Algorithm Aversion, Appreciation, and Investor Return Beliefs*

Francesco Stradi
KU Leuven

Gertjan Verdickt
University of Auckland

First draft: September 10, 2024

This draft: September 19, 2025

Abstract

Do investors trust AI analyst forecasts? Three incentivized experiments with 3,000 U.S. participants highlight that the average investor is less responsive to forecasts when AI is incorporated – suffering from algorithm aversion. The decrease in trust stems from a low perceived credibility. Interestingly, there are specific groups that exhibit algorithm appreciation, like sophisticated investors, those with higher AI literacy, and those identifying as politically progressive. Given AI's increasing role in financial markets, our results highlight the practical implications for improving trust in AI-generated forecasts.

KEYWORDS: Artificial intelligence, behavioral finance, analyst forecasts, ChatGPT, trust, experimental economics, return beliefs

JEL CLASSIFICATION: C91, G02, G11, G12, G51, G53, M41

---

* We are grateful to Daan Ballegeer, Henk Berkman, Justin Case, Simona Fabrizi, Heng Griffin Geng, Andrew Grant (discussant), Marina Gertsberg, Pablo Ruiz-Verdú, Thomas Walther, Norman Wong, Rafael Zambrana, Fernando Zapatero and all participants at the Auckland University of Technology, Massey University, New Zealand Finance Colloqium (2025), Sydney Banking and Financial Stability Conference (2024), 32nd Finance Forum, and the University of Auckland for their helpful comments. Francesco Stradi gratefully acknowledges the financial support from Research Foundation Flanders (FWO) Grant 11P7A24N. Our experiment underwent IRB review at the University of Auckland and was approved on August 13, 2025, with reference number UAHPEC28279. This paper was circulated as "Man vs. Machine: The Influence of AI Forecasts on Investor Beliefs".

Corresponding author: gertjan.verdickt@auckland.ac.nz. All remaining errors are our own.

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)

"A Chatbot Is Now Cleared to Give Buy, Sell Investment Advice"

(Bloomberg, 2024)

As artificial intelligence (AI) continues to shape financial markets, a key question remains: Do investors trust machine-generated forecasts? Historically, they exhibited algorithm aversion – a persistent reluctance to rely on machine recommendations (see Castelo et al., 2019; Dietvorst et al., 2014; Germann & Merkle, 2023; Longoni et al., 2019; Önkal et al., 2009).[1] As AI becomes (more) embedded in financial decision-making, it is unclear whether this aversion remains or whether its familiarity (Dang & Liu, 2026; Lee & See, 2004; Tsz et al., 2024) and outperformance over humans (Asparouhova et al., 2024; Bertomeu et al., 2025; Fedyk et al., 2022; Hodge et al., 2021; Lopez-Lira & Tang, 2025) might have mitigated this scepticism – potentially leading to "algorithm appreciation".

In this study, we revisit the phenomenon of algorithm aversion. More specifically, we examine whether the investors distrust AI-generated financial forecasts and identify factors that shaped this resistance. To answer this, we conduct an incentivized experiment in which 1,800 investors were asked about their return expectations before *and* after receiving a forecast for the S&P 500 (Giglio et al., 2021). However, we randomly assign participants to one of three forecast sources: (I) human analysts (i.e., Goldman Sachs), (II) human analysts using an AI model, or (III) a pure AI model. This randomized design ensures that differences in the posterior variables could be causally attributed to the source of the forecast .[2] To make our setting realistic, we use the exact wording from the Goldman Sachs' 2023 equity outlook.[3] This approach makes the information

[1] AI differs in many ways from the more traditional algorithms considered in previous studies, such as Önkal et al. (2009). Its superior performance, broader awareness, and greater adoption compared to established algorithms may have rendered algorithm aversion nonexistent.

[2] Individuals assigned to treatment I form the *Human* group, while those in treatments II and III form the *AI* group.

[3] This information is always released in November.

of the source salient and keeps the content (i.e., a 5% forecast) constant, allowing us to directly study the effect of the source on return belief updating (Beutel & Weber, 2024; Laudenbach et al., 2024).[4]

Our results show that algorithm aversion remains a significant factor in how investors process financial forecasts. Rather than observing a uniform negative reaction to AI, we highlight how investors' responsiveness to (new) information varies based on the source. We measure this by interacting the forecast provider with the *perception gap*—the distance between investors' initial belief and the forecast they received (Haaland et al., 2023; Laudenbach et al., 2024). Our key finding is that investors learn less from AI; they close this perception gap to a significantly smaller degree when the forecast comes from an AI source compared to a human one. Overall, this demonstrates a lower willingness to trust (and update from) the AI signal, especially when it challenges their pre-existing beliefs.

Next, we study what drives the (lack of) return updating. We find that the higher the perceived credibility of the information received, the greater the belief adjustment toward the provided forecast. Indeed, we show that AI aversion is closely linked to credibility perceptions: investors rated AI-powered forecasts as significantly less credible than predictions generated by human analysts.

Beyond establishing algorithm aversion, we examine the factors that shape investor responses to AI forecasts. Both cognitive and non-cognitive traits—AI literacy, investment sophistication, gender, and political affiliation—can play a significant role, though their effects are not always clear-cut, ex-ante. For instance, high AI literacy can increase trust due to (proven) performance advantages (S. Cao et al., 2024), but recent declines in model accuracy (L. Chen et al., 2024) could dampen its enthusiasm. Sophisticated investors might value AI's analytical strengths yet still prefer human analysis in complex settings. Women adopt financial technology less frequently (S. Chen et al., 2023) but are more likely to seek financial advice (Bucher-Koenen et al., 2023), potentially making them receptive to an AI tool, in particular, given evidence of biases among human advisors (Bhattacharya et al., 2024). Finally, progressives are more open

[4] Similar information-provision experiments are frequently employed in the economics literature (see, e.g., Haaland et al., 2023).

to innovation and thus more receptive to AI, while conservatives often prefer traditional solutions (Carney et al., 2008; Blank & Shaw, 2015; Cookson et al., 2020; Barrios & Hochberg, 2021).

Cognitively, investors with high AI literacy and sophistication exhibit algorithm appreciation, suggesting that familiarity with AI capabilities mitigates its scepticism. Overall, these findings extend previous studies documenting the importance of (different types of) literacy in making financial decisions (e.g., Filippini et al., 2024; Lusardi & Tufano, 2015). Non-cognitively, while gender effects are nuanced, political affiliation emerges as a significant predictor. Indeed, the Democrat-affiliated investors update their return beliefs more strongly toward the AI forecast. This shows that while the average investor exhibits algorithmic aversion, there are groups that have algorithm appreciation.

We add various robustness analyses to further strengthen confidence in our conclusions. For instance, using the posterior *Expected Sharpe ratio* as a dependent variable—which incorporates return expectations alongside risk perceptions—we show persistent algorithm aversion across all treatments involving AI. In particular, investors whose prior beliefs deviate more strongly from the AI's forecast experience a lower expected Sharpe ratio under the *AI* treatment than in the control condition (i.e., Human analysts). This indicates a lower willingness to integrate the AI's signal into posterior expectations among participants facing a larger gap between the AI projection and their initial expectations.

We also conduct two additional experiments with 1,200 new participants. First, replacing "AI" with "ChatGPT" confirms that algorihm aversion persists even when comparing two specific, named entities. Second, alternating the timing of the source revelation – by providing forecasts before disclosing their source – demonstrates that source attribution clearly impacts perceived credibility independent of information content (Johnson & Markelevich, 2025). This provides robust evidence that the average investors exhibits algorithm aversion. Moreover, it shows that the source provider plays a key role and should be an important lesson for policymakers to push for the disclosure of AI usage, for instance in analyst forecasting. It also contributes to the current debate on the risk and opportunities of incorporating AI into accounting and auditing (Eisikovits et al., 2025).

Our paper makes several contributions. First, we advance the literature on algorithm aversion in financial contexts (Dietvorst et al., 2014; Germann & Merkle, 2023; Greig et al., 2025; Longoni et al., 2019; Önkal et al., 2009) by showing its persistence in today's AI-saturated environment. Despite ever-increasing familiarity with AI tools, investors show reluctance to incorporate AI-generated forecasts into their return beliefs compared to identical human forecasts. That said, for specific groups—namely, those with higher AI literacy, greater investor sophistication, and more progressive political views—we find evidence of algorithm appreciation. Moreover, our experimental design, which distinguishes between human analysts, human analysts using AI, and pure AI forecasts, offers a more nuanced understanding of how different combinations of human and machine inputs impact investor trust, extending beyond the previous comparisons of "Man vs. Machine" (S. Cao et al., 2024; Coleman et al., 2022; Pungaliya & Wang, 2023; Rossi & Utkus, 2020).

Second, we identify both cognitive and non-cognitive factors that influence algorithm aversion in financial forecasting. Cognitively, investors with higher AI literacy and sophistication show greater receptivity to AI-generated forecasts. Non-cognitively, political affiliation consistently predicts AI receptivity, with Democrats updating beliefs more strongly toward the AI forecasts regardless of format, while gender effects are more nuanced. These findings contribute to the literature on technology adoption in financial decision-making by highlighting the pathways through which individual characteristics moderate algorithm aversion. Furthermore, we add to the literature investigating different forms of literacy (e.g., Filippini et al., 2024; Lusardi & Mitchell, 2014; Lusardi & Tufano, 2015) by showing that cognitive factors—such as AI literacy or investor sophistication—are important in mitigating algorithm aversion.

Finally, 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 analysts better dissect balance sheet information (see, e.g., Cao et al., 2024; Jha et al., 2024; Kim et al., 2024) or provide investment advice (Chak et al., 2022; Coleman et al., 2022; Hodge et al., 2021; Jiang et al., 2024). We are the first to examine investors' reactions to AI forecasts from a credibility and investor belief perspective. As such, we bring back "the human component" into the equation. Overall, this revolves around the question of trust in financial markets and in technology (e.g., Y. Cao & Zhang, 2025; Davis, 1989; Glikson & Williams Woolley, 2020; Lee & See, 2004).

2. Study design

As motivated by the introduction, we experimentally examine the impact of forecast providers on investor beliefs. The key research question is whether algorithm aversion still exists. On the one hand, there is a large (psychology) literature documenting the persistence of this aversion across different fields (Castelo et al., 2019; Dietvorst et al., 2014; Germann & Merkle, 2023). On the other hand, there is a growing amount of evidence showing that AI outperforms most tasks over humans (e.g., Asparouhova et al., 2024; Aubry et al., 2023; Fedyk et al., 2022; Hodge et al., 2021; Lopez-Lira & Tang, 2025). Furthermore, the growing familiarity can decrease reluctance to follow AI-generated advice (Lee & See, 2004; Tsz et al., 2024).

2.1. Survey flow

We conduct an incentivized experiment to examine these hypotheses.

Baseline, prior beliefs. In Part 1 of the survey, participants were asked about their traits and 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).

To ensure the integrity of the randomization of the participants in Part 2, we include a question related to AI. More specifically, we add the following measure to control for heterogeneity in AI usage:

> *"In general, how useful has ChatGPT (or similar artificial intelligence tools) been to you?"*

This question is recorded on a scale from 1 to 7 and is inspired by YouGov surveys.[5] We label this measure *ChatGPT usefulness*.

[5] To further test the integrity of the randomization, we asked individuals for their opinion on how capable artificial intelligence (AI) is at making financial forecasts compared to Goldman Sachs' analysts (*AI relative capabilities*).

To poll the expected return priors, we borrow the question of Giglio et al. (2021). Figure 1 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 1]

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:

- ***Human***: *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.*
- ***Human + Machine*** *(HM)*: *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*** *(M)***:** *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' stock market outlook introduction**.** Goldman Sachs validates our setup as a renowned investment bank that issues annual equity predictions and already incorporates AI.[6] Second, we verify that AI tools (i.e., ChatGPT) could produce a similar forecast independently, confirming the plausibility of the financial information given to the participants (cfr. Figure A.1). This approach grounds our setting in an authentic financial industry practice, while accurately representing current AI capabilities.

Posterior beliefs. After receiving one of the forecasts, respondents are again asked to provide a one-year-ahead forecast of the S&P 500 index (cfr. Figure A.2). 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 I).

2.3. Participants

In the Experiment, we recruit 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.[7]

[INSERT TABLE 1]

Sample statistics. Panel A of Table 1 presents the summary statistics, and Table A.1 provides the variable definitions.[8] Participants expect the 12-month ahead return of 8.47% for the S&P 500, with a confidence of 4.89 out of 7. They report a risk preference of 4.08 on a seven-point scale and demonstrate a higher level of AI (4.07 on a 5-point scale) and financial literacy (4.23 out of 5).

