	(0.023)	(0.037)	(0.035)	(0.028)
I_C	-0.184	-0.170	-0.191	-0.182	-0.221*	-0.197	-0.185	-0.207*
	(0.159)	(0.189)	(0.138)	(0.179)	(0.084)	(0.124)	(0.138)	(0.100)
$Perc. gap$	0.017*	0.019**	0.016*	0.015*	-0.008	-0.007	-0.003	-0.004
	(0.066)	(0.045)	(0.091)	(0.100)	(0.310)	(0.392)	(0.690)	(0.579)
Risk preference		0.136***	0.100***	0.097**		0.057	0.021	0.026
		(0.000)	(0.010)	(0.014)		(0.125)	(0.571)	(0.497)
AI literacy		-0.002	0.001	0.005		-0.032	-0.045	-0.034
		(0.970)	(0.983)	(0.917)		(0.514)	(0.354)	(0.487)
Financial literacy		-0.061	-0.017	0.004		-0.059	-0.026	-0.029
		(0.282)	(0.763)	(0.950)		(0.290)	(0.646)	(0.611)
Female			0.194*	0.211*			-0.017	-0.024
			(0.087)	(0.062)			(0.871)	(0.822)
Age			-0.001	-0.000			0.001	0.001
			(0.743)	(0.994)			(0.688)	(0.827)
AI trust			0.111*	0.110*			0.085	0.079
			(0.086)	(0.088)			(0.141)	(0.180)
AI capability			0.127*	0.133*			0.178***	0.180***
			(0.062)	(0.051)			(0.003)	(0.003)
Constant	4.550***	4.210***	3.172***	2.917***	4.377***	4.503***	3.463***	3.613***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.015	0.036	0.086	0.084	0.007	0.009	0.066	0.063
Demographics FE	No	No	No	Yes	No	No	No	Yes
Observations	558	558	555	555	557	557	557	557

Table 10. Manipulations

This table shows the coefficients of *Return Revision*, defined on a seven-point scale, on a set of independent variables. Treatments are *Man + Machine* (I_C) and *Machine* (I_{AI}). For Columns 1 and 2, manipulations are defined as the forecast, including the earning estimate (M_1) and adding a visual (M_2). For Columns 3 and 4, the manipulations are low (M_1) and high (M_2) analyst dispersion. For Columns 5 and 6, the manipulations are a downgrade without (M_1) and with an earnings estimate (M_2). All other control variables are defined in Table 1. The demographics fixed effects include education and income levels. P-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	**Return revision**					
	Information richness		Forecast dispersion		Analyst downgrade	
	1.	2.	3.	4.	5.	6.
M_1	0.257***	0.064	-0.144	-0.216*	-0.591***	-0.622***
	(0.009)	(0.664)	(0.106)	(0.086)	(0.000)	(0.000)
M_2	0.362***	0.233	0.105	0.020	-0.528***	-0.487***
	(0.000)	(0.137)	(0.134)	(0.869)	(0.000)	(0.000)
I_{AI}	0.051	-0.053	0.031	-0.104	-0.025	-0.032
	(0.592)	(0.747)	(0.656)	(0.389)	(0.744)	(0.813)
I_C	0.041	-0.147	-0.001	-0.109	-0.079	-0.064
	(0.676)	(0.375)	(0.986)	(0.382)	(0.286)	(0.625)
I_{AI} x M_1		0.140		0.215		-0.002
		(0.545)		(0.222)		(0.992)
I_{AI} x M_2		0.138		0.036		-0.073
		(0.549)		(0.836)		(0.701)
I_C x M_1		0.332		0.145		0.079
		(0.157)		(0.404)		(0.665)
I_C x M_2		0.269		0.237		-0.096
		(0.249)		(0.162)		(0.605)
Constant	-0.154	1.549***	3.039***	3.483***	3.056***	3.149***
	(0.258)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.033	0.111	0.011	0.046	0.011	0.142
Controls	No	Yes	No	Yes	No	Yes
Demographics FE	No	Yes	No	Yes	No	Yes
Observations	544	543	557	554	543	543