Integrity of the randomization. Panel B of Table 1 presents the randomization of respondents. When randomly assigning participants to one of the groups, we find no significant differences

[6] https://www.wsj.com/articles/goldman-sachs-deploys-its-first-generative-ai-tool-across-the-firm-cd94369b

[7] In total, we rejected 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.

[8] We drop participants with overly optimistic expectations, below -30% and above 30% (Merkoulova & Veld, 2022). The results do not change when we include these participants.

across the average investor characteristics. We include these control variables in our regression specifications to eliminate concerns. Doing so ensures that other characteristics are not driving the changes in investor beliefs across forecast providers.

3. Is there evidence of algorithm aversion?

Algorithm aversion—the reluctance to trust machine-generated predictions over humans – has been well-documented across various contexts (e.g., Dietvorst et al., 2014; Longoni et al., 2019). While AI has demonstrated the ability to generate accurate finance predictions, investors may still resist incorporating AI-generated insights into their return expectations, thereby revealing algorithm aversion. The presence of algorithm aversion in the context of AI financial forecasts is not assured, however, as increased exposure to AI could have led to greater acceptance over time (Lee & See, 2004; Tsz et al., 2024), specifically in the task of forecasting returns, where AI has been shown to perform very well, even in comparison to human analysts (S. Cao et al., 2024; Fedyk et al., 2022; Lopez-Lira & Tang, 2025).

For this reason, we first study whether "algorithm aversion" toward AI is present in the context of return expectation updates. Specifically, we examine the following hypothesis to determine if investors exhibit scepticism toward AI-generated forecasts:

Hypothesis 1a: Investors will update their return expectations less in response to AI than to human forecasts, consistent with the evidence of algorithm aversion.

Conversely, if increased AI awareness or adoption reduced scepticism, investors could update their return beliefs regardless of whether the forecast is AI- or human-generated. Indeed, given AI's growing importance in financial markets, algorithm aversion can differ compared to other tasks and fields.

Hypothesis 1b: Investors will update their return expectations equally for AI and human forecasts, suggesting that AI familiarity has reduced scepticism toward algorithms.

Finally, investors might also update their beliefs more when AI is involved. As shown in prior literature, the capabilities of algorithms in explaining stock market movements have increased substantially over the last few years. As such, investors might follow AI financial advice more and thus exhibit evidence of algorithm appreciation.

Hypothesis 1c: Investors will update their return expectations more in response to AI than to human forecasts, consistent with the evidence of algorithm appreciation.

3.1. Methods

In Experiment 1, we study the effect of forecast providers on investor beliefs. More specifically, we calculate differences in return updating across groups. The empirical methodology follows the one illustrated by Haaland et al. (2023).

In particular, 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 return updating if the participants believe this forecast signal is credible (Beutel & Weber, 2024; Goetzmann & Massa, 2005; Haaland et al., 2023; Laudenbach et al., 2024).

To study our research question, we use the following specifications:

$$Update_i = \beta_0 + \beta_1 Treatment_i \, x \, Perception \; gap_i + \beta_2 Treatment_i + \beta_3 Perception \; gap_i + \gamma X_i + \varepsilon_i \quad (1)$$

where $Update_i$ is the difference between posterior and prior return expectations for individual i. Since we are mostly interested in trust in forecasts involving AI, our baseline treatment group ("AI", or I_{AI} in the tables) includes both individuals who received machine-generated forecasts ("Machine," I_M in the tables) and those who received forecasts from a human powered by AI ("Human + Machine," I_{HM} in the tables). These are relative to those who only received human forecasts ("Human"), which serve as the control group. Hence, we choose an active control group, as is recommended for this type of experiment (e.g., Haaland et al., 2023). This ensures that our effect is driven mainly by the information provider rather than the information itself, as both the treatment and control groups received the same numeric forecast for S&P 500 returns.

The variable of interest is β_1, which captures the extent to which investors update their priors toward the provided forecast in the AI treatments relative to the control group. This is referred to as "the learning rate" (Haaland et al., 2023; Laudenbach et al., 2024). In the context of our research question, we would find evidence of algorithm aversion if the investors updated their beliefs less when receiving a forecast that incorporates AI—that is, if β_1 is negative.[9]

The coefficient β_2 quantifies the average amount of return updating across treatments, which is not related to the investor's expected returns priors. In our case, this could be related to prior beliefs on AI. The coefficient β_3 captures the average learning rate, unrelated to the treatments.

Finally, X_i is a vector of control variables, including demographic and socioeconomic factors such as gender, age, education, income groups, risk preference, financial literacy, and a binary variable related to investor sophistication, which equals one if the investor has experience with options, angel investing, private equity, or venture capital, and one variable related to political affiliation, where we focus on those identifing as Democrats, proxying for progressive political views. Finally, we include the individual's AI literacy and prior beliefs regarding the perceived usefulness of ChatGPT in capturing their knowledge and usage of artificial intelligence.

3.2. Main findings

[INSERT TABLE 2]

Table 2 presents our baseline empirical findings, demonstrating how an investor updates their beliefs in response to forecasts offered by different sources. It focuses on those forecasts using *AI* (*Machine* and *Human + Machine*) relative to a *Human*-only forecast (control group).

We show that the average learning rate ranges from 0.515 to 0.524 (significant at the 1% level). This means that the average investor adjusts their return expectations by approximately 52% toward the forecasted value after receiving the information. These magnitudes are comparable to previous studies (e.g., Beutel & Weber, 2024), indicating that the investors generally respond

[9] Cavallo et al. (2017) and Cullen and Perez-Truglia (2018) demonstrate that the coefficient β_1 can also be interpreted in a more structural manner:, under a Bayesian updating framework with normally distributed priors and signals—where the moments of the distributions are independent—and assuming a quadratic loss function, belief updating follows a linear pattern relative to the perception gap. The coefficient β_1 represents the weight individuals assign to the new signal, while the remaining weight, $1-\beta_1$, is attributed to their prior belief.

to forecast information. Furthermore, we highlight that the interaction between the perception gap (*Perc. gap*) and AI treatment is negative across all specifications (all significant at the 10% level). This indicates that investors update their beliefs less toward the forecasts incorporating AI, particularly when their prior return expectations differed more from the provided forecast. This, therefore, supports the hypothesis of algorithm aversion.

The regression coefficient estimates of the interaction range around -0.040. In economic terms, this implies that for each percentage point difference between prior return belief and the given forecast, an investor updates its expectation four percentage points *less* when AI is involved in the forecast generation. For an average investor in our sample – with a gap of 3.287 percentage points between their prior belief (8.287%) and the forecast (5%) – this translates to roughly 0.13 percentage points less adjustment when receiving AI-involved forecasts compared to human-only forecasts. In other words, a reduction in return updating of approximately 7.6%, which is economically meaningful.

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.

Table A.2. finds that the results are driven by the *Human+Machine* interaction. In other words, investors appear more skeptical of those forecasts. This finding challenges the recent literature asserting superior performance of Human-AI combinations (Cao et al., 2024; Kim et al., 2024; Lopez-Lira & Tang, 2025). It may seem surprising that, on average, forecasts made by humans using AI are less credible than "pure human" ones. However, this pattern appears consistent with insights from psychology literature (Dang & Liu, 2026) and recent findings from finance (Walther & Dutordoir, 2025).

Given that the ability to form reasonable stock return expectations is crucial for the reliability of our empirical exercise, we excluded from our analysis stock return-ignorant individuals—defined as those estimating the annual stock market return to be higher than 30% in absolute value, in line with Merkoulova & Veld (2022). However, our baseline findings on algorithm aversion are likely to be a lower-bound estimate, as including these investors makes our results considerably stronger in both magnitude and statistical significance (cfr. Table A.3).

3.3. Does credibility play a role?

The documented algorithm aversion raises a question about the mechanisms underlying this behavior. While the empirical evidence demonstrates reduced return belief updating when AI is involved, the processes driving this remain unclear. To study this, we examine an investor's subjective evaluations collected after exposure to the forecast. Specifically, we introduce a new variable, *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).

The first question is whether the provider's credibility is important in updating return beliefs. We use the following specification to formally examine whether the general level of credibility can explain variation in return belief updating:

$$Update_i = \beta_0 + \beta_1 Credibility_i \, x \, Perception\ gap_i + \beta_2 Credibility_i + \beta_3 Perception\ gap_i + \gamma X_i + \varepsilon_i \quad (2)$$

where β_1 captures the extent to which an investor with a large difference between prior beliefs and its credibility level update its beliefs toward the forecast. Similar to Equation 1, we include a number of control variables.

[INSERT TABLE 3]

Table 3 reports the regression coefficients. Overall, we document that the higher the perceived credibility of the information received, the greater the belief adjustment is toward the forecast. Therefore, since return belief updating is significantly driven by signal credibility, we examine differences in the perceived credibility of the signals we provided.

[INSERT FIGURE 2]

Figure 2 illustrates the respondents' average beliefs regarding *Credibility* across the treatment and control groups. Panel A highlights a statistically significant difference between treatments involving *AI* (*Machine* or *Human + Machine*) and the control group, which features a *Human*-only forecast. In particular, the *Human*-only forecast is perceived as more credible.

In Panel B, we decompose the AI treatments into *Machine* and *Human + Machine*. We document statistically significant differences between treatments (significant at 5% level), with *Human* being the most credible compared to *Machine* being the least credible forecast provider. This can explain the observed differences in return updating, since we find a pattern indicating that individuals seem to perceive forecasts involving AI as less credible.

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

$$Credibility_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Perception\ gap_i + \gamma X_i + \varepsilon_i \quad (2)$$

where β_1 captures the extent to which investors perceived the two forecasts involving AI (i.e., *Machine* or *Man+Machine*) as having different credibility than *Human,* and the control variables are similar to Equation 1.

[INSERT TABLE 4]

Table 4 presents the results from this regression. In columns 1-3, we highlight a consistent and statistically significant negative coefficient on the AI treatment indicator (significant at the 1% level). Specifically, forecasts involving AI are perceived as approximately 0.17-0.20 points less credible on our 7-point scale relative to human-generated forecasts.[10] This effect remains robust when controlling for demographic characteristics and investor-specific traits.

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

4. Why do investors not update their beliefs more toward the AI forecast?

In the previous section, we show significant evidence of algorithm aversion among the average investors. However, trust in AI-generated forecasts might be shaped by two broad categories: cognitive, from individual knowledge or experience, and non-cognitive factors, such as gender or political beliefs. Indeed, prior literature showed that those factors are related to investment behavior (Bonaparte et al., 2017; Bucher-Koenen et al., 2023; S. Chen et al., 2023; Cookson et al., 2020; Meeuwis et al., 2022; Rossi & Utkus, 2020).

[10] When separating the AI treatments in Table A.4, a nuanced pattern emerges: purely algorithmic (Machine-only) forecasts show the largest credibility decline (–0.224, significant at the 1% level).

This section studies how return updating varies depending on these factors. We use AI literacy and investor sophistication as proxies for cognitive abilities. For non-cognitive characteristics, we focus on gender and political affiliation.

4.1. Cognitive factors

4.1.1. 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 Filippini et al., 2024; Lusardi & Mitchell, 2014; Lusardi & Tufano, 2015).

Given the proven capabilities of AI in forecasting returns, one expects that if cognitive factors contribute to algorithm aversion, individuals with greater knowledge of AI exhibit higher trust in AI-generated forecasts. Conversely, prior research has shown a decrease in the performance of large language models, such as ChatGPT (e.g., L. Chen et al., 2024). This might lead to slower adoption of AI in financial decision-making.

Hypothesis 2a: Investors with higher levels of AI literacy update their return belief toward the signal relatively more when receiving AI-generated forecasts.

Hypothesis 2b: Investors with higher levels of AI literacy update their return belief away from the signal relatively more when receiving AI-generated forecasts.

4.1.2. Sophistication

Beyond AI literacy, an investor's financial sophistication could influence their receptiveness to AI-generated forecasts. Sophisticated investors, with broader exposure to various investment strategies and technologies, might evaluate AI forecasts differently than their less experienced counterparts. These investors could have had more opportunities to compare AI performance with human forecasts in real-world settings, potentially affecting their algorithm aversion.

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

The theoretical relationship between sophistication and algorithm aversion is ambiguous. On one hand, sophisticated investors can better appreciate AI's computational advantages in data processing or pattern recognition, potentially increasing receptivity to AI-generated forecasts.

Less sophisticated market participants, lacking this nuanced perspective, can view humans as more trustworthy information sources.

On the other hand, the experience of sophisticated investors may reinforce the value of human judgment in complex market environments, potentially strengthening their preference for human analysts.

Hypothesis 3a: When receiving an AI-generated forecast, sophisticated investors update their return beliefs toward the signal relatively more.

Hypothesis 3b: When receiving an AI-generated forecast, sophisticated investors update their return beliefs away from the signal.

We classify investors as "sophisticated" if they reported experience with advanced investment vehicles such as options, private equity, venture capital, or angel investing. This classification aligns with previous literature that uses investment diversification and complexity as proxies for sophistication (Barber et al., 2016; Bucher-Koenen et al., 2023; Dumitrescu & Gil-bazo, 2018; Goetzmann & Kumar, 2008).

4.1.3. Empirical results

To test the impact of cognitive factors on return updating, we run the following model:

$$\begin{aligned} Update_i = \beta_0 + \beta_1 \cdot Treatment_i \, x \, Perception\ gap_i \, x \, Cog_i + \beta_2 \\ \cdot Perception\ gap_i \, x \, Cog_i + \beta_3 \cdot Treatment_i \, x \, Cog_i + \beta_4 \\ \cdot Treatment_i \, x \, Perception\ gap_i + \beta_5 \cdot Treatment_i \, + \beta_6 \\ \cdot Perception\ gap_i + \beta_7 \cdot Cog_i + \gamma X_i + \varepsilon_i \end{aligned} \quad (3)$$

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, $Treatment_i$ is a binary variable equal to one for individuals receiving the treatments *Machine* or *Human + Machine*, while those receiving the *Human* message are the control group. Cog_i is a variable that is either AI literacy or Sophistication, as defined above.

[INSERT TABLE 5]

Table 5 reports the findings. The coefficient on the triple interaction term (Perception gap x AI x AI Literacy) equals 0.043 (significant at the 5% level), implying that investors with higher AI literacy update their beliefs relatively more toward AI-generated forecasts relative to humans-only. For Sophistication, the interaction coefficient is equal to 0.167 (significant at the 1% level), with a larger magnitude. Overall, the results highlight that both AI literacy and sophistication mitigate algorithm aversion. In terms of economic magnitude, with an average perception gap of 3.287, sophisticated (higher AI literate) investors updated their return beliefs with 0.55 (0.14) percentage points relatively more, which is economically meaningful. Therefore, we argue that this is evidence in favor of "algorithm appreciation".[12]

Result 3. Investors with high levels of AI literacy update their return beliefs toward the signal relatively more when receiving AI forecasts, consistent with algorithm appreciation.

Result 4. More sophisticated investors update their return beliefs toward the signal relatively more when receiving AI forecasts, consistent with algorithm appreciation.

The conclusions align with the recent literature on the role of domain knowledge in technology adoption. Similar to how financial literacy improves financial decision-making (see Lusardi & Mitchell, 2014; Filippini et al., 2024), our results show that AI literacy enables investors to better evaluate the merits of AI-generated forecasts. This knowledge advantage likely helps investors overcome the initial scepticism that characterizes algorithm aversion. The stronger impact of *Human + Machine* combinations is particularly interesting given the large evidence on human-AI collaboration in finance (e.g., Cao et al., 2024; Kim et al., 2024).

For sophistication, our findings extend previous research on sophisticated investors' adoption of technological innovations (e.g., Goetzmann & Kumar, 2008). The positive effect implies that sophisticated investors are more open to computational approaches. However, this openness varies across contexts and presentation formats. Overall, this response heterogeneity parallels

[12] Table A.5 distinguishes between the two AI treatments. Significant interaction effects for both AI literacy (0.066) and sophistication (0.253) arise specifically in the *Human + Machine* condition (both significant at the 1% level), indicating that investors with higher cognitive ability are more responsive to forecasts that integrate human and AI inputs.

results from other fields where expertise influences technology adoption patterns (Venkatesh et al., 2003).

Interestingly, we introduce a paradox in our understanding of algorithm aversion. While the general investor population exhibits scepticism toward AI, those with greater AI knowledge—who presumably better understand AI's capabilities or limitations—show great receptiveness. This implies that algorithm aversion may partly stem from unfamiliarity rather than a rational evaluation of AI's performance, pointing toward educational interventions as one approach to increase appropriate trust in AI-generated financial forecasts.

4.2. Non-cognitive factors

4.2.1. Gender

There is a growing literature on gender differences in financial decision-making or technology adoption. However, the relationship between gender and algorithm aversion remains unclear. For instance, women are sometimes found to adopt financial technology at sizeably lower rates than men (S. Chen et al., 2023), which can indicate lower receptivity to AI forecasts (i.e., higher algorithm aversion). In contrast, prior research shows that women seek financial advice more frequently (Bucher-Koenen et al., 2023) and could be more receptive to alternative information sources if they perceive it as unbiased. This is particularly relevant given evidence that human financial advice could vary in quality (Bhattacharya et al., 2024) and could be subject to gender-related biases (Bucher-Koenen et al., 2023). As such, we examine the following two hypotheses:

Hypothesis 4a: When receiving an AI-generated forecast, women update their beliefs toward the signal relatively more.

Hypothesis 4b: When receiving an AI-generated forecast, women update their beliefs toward the signal relatively less.

4.2.2. Political affiliation

Like gender, a large body of research demonstrates how political affiliation represents another non-cognitive factor shaping financial market beliefs (e.g., Barrios & Hochberg, 2021; Cookson et al., 2020). Its psychological foundation stems from stable personality traits creating left-right differences, where liberals are more open-minded, curious, and novelty-seeking relative to the

conservatives, who are more orderly and conventional (e.g., Carney et al., 2008). For instance, liberals (Democrats) exhibit greater openness to innovation, while conservatives (Republicans) often prefer an established method (e.g., Blank & Shaw, 2015; Cookson et al., 2020).

Such differences have implications for algorithm aversion in financial forecasting—if political ideology influences attitudes toward technological automation, Democrats might demonstrate relatively lower algorithm aversion. Therefore, we study the following hypothesis:

Hypothesis 5: Individuals with more progressive political beliefs, such as Democrats, update their return beliefs more toward the signal when receiving AI-generated forecasts.

4.2.3. Empirical results

To test the impact of non-cognitive factors on return updating, we run the following model:

$$\begin{aligned} Update_i = {} & \beta_0 + \beta_1 \cdot Treatment_i \; x \; Perception\ gap_i \; x \; NCog_i + \beta_2 \\ & \cdot Perception\ gap_i \; x \; NCog_i + \beta_3 \cdot Treatment_i \; x \; NCog_i + \beta_4 \\ & \cdot Treatment_i \; x \; Perception\ gap_i + \beta_5 \cdot Treatment_i \\ & + \beta_6 \cdot Perception\ gap_i + \beta_7 \cdot NCog_i + \gamma X_i + \varepsilon_i \end{aligned} \tag{4}$$

where $Update_i$ is the difference between posterior and prior return expectations for individual i. Since we are interested in the impact of receiving a forecast from AI versus (human) financial analysts, $Treatment_i$ is a binary variable equal to one for individuals receiving the treatments *Machine* or *Human + Machine*, while those receiving the *Human* message are the control group. $NCog_i$ is a variable that is either *Female* or *Democrat*, as defined above.

[INSERT TABLE 6]

Table 6 reports the findings for the perception gap interactions with the AI forecast and non-cognitive factors. The coefficients on the triple interaction term *Perception gap* × *AI* × *Female* are positive but statistically insignificant, implying that female investors do not significantly differ from male investors in updating their return belief toward the AI-generated forecasts. For Democrats, the interaction coefficient *Perception gap* × *AI* × *Democrats* equals 0.143 in column 2 (significant at the 1% level) and 0.131 in column 3 (significant at the 1% level). Regarding economic magnitudes, Democratic investors updated their return beliefs with 0.47 percentage

points relatively more toward AI forecasts, signalling an economic meaningful effect. Hence, this points to algorithm appreciation.[13] As such, we find strong support for Hypothesis 5.

Result 5. Political affiliation affects algorithm aversion, with Democrats updating their return beliefs relatively more toward the new signal when receiving AI forecasts, consistent with algorithm appreciation.

We contribute to the growing literature on gender differences in advice-taking and technology adoption. While Brenner and Meyll (2020) document women's preference for human financial advisors over robo-advisors, Scherer and Lehner (2025) do not find any significant association between gender and the likelihood of using robo-advisory services. Our results extend this understanding to return forecast domains, showing more complex patterns of algorithm appreciation. The significant positive interaction between females and purely AI-generated forecasts, coupled with the non-significant response to human-AI combinations, suggests that the presentation and framing of algorithmic advice critically influences its reception.

Our results present an interesting counterpoint to Chen et al. (2023), who show a lower fintech adoption rate among women. Indeed, our findings indicate that when AI-generated advice is clearly delineated as such, without human involvement, female investors demonstrate higher receptivity.

5. Robustness

To validate the robustness of our findings, we employ two approaches. First, we introduce an alternative outcome measure—Expected Sharpe Ratio (ESR)—that incorporates investors' risk perceptions alongside return expectations. Second, we conduct two experiments examining: (i) provider familiarity by replacing "AI" with "ChatGPT", and (ii) how provider revelation timing affects credibility judgments. These checks confirm our core finding that investors process financial information differently based on its source.

5.1. Alternative return expectation measure

[13] Table A.6 shows that the triple interaction for Democrats is significant in both the Machine-only (0.162, 1% level) and Human + Machine (0.126, 5% level) treatments, indicating receptiveness to AI-generated forecasts regardless of framing.

We employ the following alternative specification to assess the robustness of the findings:

$$ESR_i = \beta_0 + \beta_1 \cdot Treatment_i \, x \, Perception\ gap_i + \beta_2 \cdot Treatment_i + \beta_3 \cdot Perception\ gap_i + \gamma X_i + \varepsilon_i \quad (5)$$

where ESR_i is a ratio between the *posterior return* point estimate and risk perception measure for individual i. The risk perception measure is retrieved through the following question, as in Borsboom and Zeisberger (2020) and Holzmeister et al. (2024):

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

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 5.

A negative β_1 indicates that participants – whose prior beliefs deviate more strongly from AI's forecast – experience a lower Sharpe ratio under the AI treatment than in the control condition. Hence, this suggests a lower willingness to integrate the AI's signal into posterior expectations among participants facing larger gaps between the AI projection and their initial expectation.[14]

[INSERT TABLE 7]

Table 7 presents the results using Expected Sharpe Ratio (ESR) as the outcome. Consistent with our conclusions, we observe a negative relationship of AI forecasts on investors' risk-adjusted return expectations. More specifically, the AI treatment coefficient is negative and statistically significant at the 1% level across columns 1-3, indicating that algorithm aversion persists even when accounting for risk perceptions.[15] Overall, these results reinforce our key conclusion that investors process the AI-generated forecasts differently than human forecasts, with algorithm aversion evident even when measuring risk-adjusted return expectations.

5.2. Additional experiments

[14] For instance, an investor who initially expects low returns (e.g., 0.5%) will, in the presence of algorithm aversion, have a lower posterior risk-adjusted return (i.e., expected Sharpe ratio) after receiving the AI's 5% forecast than an investor with the same expectation who receives the forecast from humans. This occurs because, under algorithm aversion, the forecast leads to a higher Sharpe ratio for individuals in the control group (a human forecast) than for those in the treatment groups, despite the same initial belief in low expected returns.

[15] When examining the specific treatment effects in Table A.7, we find that both the *Machine*-only treatment and the *Human + Machine* treatment have strong and statistically significant negative effects.

We conduct two new experiments to address any potential concerns with our main conclusion and strengthen its validity. This includes provider familiarity, and provider revelation timing. For all experiments, we recruit 600 participants who had not participated in Experiment 1 on September 2, 2024. This corresponds to 1,200 new individuals.[16]

5.2.1. Experiment 2: provider familiarity

A concern relates to the comparison between a specific human source (Goldman Sachs) and a generic AI system. To address this, we replace the "AI" provider with "ChatGPT", reflecting the growing prominence of large language models (J. Chen et al., 2024; Jha et al., 2024; Lopez-Lira & Tang, 2025) and to directly compare two named entities (Goldman Sachs and ChatGPT).