Table 11. AI model disclosure

This table shows the coefficients of *Update* (the difference between posterior and prior expectation for the 12-month-ahead S&P500 index return) and *Expected Sharpe Ratio* (the ratio of posterior return expectation and risk perception) on a set of independent variables for Experiment 4 (AI model disclosure). Perception gap (*Perc. gap*) is the difference between the analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. Treatments are "an AI model using BLUE (I_{BLUE}), OLS (I_{OLSI}), or deep learning methods" (I_{DLI}) . All other control variables are defined in Table 1. Demographics fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable:	Update				Expected Sharpe ratio			
	1.	2.	3	4.	5.	6.	7.	8.
$Perc. gap\ x\ I_{OLS}$	0.184***	0.189***	0.180***	0.179***	0.004	0.009	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.885)	(0.750)	(0.959)	(0.976)
$Perc. gap\ x\ I_{BLUE}$	-0.008	-0.012	-0.016	-0.018	-0.042	-0.041	-0.476*	-0.049*
	(0.879)	(0.828)	(0.770)	(0.747)	(0.150)	(0.151)	(0.094)	(0.086)
$Perc. gap\ x\ I_{DL}$	-0.019	-0.028	-0.034	-0.025	-0.030	-0.029	-0.040	-0.039
	(0.706)	(0.580)	(0.502)	(0.626)	(0.250)	(0.273)	(0.121)	(0.132)
I_{OLS}	0.409	0.334	-0.342	0.277	-0.094	-0.143	-0.175	-0.187
	(0.336)	(0.431)	(0.421)	(0.516)	(0.675)	(0.519)	(0.423)	(0.391)
I_{BLUE}	0.147	0.091	0.152	0.106	-0.243	-0.291	-0.286	-0.255
	(0.732)	(0.831)	(0.724)	(0.807)	(0.282)	(0.194)	(0.195)	(0.248)
I_{DL}	-0.301	-0.477	-0.456	-0.489	-0.049	-0.144	-0.162	-0.173
	(0.453)	(0.235)	(0.258)	(0.228)	(0.818)	(0.492)	(0.432)	(0.404)
$Perc. gap$	0.381***	0.379***	0.383***	0.377***	-0.172***	-0.176***	-0.169***	-0.169***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Risk preference		0.066	0.072	0.044		-0.016	-0.027	-0.045
		(0.450)	(0.433)	(0.634)		(0.736)	(0.572)	(0.346)
AI literacy		-0.089	-0.105	-0.126		0.024	-0.001	-0.009
		(0.424)	(0.352)	(0.268)		(0.682)	(0.986)	(0.864)
Financial literacy		0.459***	0.438***	0.437***		0.262***	0.266***	0.243***
		(0.001)	(0.001)	(0.001)		(0.000)	(0.000)	(0.000)
Female			-0.346	-0.296			-0.309**	-0.299**
			(0.194)	(0.273)			(0.024)	(0.030)
Age			0.001	0.004			-0.021***	-0.020***
			(0.936)	(0.755)			(0.000)	(0.001)
AI trust			0.031	0.084			0.063	0.060
			(0.834)	(0.578)			(0.408)	(0.433)
AI capability			-0.202	-.231			-0.175**	-0.165**
			(0.193)	(0.140)			(0.028)	(0.040)
Constant	-0.232	-1.933**	-0.979	-0.549	1.730***	0.685*	2.264***	3.330***
	(0.432)	(0.011)	(0.307)	(0.646)	(0.000)	(0.084)	(0.000)	(0.000)
R-squared	0.455	0.466	0.467	0.470	0.387	0.401	0.427	0.435
Demographics FE	No	No	No	Yes	No	No	No	Yes
Observations	571	571	570	570	571	571	571	571