[INSERT TABLE 8]

Table 8 presents these results. Overall, this largely mirrors our main findings: investors update return beliefs in response to ChatGPT-generated return forecasts differently than to forecasts generated solely by humans. The direction of the updating is consistent with the earlier results. In particular, when individuals receive forecasts incorporating *AI*, they adjust their beliefs less in the direction of the signal (significant at the 10% level), and this result holds regardless of whether we use *Update* (Table 8, Column 1) or *Expected Sharpe Ratio* (Table 8, Column 2) as the dependent variable.[17]

Since reliable stock return expectations are essential for our analysis, we also followed Merkoulova & Veld (2022) in this experiment and excluded participants who reported annual stock market return estimates exceeding 30% in absolute terms, classifying them as stock return-ignorant. That said, once again, our results likely represent a lower bound, as including the individuals leads to stronger algorithm aversion effects—both in magnitude and statistical significance—as shown in Table A.9, similar to what we observed in our first experiment.

There are two takeaways. First, it highlights that there is a negative perception of ChatGPT's capabilities, which may be due to its declining performance shown in other papers (L. Chen et

[16] As in Experiment 1, we used Prolific to recruit participants with investment experience in exchange-traded funds or stocks at least 18 years of age. We paid participants £9 per hour, next to an additional €50 for five participants. We reject 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 (per experiment). We add the summary statistics of these participants in Table A.10.

[17] When separating the AI treatments, the Machine-only coefficient exhibits the strongest magnitude (Table A.8).

al., 2024). Second, and more importantly, the results indicate that our conclusion of algorithm aversion is not driven by differences in source specificity or familiarity but reflects differences in how investors process information from human versus algorithmic sources.

5.2.2. Experiment 3: source revelation timing

To address the potential framing effects, we change the timing when the source information is revealed. In this experiment, we first provided a generic forecast without revealing its source. Specifically,

Assume you received the following message: "The S&P 500 12-month index return is forecasted at 5%. The baseline assumption during the next year is that the US economy will continue to expand at a modest pace and avoid a recession".

Then, we ask about the pariticpant's posterior S&P500 return expectation and only afterwards reveal the source. After this revelation, we question investors on source credibility.

This sequencing allows us to study how source attribution affects credibility judgments when the information content is held constant across conditions. By making the source information more salient after initial belief formation, we can better isolate the impact of source attribution on perceived credibility.

[INSERT FIGURE 3]

Figure 3 summarizes the results across the three experiments. A clear trend emerges: investors consistently find human analysts (*Human*) more credible than the forecasts involving *AI*. The difference is statistically significant across all experiments (Subfigure 3.1). Moreover, across experiments, *Human* is perceived as the most credible source, followed by *Human + Machine*, and lastly *Machine* (Subfigure 3.2). Even in an (increasingly) AI-driven financial landscape, this preference implies that the human component in financial advice remains crucial. We use the regression specification of Equation 2, with *Credibility* as the dependent variable, to explore this more.

[INSERT TABLE 9]

Table 9 presents the findings. Columns 1 shows the results using the sample from Experiment 2, where ChatGPT was used instead of a generic AI forecast, while Columns 2 use the sample from Experiment 3, "the reframing experiment". Regardless of the experiment, the results are similar to those in Table 4. *AI*-generated forecasts are perceived as significantly less credible.[18] These findings further support our conclusion that the forecast source significantly impacts an investor's trust, independent of information content.

6. Conclusion

Our study provides novel insights into the complex relationship between artificial intelligence, financial analyst forecasts, and investor trust. Through three 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. In other words, there is evidence of algorithm aversion.

Beyond documenting the existence of algorithm aversion, we highlight that both cognitive and non-cognitive factors significantly influence how investors respond to AI-generated financial advice. Market participants with higher AI literacy and investment sophistication demonstrate reduced algorithm aversion, while political affiliation emerged as a main non-cognitive factor, with Democratic investors showing greater receptiveness to AI forecasts. In other words, there is evidence of algorithm appreciation for specific groups.

The findings have the following implications for financial institutions and policymakers. First, the persistent algorithm aversion among specific investor groups indicates that increasing AI literacy has to be a main priority, particularly as financial markets keep integrating algorithmic decision-making.[19] Second, differences in AI receptivity between political affiliations indicate that financial advice communication strategies may need to be tailored to diverse audiences, acknowledging that factors beyond just understanding AI's capabilities shape predispositions toward technology.

---

[18] Table A.11 also shows that the *Machine*-only treatment exhibits the lowest credibility in regression results.

[19] "*Morgan Stanley to cut 2,000 employees as AI reshapes Wall Street*" (Investment News, 2025); "*J.P. Morgan Chase is giving its employees AI assistance powered by ChatGPT maker OpenAI*" (CNBC, 2024); and "*Goldman Sachs rolls out an AI assistant for its employees as artificial intelligence sweeps Wall Street*"(CNBC, 2024).

References

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.

Barber, B. M., Huang, X., & Odean, T. (2016). Which Factors Matter to Investors? Evidence from Mutual Fund Flows. *Review of Financial Studies*, *29*(10), 2601–2642.

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

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).

Bertomeu, J., Lin, Y., Liu, Y., & Ni, Z. (2025). The impact of generative AI on information processing: Evidence from the ban of ChatGPT in Italy. *Journal of Accounting and Economics*, *80*(1), 101782.

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

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.

Bonaparte, Y., Kumar, A., & Page, J. K. (2017). Political climate, optimism, and investment decisions. *Journal of Financial Markets*, *34*, 69–94.

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.

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).

Cao, Y., & Zhang, W. (2025). How AI is shaping accounting and finance. *The British Accounting Review*, *January*, 101650. https://doi.org/10.1016/j.bar.2025.101650

Carney, D. R., Jost, J. T., Gosling, S. D., & Potter, J. (2008). The Secret Lives of Liberals and Conservatives: Personality Profiles , Interaction Styles, and the Things They Leave Behind. *Political Psychology*, *29*(6), 807–840.

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, *56*(5), 809–825.

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

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

Chen, S., Doerr, S., Frost, J., Gambacorta, L., & Song, H. (2023). The fintech gender gap. *Journal of Financial Intermediation*, *54*(April 2022), 101026.

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.

Dang, J., & Liu, L. (2026). Public trust in artificial intelligence users. *Current Opinion in Psychology*, *66*(August 2025), 102148.

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

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.

Dumitrescu, A., & Gil-bazo, J. (2018). Market frictions, investor sophistication, and persistence in mutual fund performance. *Journal of Financial Markets*, *40*, 40–59.

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.

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

Giglio, S., Maggiori, M., Stroebel, J., & Utkus, S. (2021). Five Facts about Beliefs and Portfolios. *American Economic Review*, *111*(5), 1481–1522.

Glikson, E., & Williams Woolley, A. (2020). Human Trust in Artificial Intelligence: Review of

Empirical Research. *Academy of Management*, *14*(2), 627–660.

Goetzmann, W. N., & Kumar, A. (2008). Equity Portfolio Diversification. *Review of Finance*, *12*(March), 433–463. https://doi.org/10.1093/rof/rfn005

Goetzmann, W. N., & Massa, M. (2005). Dispersion of opinion and stock returns. *Journal of Financial Markets*, *8*, 325–350.

Greig, F., Rossi, A. G., Utkus, S. P., & Walther, A. (2025). *Human Financial Advice in the Age of Automation*.

Haaland, I., Roth, C., & Wohlfart, J. (2023). Designing Information Provision Experiments. *Journal of Economic Literature*, *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.

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

Jiang, F., Ma, T., & Zhu, F. (2024). Fundamental characteristics, machine learning, and stock price. *Journal of Financial Markets*, *69*(April), 100908.

Johnson, W. C., & Markelevich, A. (2025). Should Accountants Be Afraid of AI? Risks and Opportunities of Incorporating Artificial Intelligence into Accounting and Auditing. *Accounting Horizons*, *39*(2), 117–123.

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.

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).

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

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. (2025). *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.

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.

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

Pungaliya, R. S., & Wang, Y. (2023). Machine invasion: Automation in information acquisition and the cross-section of stock returns. *Journal of Financial Markets*, *64*(October 2022), 100788. https://doi.org/10.1016/j.finmar.2022.100788

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

Scherer, B., & Lehner, S. (2025). What drives robo-advice? *Journal of Empirical Finance*, *80*, 101574.

Tsz, T., Tse, K., Hanaki, N., & Mao, B. (2024). Beware the performance of an algorithm before relying on it: Evidence from a stock price forecasting experiment. *Journal of Economic Psychology*, *102*, 102727.

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.

Walther, T., & Dutordoir, M. (2025). *Certainly! Generative AI and its Impact on Academic Writing (in Finance)*.

Figures

Figure 1. Incentivized question (prior)

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

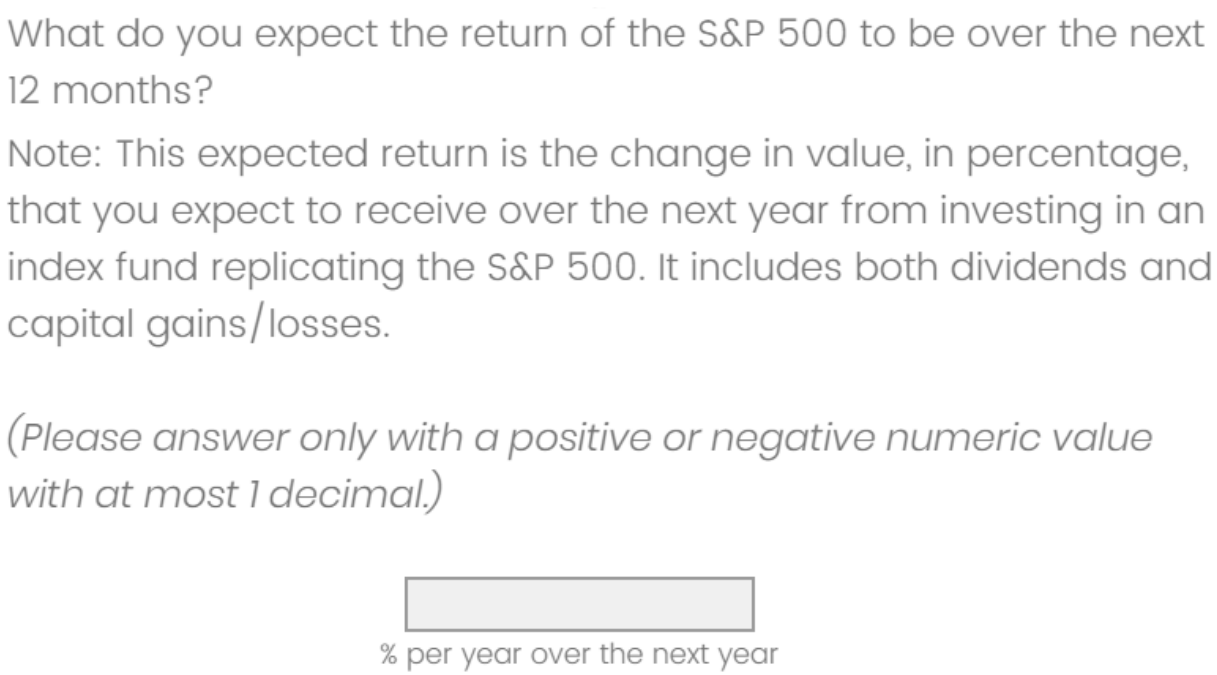
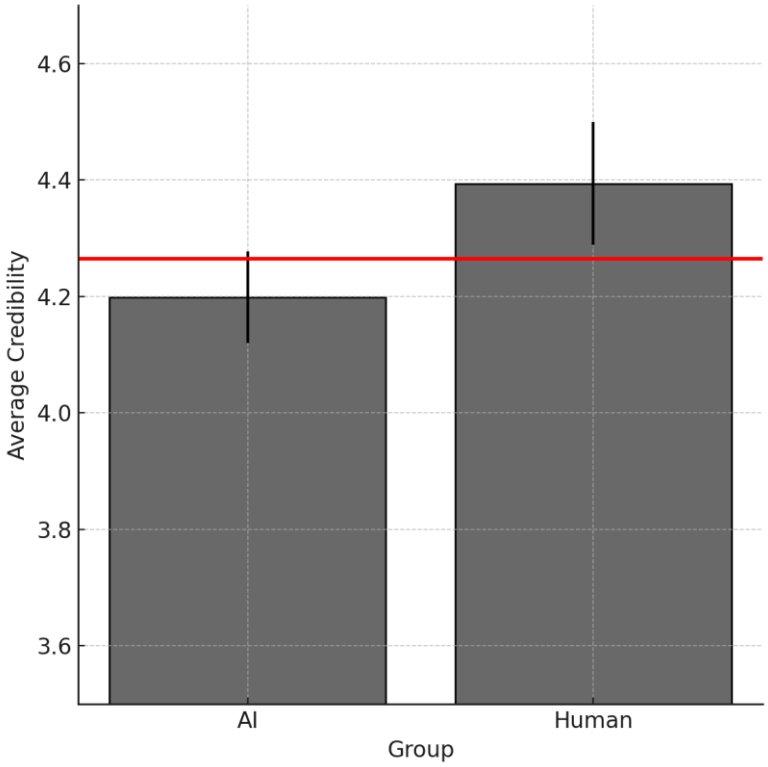


Figure 2. Average credibility

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.

Panel A



Panel B

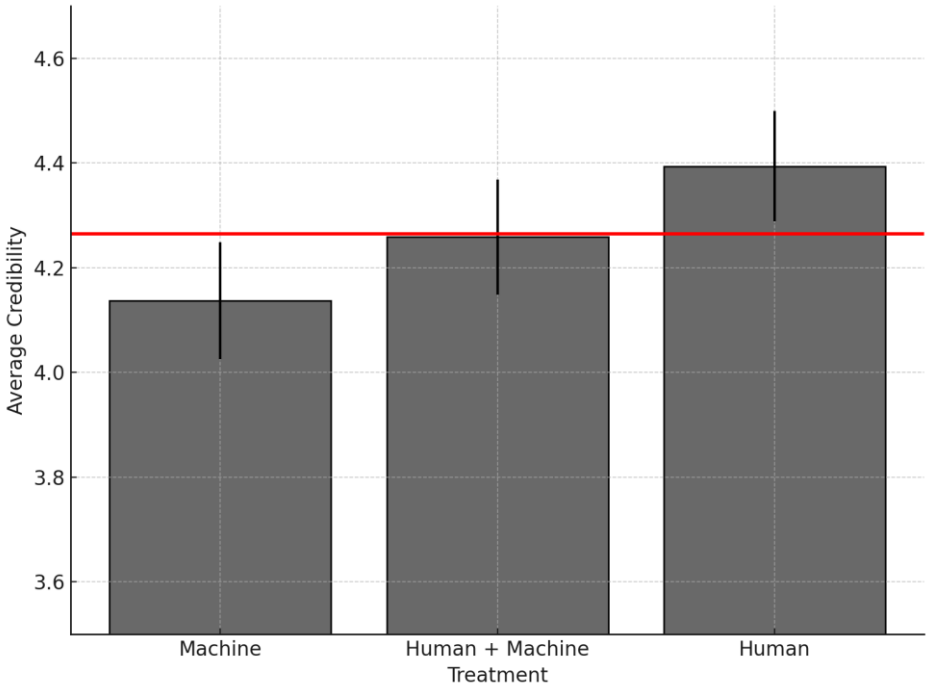
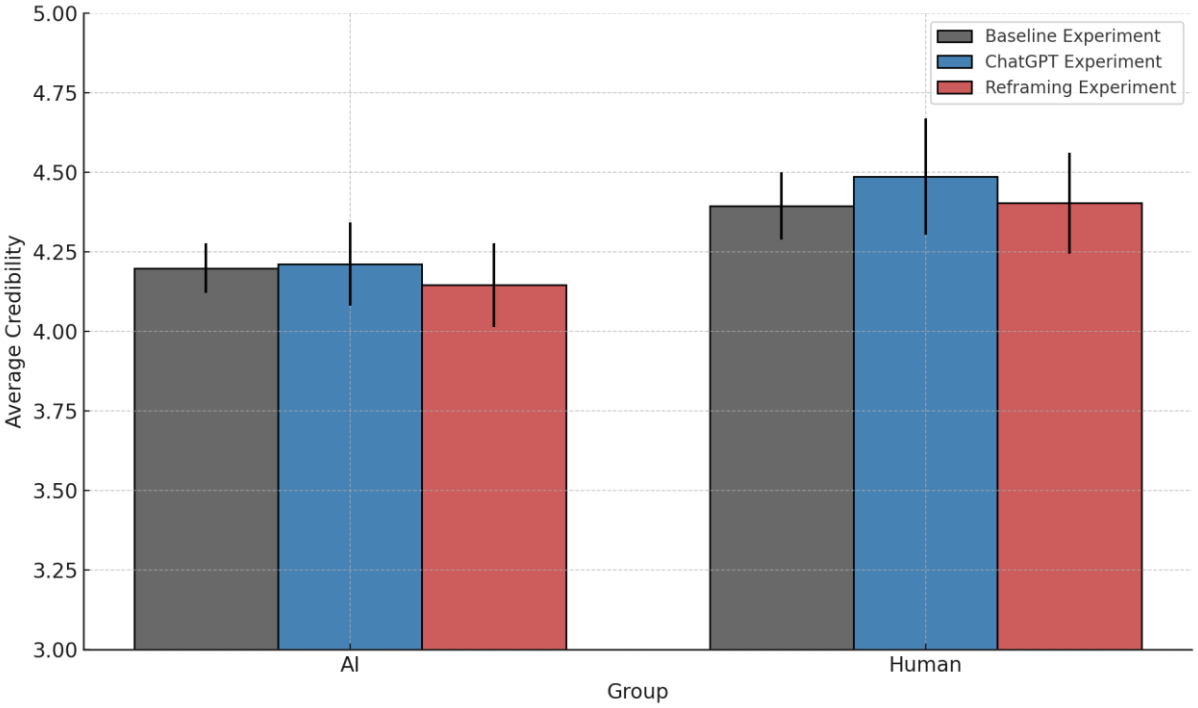


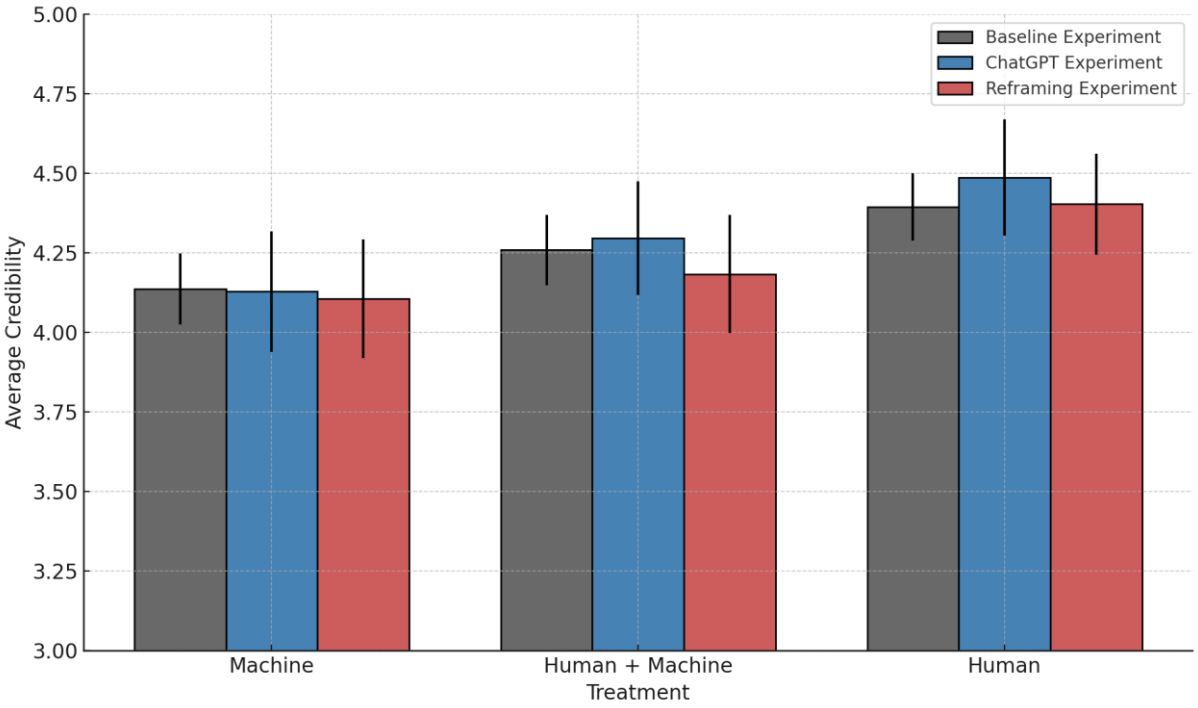
Figure 3. Summary of the results: Credibility

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

Subfigure 3.1



Subfigure 3.2



Tables

Table 1. Summary statistics

This table reports summary statistics for variables collected before participants received the treatment. 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). Panel B tests the balance of means across the control and treatment groups, as well as the p-values between the control and treatment groups (Columns 5, 6 and 7). Belief variables refer to prior elicited before the information manipulation. The return expectation prior is incentivized. We drop all the participants who exhibit return expectations below -30% and above 30%.

Panel A: Summary statistics

	1.	2.	3.	4.	5.
	Mean	SD	p10	p50	p90
Female	0.444	0.497	0	0	1
Age	41.451	12.729	26	40	60
Risk preference (1-7)	4.081	1.411	2	4	6
Financial literacy (1-5)	4.231	0.871	3	4	5
Sophistication (0-1)	0.154	0.361	0	0	1
Democrat (0-1)	0.414	0.492	0	0	1
AI literacy (1-5)	4.070	1.068	3	4	5
ChatGPT usefulness (1-7)	4.356	1.832	2	5	7
Return expectation (prior)	8.471	7.281	3	8	18
Confidence (1-7)	4.893	1.435	3	5	7
AI relative capabilities (1-7)	4.581	1.198	3	5	6

Panel B: Treatment differences

	1.	2.	3.	4.	5.	6.	7.
	Human	*AI*	*Human + Machine*	*Machine*	p-value (2 = 1)	p-value (3 = 1)	p-value (4 = 1)
Female	0.464	0.433	0.426	0.441	0.229	0.196	0.436
Age	41.053	41.656	41.926	41.380	0.359	0.251	0.668
Risk preference (1-7)	4.045	4.099	4.059	4.141	0.460	0.873	0.266
Financial literacy (1-5)	4.238	4.227	4.248	4.206	0.811	0.848	0.548
Sophistication (0-1)	0.137	0.163	0.162	0.164	0.172	0.259	0.208
Democrat (0-1)	0.432	0.404	0.397	0.412	0.269	0.229	0.483
AI literacy (1-5)	4.077	4.066	4.031	4.102	0.835	0.464	0.696
ChatGPT usefulness (1-7)	4.459	4.303	4.277	4.329	0.099	0.097	0.234
Return expectation (prior)	8.468	8.472	8.338	8.610	0.990	0.767	0.740
Confidence (1-7)	4.961	4.859	4.826	4.893	0.170	0.109	0.433
AI relative capabilities (1-7)	4.582	4.581	4.611	4.551	0.988	0.689	0.667

Table 2. Baseline result

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) 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 treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). The control variables are defined in Table A.1. Fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.	4.
$Perc.\,gap\ x\ I_{AI}$	-0.042*	-0.042*	-0.040*	-0.040*
	(0.073)	(0.078)	(0.087)	(0.094)
$Perc.\,gap$	0.524***	0.520***	0.517***	0.515***
	(0.000)	(0.000)	(0.000)	(0.000)
I_{AI}	0.155	0.140	0.163	0.163
	(0.414)	(0.462)	(0.391)	(0.390)
Female		-0.272	-0.069	-0.056
		(0.103)	(0.690)	(0.748)
Age		-0.001	-0.006	-0.006
		(0.851)	(0.400)	(0.374)
Risk preference			0.070	0.085
			(0.238)	(0.162)
Financial literacy			0.470***	0.444***
			(0.000)	(0.000)
Sophistication			0.137	0.137
			(0.557)	(0.555)
Democrat			0.042	0.038
			(0.802)	(0.822)
AI literacy				0.114
				(0.148)
ChatGPT usefulness				-0.022
				(0.621)
Constant	-0.340**	-0.722	-3.011***	-3.314***
	(0.028)	(0.323)	(0.001)	(0.000)
R-squared	0.543	0.543	0.549	0.550
Income FE	No	Yes	Yes	Yes
Education FE	No	Yes	Yes	Yes
Observations	1,668	1,668	1,668	1,668

Table 3. The role of credibility in belief updating

This table shows the coefficients of *Update* (the difference between posterior and prior expectation for the 12-month-ahead S&P500 index return) on a set of independent variables. *Credibility* is measured as posterior perceptions. The perception gap (*Perc. gap*) is the difference between the analyst forecast and prior expectations about the 12-month-ahead S&P500 returns. The control variables are defined in Table A.1. Fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.	4.
$Perc.gap\ x\ Credibility$	0.137***	0.137***	0.135***	0.135***
	(0.000)	(0.000)	(0.000)	(0.000)
$Credibility$	0.182***	0.181***	0.162***	0.174***
	(0.003)	(0.003)	(0.008)	(0.006)
$Perc.gap$	-0.022	-0.023	-0.018	-0.020
	(0.452)	(0.452)	(0.550)	(0.504)
Female		-0.293*	-0.097	-0.085
		(0.053)	(0.536)	(0.591)
Age		-0.004	-0.007	-0.007
		(0.465)	(0.248)	(0.226)
Risk preference			0.137**	0.150***
			(0.012)	(0.007)
Financial literacy			0.334***	0.310***
			(0.000)	(0.001)
Sophistication			0.131	0.133
			(0.534)	(0.529)
Democrat			-0.013	-0.018
			(0.935)	(0.907)
AI literacy				0.106
				(0.139)
ChatGPT usefulness				-0.023
				(0.591)
Constant	-0.824***	-0.873	-2.817***	-3.126***
	(0.002)	(0.205)	(0.000)	(0.000)
R-squared	0.623	0.623	0.627	0.628
Income FE	No	Yes	Yes	Yes
Education FE	No	Yes	Yes	Yes
Observations	1,668	1,668	1,668	1,668

Table 4. Credibility of the treatments

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. The treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). All other control variables are defined in Table A.1. Fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3
I_{AI}	-0.195***	-0.191***	-0.169**
	(0.004)	(0.005)	(0.012)
$Perc.gap$	0.006	0.005	0.008*
	(0.175)	(0.242)	(0.056)
Female		-0.044	-0.012
		(0.503)	(0.854)
Age		0.003	0.005*
		(0.197)	(0.054)
AI literacy			-0.063**
			(0.039)
Risk preference			0.058**
			(0.014)
Financial literacy			-0.002
			(0.967)
Sophistication			-0.078
			(0.388)
Democrat			0.123*
			(0.062)
ChatGPT usefulness			0.134***
			(0.000)
Constant	4.415***	3.700***	2.998***
	(0.000)	(0.000)	(0.000)
R-squared	0.005	0.002	0.046
Income FE	No	Yes	Yes
Education FE	No	Yes	Yes
Exclude	No	No	No
Observations	1,668	1,668	1,668

Table 5. Cognitive factors

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) 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 treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). *Sophistication* equals one if the individual has experience in options, private equity, angel investing, or venture capital. Control variables are defined in Table A.1. *P*-values are in parentheses. Fixed effects include education and income levels. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.
$Perc.gap\ x\ I_{AI}\ x\ AI\ literacy$	0.059***		0.043**
	(0.007)		(0.049)
$Perc.gap\ x\ I_{AI}\ x$ Sophistication		0.205***	0.167***
		(0.001)	(0.007)
$Perc.gap\ x\ I_{AI}$	-0.283***	-0.077***	-0.251***
	(0.002)	(0.003)	(0.006)
$Perc.gap\ x\ AI\ literacy$	-0.089***		-0.073***
	(0.000)		(0.000)
$Perc.gap\ x$ Sophistication		-0.276***	-0.235***
		(0.000)	(0.000)
$I_{AI}\ x\ AI\ literacy$	0.478***		0.415**
	(0.009)		(0.022)
$I_{AI}\ x$ Sophistication		1.240**	1.146**
		(0.014)	(0.022)
I_{AI}	-1.864**	-0.106	-1.836**
	(0.017)	(0.607)	(0.018)
$AI\ literacy$	-0.404***		-0.342**
	(0.007)		(0.022)
Sophistication		-1.041**	-0.957**
		(0.013)	(0.022)
$Perc.gap$	0.882***	0.565***	0.859***
	(0.000)	(0.000)	(0.000)
Constant	-1.142	-2.017***	-0.463
	(0.283)	(0.000)	(0.576)
R-squared	0.558	0.558	0.556
Controls	Yes	Yes	Yes
Income FE	Yes	Yes	Yes
Education FE	Yes	Yes	Yes
Observations	1,668	1,668	1,663

Table 6. Non-cognitive factors

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) 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 treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). Control variables are defined in Table A.1. *P*-values are in parentheses. Fixed effects include education and income levels. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.
$Perc.gap\ x\ I_{AI}\ x\ Female$	0.029		0.016
	(0.527)		(0.716)
$Perc.gap\ x\ I_{AI}\ x\ Democrat$		0.143***	0.131***
		(0.004)	(0.008)
$Perc.gap\ x\ I_{AI}$	-0.044	-0.088***	-0.083**
	(0.165)	(0.002)	(0.017)
$Perc.gap\ x\ Female$	0.168***		0.166***
	(0.000)		(0.000)
$Perc.gap\ x\ Democrat$		0.027	0.017
		(0.503)	(0.663)
$I_{AI}\ x\ Female$	-0.348		-0.339
	(0.357)		(0.369)
$I_{AI}\ x\ Democrat$		-0.024	-0.034
		(0.951)	(0.929)
I_{AI}	0.315	0.254	0.351
	(0.202)	(0.290)	(0.205)
$Female$	0.835***		0.764**
	(0.007)		(0.014)
$Democrat$		0.512***	0.437
		(0.107)	(0.162)
$Perc.gap$	0.424***	0.507***	0.432***
	(0.000)	(0.000)	(0.000)
Constant	-3.767***	-3.259***	-3.328***
	(0.000)	(0.000)	(0.000)
R-squared	0.569	0.556	0.576
Controls	Yes	Yes	Yes
Income FE	Yes	Yes	Yes
Education FE	Yes	Yes	Yes
Observations	1,668	1,668	1,668

Table 7. Expected Sharpe Ratio

This table shows the coefficients of *Expected Sharpe Ratio* (defined as the ratio of posterior to perceived riskiness) 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 treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). Control variables are defined in Table A.1. *P*-values are in parentheses. Fixed effects include education and income levels. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.
$Perc.\,gap\ x\ I_{AI}$	-0.040***	-0.040***	-0.040***
	(0.004)	(0.004)	(0.003)
$Perc.\,gap$	-0.144***	-0.146***	-0.148***
	(0.000)	(0.000)	(0.000)
I_{AI}	-0.033	-0.042	-0.035
	(0.764)	(0.700)	(0.744)
Female		-0.203**	-0.182*
		(0.034)	(0.067)
Age		-0.006	-0.009**
		(0.112)	(0.014)
AI literacy			0.076*
			(0.091)
Risk preference			-0.087**
			(0.012)
Financial literacy			0.212***
			(0.000)
Sophistication			0.044
			(0.738)
Democrat			0.091
			(0.345)
ChatGPT usefulness			-0.054**
			(0.037)
Constant	1.488***	1.805***	1.262**
	(0.000)	(0.000)	(0.017)
R-squared	0.301	0.306	0.318
Income FE	No	Yes	Yes
Education FE	No	Yes	Yes
Exclude	No	No	No
Observations	1,668	1,668	1,668

Table 8. ChatGPT

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) and *Expected Sharpe Ratio* (defined as the ratio of posterior to perceived riskiness) on a set of independent variables. The perception gap (*Perc. gap*) is the difference between analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. The treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). Control variables are defined in Table A.1. 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.
$Perc.gap\ x\ I_{AI}$	-0.081*	-0.036*
	(0.095)	(0.075)
I_{AI}	-0.063	0.020
	(0.863)	(0.896)
$Perc.gap$	0.556***	-0.110***
	(0.000)	(0.000)
Constant	-0.997	1.391**
	(0.506)	(0.026)
R-squared	0.464	0.297
Controls	Yes	Yes
Income FE	Yes	Yes
Education FE	Yes	Yes
Observations	558	558

Table 9. Reframing

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 treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). Control variables are defined in Table A.1. *P*-values are in parentheses. Fixed effects include education and income levels. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.
I_{AI}	-0.240**	-0.241**
	(0.036)	(0.029)
Constant	3.050***	4.222***
	(0.000)	(0.000)
R-squared	0.083	0.059
Controls	Yes	Yes
Education FE	Yes	Yes
Income FE	Yes	Yes
Observations	558	557
Experiment	ChatGPT	Reframing

Appendix I: Questionnaire

Financial literacy

On a scale from 1 to 7, where 1 means very low and 7 means very high, how would you assess your overall financial knowledge? (On a scale from 1-7; 1, not at all useful; 7, extremely useful)

(item from van Rooij et al., 2011)

Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow? Please select one.

- More than $102
- Exactly $102
- Less than $102
- Don't know
- Prefer not to say

Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account? Please select one.

- More than today
- Exactly the same as today
- Less than today
- Don't know
- Prefer not to say

Buying a single company's stock usually provides a safer return than a stock mutual fund. Please select one.

- True
- False
- Don't know

- Prefer not to say

A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less. Please select one.

- True
- False
- Don't know
- Prefer not to say

If interest rates fall, what should happen to bond prices? Please select one.

- They will rise
- They will fall
- They will stay the same
- There is no relationship between bond prices and the interest rate
- Don't know
- Prefer not to say

(all items from van Rooij et al., 2011)

AI literacy

Artificial intelligence (AI) is designed to learn tasks that humans typically do. How much have you heard or read about AI? (On a scale from 1-7; 1, not at all useful; 7, extremely useful)

Thinking about customer service, which of the following uses artificial intelligence (AI)? Please select one.

- A chatbot that immediately answers customer questions
- An online survey sent to customers that allows them to provide feedback
- A contact page with a form available to customers to provide feedback
- A detailed Frequently Asked Questions webpage
- Not sure
- No answer

Which of the following uses artificial intelligence (AI) when playing music? Please select one.

- Using Bluetooth to connect to wireless speakers
- A playlist recommendation
- A wireless internet connection to stream music
- Shuffle play from a chosen playlist
- Not sure
- No answer

Which of the following uses artificial intelligence (AI) when using email? Please select one.

- The email service marking an email as read after the user opens it
- The email service allowing the user to schedule an email to send at a specific time in the future
- The email service categorizing an email as spam
- The email service sorting emails by time and date
- Not sure
- No answer

Thinking about health products, which of the following uses artificial intelligence (AI)? Please select one.

- Wearable fitness trackers that analyze exercise and sleeping patterns
- Thermometers that are placed under someone's tongue to detect a fever
- At-home COVID-19 tests
- Pulse oximeters that measure a person's oxygen level of the blood
- Not sure
- No answer

Thinking about devices in the home, which of the following uses artificial intelligence (AI)? Please select one.

- Programming a home thermostat to change temperatures at certain times
- A security camera that sends an alert when there is an unrecognized person at the door

- Programming a timer to control when lights in a home turn on and off
- An indicator light that turns red when a water filter needs to be replaced
- Not sure
- No answer

(all items from PewResearch)

Appendix II: Additional information

Figure A.1. ChatGPT's forecast estimate for the S&P 500

This figure illustrates an example of ChatGPT forecasting returns for the S&P 500 over the next 12 months.

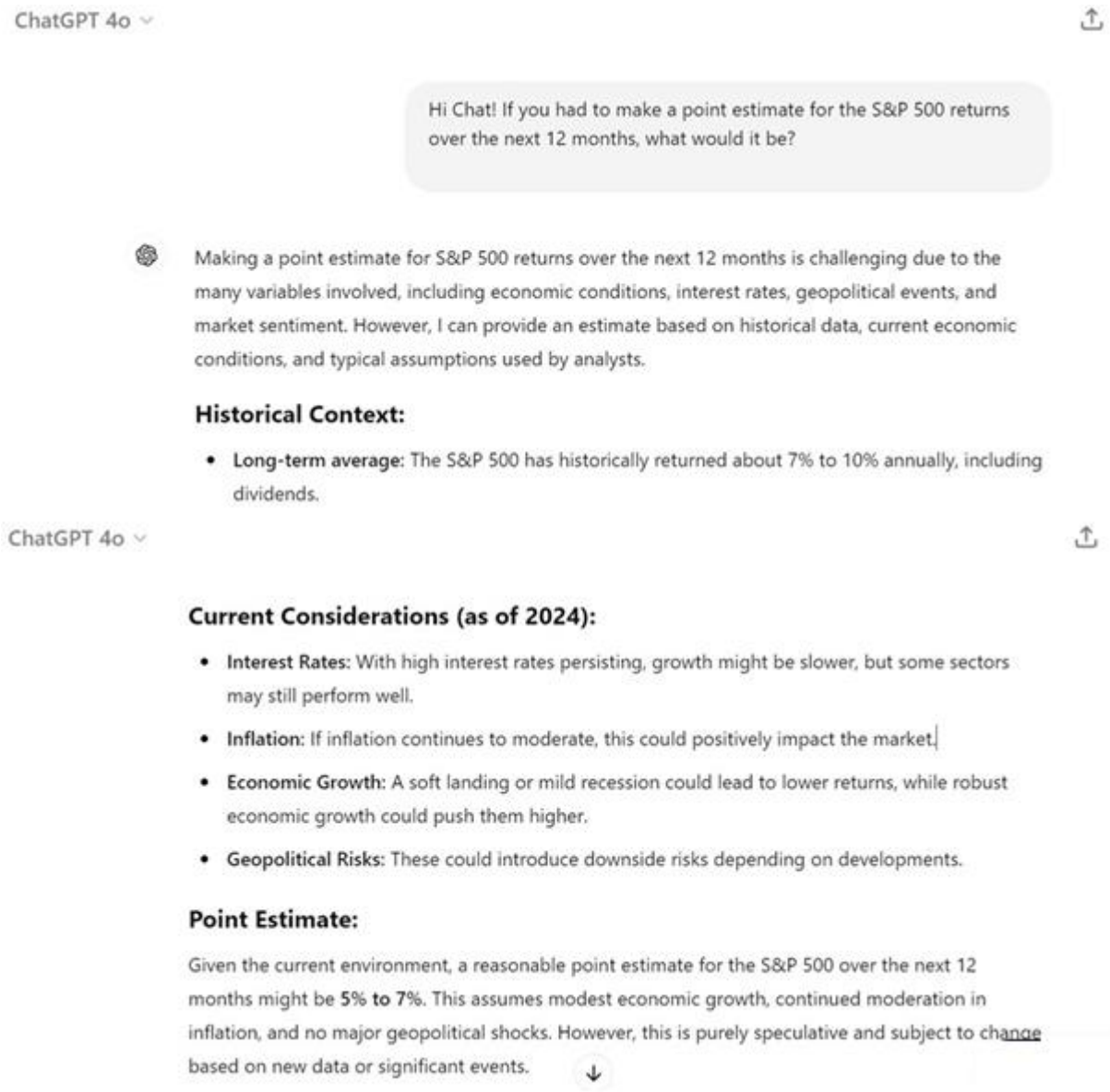


Table A.1. Variable definitions

This table reports the variable definitions and the experiment in which we obtained said variable.

Variable Name	Definition	Experiment
Dependent variables		
Update	The difference in return expectations before and after receiving the forecast	1,2
Credibility (1-7)	How credible do you think the information containing the 5% forecast you previously received is?	1,2,3
Expected Sharpe Ratio	The ratio between the posterior return point estimate and the posterior risk perception measure for the S&P 500	1,2
Key independent variables		
I_{AI} (0-1)	An indicator variable when the participant is randomly assigned to receive a forecast involving AI	1,2,3
I_{HM} (0-1)	An indicator variable when the participant is randomly assigned to receive the *Human + Machine* treatment	1,2,3
I_{M} (0-1)	An indicator variable when the participant is randomly assigned to receive the *Machine* treatment.	1,2,3
Other variables		
Age	Age of the individual (in years)	1,2,3
AI literacy (1-5)	Number of correct answers to AI literacy questions	1,2,3
ChatGPT usefulness (1-7)	In general, how useful has ChatGPT (or similar artificial intelligence tools) been to you?	1,2,3
Confidence (prior) (1-7)	How confident are you with your answer?	1,2
Democrat (0-1)	The individual politically identifies as a Democrat	1
Education	A categorical variable for the highest level of education completed, ranging from secondary education to Ph.D.	1,2,3
Female (0-1)	The individual is a female	1,2,3
Financial literacy (1-5)	Number of correct answers to financial literacy questions	1,2,3
Income	A categorical variable for annual household income (USD), ranging from less than $10,000 to over $150,000.	1,2,3
Return expectation (prior)	What do you expect the return of the S&P 500 to be over the next 12 months?	1,2,4
Risk preference (1-7)	Generally speaking, are you the kind of person who is willing to take risks or who prefers to avoid risks?	1,2,3
Sophistication (0-1)	The individual has experience with options, private equity, venture capital, or angel investing.	1,2,3

Table A.2. Baseline results: disentangling the AI treatment

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) 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 treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). The treatment variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_{M} equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. Column 1 includes individuals receiving all treatments. Column 2 excludes individuals receiving the *Machine* forecast from the sample, while Column 3 excludes individuals receiving the *Human + Machine* forecast. Column 4 includes the same individuals as Column 1. The control variables are defined in Table A.1. Fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.	4.
$Perc.\,gap\ x\ I_{AI}$	-0.040*			
	(0.094)			
$Perc.\,gap\ x\ I_{HM}$		-0.050*		-0.051*
		(0.073)		(0.056)
$Perc.\,gap\ x\ I_{M}$			-0.028	-0.026
			(0.303)	(0.346)
$Perc.\,gap$	0.515***	0.516***	0.518***	0.515***
	(0.000)	(0.000)	(0.000)	(0.000)
I_{AI}	0.163			
	(0.390)			
I_{HM}		0.187		0.166
		(0.405)		(0.446)
I_{M}			0.151	0.169
			(0.491)	(0.446)
Female	-0.056	-0.059	0.049	-0.052
	(0.748)	(0.786)	(0.817)	(0.765)
Age	-0.006	-0.004	0.003	-0.006
	(0.374)	(0.617)	(0.728)	(0.371)
AI literacy	0.114	0.085	0.076	0.114
	(0.148)	(0.386)	(0.430)	(0.150)
Risk preference	0.085	0.074	0.158**	0.084
	(0.162)	(0.334)	(0.030)	(0.166)
Financial lit.	0.444***	0.331**	0.518***	0.441***
	(0.000)	(0.011)	(0.000)	(0.000)
Sophistication	0.137	0.060	0.004	0.137
	(0.555)	(0.839)	(0.987)	(0.555)
Democrat	0.038	0.094	0.169	0.038
	(0.822)	(0.658)	(0.407)	(0.822)
ChatGPT usefulness	-0.022	-0.041	-0.002	-0.021
	(0.621)	(0.481)	(0.970)	(0.638)
Constant	-3.314***	-3.083***	-3.958***	-3.297***
	(0.000)	(0.007)	(0.000)	(0.000)
R-squared	0.550	0.542	0.564	0.549
Income FE	Yes	Yes	Yes	Yes
Education FE	Yes	Yes	Yes	Yes
Exclude	No	*M*	*HM*	No
Observations	1668	1122	1112	1668

Table A.3. Baseline results including stock return-ignorant investors

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) 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. In Panel A, the treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). In Panel B, the treatment variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_M equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. The control variables are defined in Table A.1. Fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Panel A

	1.	2.	3.	4.
$Perc.\,gap\ x\ I_{AI}$	-0.080***	-0.077***	-0.075***	-0.075***
	(0.001)	(0.001)	(0.002)	(0.002)
$Perc.\,gap$	-1.468	-1.217	-1.141	-1.176
	(0.339)	(0.429)	(0.459)	(0.445)
I_{AI}	0.329***	0.328***	0.322***	0.321***
	(0.000)	(0.000)	(0.000)	(0.000)
Female		-0.961	0.016	0.043
		(0.509)	(0.992)	(0.978)
Age		-0.029	-0.052	-0.056
		(0.621)	(0.377)	(0.344)
Risk preference			-0.059	0.046
			(0.910)	(0.932)
Financial literacy			2.210***	2.115**
			(0.007)	(0.012)
Sophistication			1.908	1.954
			(0.354)	(0.343)
Democrat			0.648	0.685
			(0.662)	(0.645)
AI literacy				0.417
				(0.539)
ChatGPT usefulness				-0.289
				(0.473)
Constant	-0.607	12.757**	3.472	3.202
	(0.627)	(0.046)	(0.648)	(0.690)
R-squared	0.270	0.271	0.273	0.273
Income FE	No	Yes	Yes	Yes
Education FE	No	Yes	Yes	Yes
Observations	1794	1794	1794	1794

(Table A.3 continued)

Panel B

	1.	2.	3.	4.
$Perc.\,gap\ x\ I_{AI}$	-0.075***			
	(0.002)			
$Perc.\,gap\ x\ I_{HM}$		-0.077***		-0.080***
		(0.002)		(0.003)
$Perc.\,gap\ x\ I_{M}$			-0.067**	-0.069**
			(0.014)	(0.012)
$Perc.\,gap$	0.321***	0.319***	0.320***	0.321***
	(0.000)	(0.000)	(0.000)	(0.000)
I_{AI}	-1.176			
	(0.445)			
I_{HM}		-1.261		-1.305
		(0.446)		(0.464)
I_{M}			-0.826	-1.035
			(0.642)	(0.562)
Female	0.043	-1.944	0.524	0.048
	(0.978)	(0.263)	(0.777)	(0.975)
Age	-0.056	-0.033	-0.041	-0.055
	(0.344)	(0.628)	(0.561)	(0.348)
AI literacy	0.417	0.260	0.608	0.405
	(0.539)	(0.735)	(0.467)	(0.552)
Risk preference	0.046	0.095	-0.062	0.039
	(0.932)	(0.877)	(0.923)	(0.943)
Financial lit.	2.115**	2.683***	1.905*	2.115**
	(0.012)	(0.007)	(0.060)	(0.012)
Sophistication	1.954	0.895	1.491	1.959
	(0.343)	(0.709)	(0.554)	(0.342)
Democrat	0.685	0.655	1.564	0.664
	(0.645)	(0.699)	(0.386)	(0.656)
ChatGPT usefulness	-0.289	-0.031	-0.215	-0.280
	(0.473)	(0.947)	(0.664)	(0.487)
Constant	3.202	2.743	8.736	3.267
	(0.690)	(0.762)	(0.367)	(0.684)
R-squared	0.273	0.314	0.284	0.272
Income FE	Yes	Yes	Yes	Yes
Education FE	Yes	Yes	Yes	Yes
Exclude	No	*M*	*HM*	No
Observations	1794	1193	1200	1794

Table A.4. Credibility of the treatments

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. The variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_M equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. All other control variables are defined in Table A.1. Fixed effects include education and income levels. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1	2	3
I_{HM}	-0.114		-0.105
	(0.141)		(0.172)
I_M		-0.224***	-0.233***
		(0.004)	(0.003)
$Perc.gap$	0.007	0.008	0.008*
	(0.190)	(0.142)	(0.058)
Female	-0.094	0.037	-0.011
	(0.254)	(0.657)	(0.869)
Age	0.002	0.004	0.005*
	(0.473)	(0.230)	(0.058)
AI literacy	-0.071*	-0.061	-0.061**
	(0.056)	(0.107)	(0.045)
Risk preference	0.073**	0.057**	0.059**
	(0.012)	(0.046)	(0.012)
Financial literacy	0.054	0.001	-0.003
	(0.271)	(0.988)	(0.933)
Sophistication	-0.079	-0.156	-0.078
	(0.483)	(0.162)	(0.388)
Democrat	0.114	0.073	0.124*
	(0.155)	(0.363)	(0.060)
ChatGPT usefulness	0.107***	0.152***	0.134***
	(0.000)	(0.000)	(0.000)
Constant	2.815***	3.043***	2.987***
	(0.000)	(0.000)	(0.000)
R-squared	0.037	0.055	0.047
Income FE	Yes	Yes	Yes
Education FE	Yes	Yes	Yes
Exclude	*M*	*HM*	No
Observations	1,122	1,112	1,668

Table A.5. Cognitive factors: disentangling the AI treatment

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) 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 treatment variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_M equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. *Sophistication* equals one if the individual has experience in options, private equity, angel investing, or venture capital. Control variables are defined in Table A.1. *P*-values are in parentheses. Fixed effects include education and income levels. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.
$Perc.gap\ x\ I_{HM}\ x\ AI\ literacy$	0.083***		0.066***
	(0.001)		(0.008)
$Perc.gap\ x\ I_M\ x\ AI\ literacy$	0.033		0.024
	(0.190)		(0.341)
$Perc.gap\ x\ I_{HM}\ x\ Sophistication$		0.294***	0.253***
		(0.000)	(0.000)
$Perc.gap\ x\ I_M\ x\ Sophistication$		0.067	0.053
		(0.376)	(0.486)
$Perc.gap\ x\ I_{HM}$	-0.388***	-0.106***	-0.369***
	(0.000)	(0.000)	(0.000)
$Perc.gap\ x\ I_M$	-0.173*	-0.046	-0.153
	(0.101)	(0.125)	(0.143)
$Perc.gap\ x\ AI\ literacy$	-0.089***		-0.073***
	(0.000)		(0.000)
$Perc.gap\ x\ Sophistication$		-0.278***	-0.236***
		(0.000)	(0.000)
$I_{HM}\ x\ AI\ literacy$	0.558***		0.502**
	(0.007)		(0.015)
$I_M\ x\ AI\ literacy$	0.382*		0.333
	(0.075)		(0.119)
$I_{HM}\ x\ Sophistication$		1.457**	1.330**
		(0.012)	(0.020)
$I_M\ x\ Sophistication$		1.007*	0.891
		(0.086)	(0.125)
I_{HM}	-2.184**	-0.161	-2.235**
	(0.013)	(0.498)	(0.011)
I_M	-1.468	-0.085	-1.457
	(0.111)	(0.724)	(0.111)
$AI\ literacy$	-0.405***		-0.343**
	(0.007)		(0.021)
$Sophistication$		-1.003**	-0.955**
		(0.018)	(0.022)
$Perc.gap$	0.882***	0.567***	0.859***
	(0.000)	(0.000)	(0.000)
Constant	-1.188	-1.755***	-1.275
	(0.264)	(0.042)	(0.227)
R-squared	0.558	0.556	0.566
Controls	Yes	Yes	Yes
Income FE	Yes	Yes	Yes
Education FE	Yes	Yes	Yes
Observations	1,668	1,668	1,668

Table A.6. Non-cognitive factors: disentangling the AI treatment

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) 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 treatment variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_M equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. Control variables are defined in Table A.1. *P*-values are in parentheses. Fixed effects include education and income levels. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.
$Perc.gap\ x\ I_{HM}\ x\ Female$	-0.052		-0.048
	(0.320)		(0.354)
$Perc.gap\ x\ I_{M}\ x\ Female$	0.125**		0.098*
	(0.023)		(0.077)
$Perc.gap\ x\ I_{HM}\ x\ Democrat$		0.126**	0.136**
		(0.034)	(0.020)
$Perc.gap\ x\ I_{M}\ x\ Democrat$		0.162***	0.111**
		(0.005)	(0.049)
$Perc.gap\ x\ I_{HM}$	-0.018	-0.080**	-0.053
	(0.597)	(0.011)	(0.168)
$Perc.gap\ x\ I_{M}$	-0.075**	-0.101***	-0.115***
	(0.044)	(0.005)	(0.005)
$Perc.gap\ x\ Female$	0.168***		0.166***
	(0.000)		(0.000)
$Perc.gap\ x\ Democrat$		0.026	0.017
		(0.507)	(0.663)
$I_{HM}\ x\ Female$	-0.680		-0.607
	(0.114)		(0.158)
$I_{M}\ x\ Female$	0.073		0.011
	(0.867)		(0.980)
$I_{HM}\ x\ Democrat$		-0.212	-0.120
		(0.642)	(0.789)
$I_{M}\ x\ Democrat$		0.176	-0.016
		(0.698)	(0.971)
I_{HM}	0.444	0.365	0.571*
	(0.116)	(0.181)	(0.071)
I_{M}	0.172	0.120	0.219
	(0.548)	(0.669)	(0.494)
$Female$	0.839***		0.768**
	(0.007)		(0.014)
$Democrat$		0.512	0.444
		(0.108)	(0.155)
$Perc.gap$	0.434***	0.507***	0.432***
	(0.000)	(0.000)	(0.000)
Constant	-3.638***	-3.286***	-3.251***
	(0.000)	(0.000)	(0.000)
R-squared	0.571	0.571	0.570
Controls	Yes	Yes	Yes
Income FE	Yes	Yes	Yes
Education FE	Yes	Yes	Yes
Observations	1,668	1,668	1,668

Table A.7. Expected Sharpe Ratio: disentangling the AI treatment

This table shows the coefficients of *Expected Sharpe Ratio* (defined as the ratio of posterior to perceived riskiness) 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 variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_M equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. Control variables are defined in Table A.1. *P*-values are in parentheses. Fixed effects include education and income levels. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.	3.
$Perc.gap\ x\ I_{HM}$	-0.033**		-0.033**
	(0.028)		(0.032)
$Perc.gap\ x\ I_M$		-0.050***	-0.050***
		(0.001)	(0.002)
$Perc.gap$	-0.148***	-0.146***	-0.148***
	(0.000)	(0.000)	(0.000)
I_{AI}			
I_{HM}	0.109		
	(0.121)		
I_M		-0.188	-0.196
		(0.128)	(0.120)
Female	-0.308***	-0.034	-0.181*
	(0.008)	(0.773)	(0.067)
Age	-0.009**	-0.004	-0.010**
	(0.040)	(0.322)	(0.012)
AI literacy	0.056	0.097*	0.080*
	(0.284)	(0.075)	(0.074)
Risk preference	-0.115***	-0.049	-0.085**
	(0.006)	(0.229)	(0.015)
Financial literacy	0.152**	0.255***	0.210***
	(0.030)	(0.000)	(0.000)
Sophistication	0.007	0.027	0.044
	(0.966)	(0.866)	(0.737)
Democrat	0.046	0.249**	0.094
	(0.685)	(0.030)	(0.329)
ChatGPT usefulness	-0.062**	-0.033	-0.055**
	(0.046)	(0.296)	(0.035)
Constant	1.545**	0.632	
	(0.013)	(0.317)	
R-squared			
Income FE	Yes	Yes	Yes
Education FE	Yes	Yes	Yes
Exclude	M	HM	No
Observations	1,112	1,122	1,668

Table A.8. ChatGPT: disentangling the AI treatment

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) and *Expected Sharpe Ratio* (defined as the ratio of posterior to perceived riskiness) on a set of independent variables. The perception gap (*Perc. gap*) is the difference between analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. The variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_M equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. Control variables are defined in Table A.1. 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.
$Perc.gap\ x\ I_{HM}$	-0.062	-0.022
	(0.262)	(0.326)
$Perc.gap\ x\ I_M$	-0.109*	-0.056**
	(0.066)	(0.022)
I_{HM}	-0.062	-0.022
	(0.262)	(0.326)
I_M	-0.071	0.048
	(0.864)	(0.781)
$Perc.gap$	0.556***	-0.110***
	(0.000)	(0.000)
Constant	-0.951	1.435**
	(0.526)	(0.021)
R-squared	0.462	0.299
Controls	Yes	Yes
Income FE	Yes	Yes
Education FE	Yes	Yes
Observations	558	558

Table A.9. ChatGPT results including stock return-ignorant investors

This table shows the coefficients of *Update* (defined as the difference between posterior and prior return expectation) and *Expected Sharpe Ratio* (defined as the ratio of posterior to perceived riskiness) on a set of independent variables. The perception gap (*Perc. gap*) is the difference between analyst forecast and prior expectations about the 12-month-ahead S&P500 index returns. The treatment variable I_{AI} equals one if the participant received a forecast involving *AI* (*Machine* or *Human + Machine*), and zero otherwise (*Human* condition). The variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_M equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. Control variables are defined in Table A.1. 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.
$Perc.\,gap\ x\ I_{AI}$	-0.626***		-0.127***	
	(0.000)		(0.000)	
$Perc.\,gap\ x\ I_{HM}$		-0.112**		-0.029**
		(0.046)		(0.044)
$Perc.\,gap\ x\ I_M$		-0.720***		-0.145***
		(0.000)		(0.000)
I_{AI}	-2.801***		-0.477**	
	(0.002)		(0.026)	
I_{HM}		-0.429		-0.108
		(0.643)		(0.644)
I_M		-2.974***		-0.424*
		(0.001)		(0.061)
$Perc.\,gap$	0.659***	0.663***	-0.071***	-0.070***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-3.232	-2.308	1.413	1.607*
	(0.402)	(0.503)	(0.131)	(0.066)
R-squared	0.369	0.497	0.744	0.777
Controls	Yes	Yes	Yes	Yes
Income FE	Yes	Yes	Yes	Yes
Education FE	Yes	Yes	Yes	Yes
Observations	601	601	601	601

Table A.10. Summary statistics: Experiments 2 and 3

This table reports the summary statistics for Experiments 2 (Provider familiarity) and 3 (Source revelation). Numbers in parentheses imply the range of possible values. Columns 6-10 check the balance of means across the control and two treatment groups, and *p*-values between control and treatment groups (Columns 10, 11 and 12). Belief variables reported in the table refer to individuals' prior elicited before the information manipulation. Return expectation prior is incentivized. *P*-values are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. We drop participants who exhibit return expectations below -30% and above 30%.

	1. Mean	2. SD	3. p10	4. p50	5. p90	6. Human	7 AI	8. Human + Machine	9. Machine Mean	10. p-value (7 = 6)	11. p-value (8 = 6)	12. p-value (9 = 6)
Panel A: Provider familiarity (558 participants)												
Female	0.478	0.499	0	0	1	0.481	0.477	0.440	0.513	0.931	0.438	0.534
Age	41.173	13.176	26	39	61	41.210	41.155	41.376	40.935	0.962	0.903	0.842
Risk preference (1-7)	4.342	1.480	2	5	6	4.427	4.300	4.370	4.229	0.341	0.710	0.200
Financial literacy (1-5)	4.116	0.935	3	4	5	4.070	4.139	4.172	4.106	0.411	0.299	0.705
Sophistication (0-1)	0.146	0.354	0	0	1	0.135	0.152	0.139	0.165	0.579	0.896	0.410
AI literacy (1-5)	4.693	1.771	2	5	7	3.935	3.815	3.725	3.903	0.267	0.102	0.798
ChatGPT usefulness (1-7)	8.748	6.382	3	7.8	17	4.762	4.659	4.526	4.791	0.519	0.195	0.875
Return expectation (prior)	4.980	1.399	3	5	7	8.717	8.764	9.111	8.420	0.934	0.565	0.647
Confidence (1-7)	0.478	0.499	0	0	1	5.048	4.946	5.032	4.860	0.416	0.912	0.174
Panel B: Source revelation (557 participants)												
Female	0.483	0.500	0	0	1	0.435	0.507	0.518	0.494	0.112	0.108	0.259
Age	43.260	14.517	25	41	63	42.220	43.782	44.105	43.439	0.231	0.199	0.421
Risk preference (1-7)	4.104	1.408	2	4	6	4.161	4.075	3.995	4.161	0.497	0.262	0.999
Financial literacy (1-5)	4.072	9.985	3	4	5	3.952	4.132	4.089	4.178	0.041	0.194	0.028
Sophistication (0-1)	0.133	0.339	0	0	1	0.124	0.137	0.168	0.106	0.651	0.228	0.588
AI literacy (1-5)	4.092	1.116	2	4	5	3.935	4.169	4.120	4.222	0.019	0.118	0.015
ChatGPT usefulness (1-7)	4.386	1.829	2	5	7	4.435	4.361	4.351	4.372	0.651	0.644	0.741
Return expectation (prior)	8.116	6.592	3	7.5	15	8.229	8.059	8.309	7.793	0.773	0.900	0.532
Confidence (1-7)	4.731	1.411	3	5	6	4.855	4.668	4.675	4.661	0.141	0.205	0.187

Table A.11. Reframing: disentangling the AI treatment

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 variable I_{HM} equals one if the individual is assigned to the *Human + Machine* forecast, and the treatment variable I_M equals one if the individual is assigned to the *Machine* condition; both are zero if the individual received the *Human* forecast. Control variables are defined in Table A.1. *P*-values are in parentheses. Fixed effects include education and income levels. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	1.	2.
I_{HM}	-0.132	-0.201
	(0.316)	(0.112)
I_M	-0.345***	-0.285**
	(0.009)	(0.027)
Constant	3.020***	4.206***
	(0.000)	(0.000)
R-squared	0.086	0.058
Controls	Yes	Yes
Education FE	Yes	Yes
Income FE	Yes	Yes
Observations	558	557
Experiment	ChatGPT	Reframing