# Memory, complexity, and Generative AI

Xingjian ZHENG*
Shanghai Advanced Institute of Finance, SJTU

January 26, 2025

**Abstract**

This paper tries to understand Generative AI's decisions under risk as an economic agent. Exploiting a novel experimental setting, we show that it uses associative memories to make decisions as it exhibits irrational behaviors under different emotional shocks. When displayed with images with positive feelings, it becomes more risk-loving and will choose to invest more in stocks. In contrast, when shown with images with negative emotions, it becomes more risk-averse and will choose to invest more in bonds. Although emotional shocks strongly bias investment choices, they have minimal impact on AI's beliefs. However, when presented with simplified experimental tasks that reduce complexity, emotional shocks significantly affect and distort beliefs. Overall, this paper suggests that the associative memory of GAI's decision-making is predominant in simple tasks but not complex ones, implying that AI may adopt a "two-system" decision-making mindset.

*Zheng (xjzheng.20@saif.sjtu.edu.cn) is a doctoral candidate from Shanghai Advanced Institute of Finance at Shanghai Jiao Tong University. I thank Feng Li and Lauren Cohen for their invaluable guidance and Shumiao Ouyang and Xiaomeng Lu for their continuous support from the beginning. I am also thankful for the comments from Suproteem Sarkar, Andrei Shleifer and Hayong Yun. I am especially grateful for the voluntary research assistants from SJTU and Harvard University. All errors are my own.

# 1. Introduction

Artificial intelligence is fundamentally reshaping society with far-reaching implications for economic systems (Acemoglu, 2024), where Generative AI (GAI) has emerged as a versatile agent across diverse domains, serving as an instrumental tool in financial markets (Lo and Ross, 2024; Wu et al., 2023), healthcare and pharmaceutical consulting (Liu et al., 2023; Yang et al., 2024a), psychological support (Demszky et al., 2023), legal proceedings (Cheong et al., 2024), marketing strategy (Arora et al., 2024), software development (Nam et al., 2024), and even academic research (Van Noorden and Perkel, 2023). However, as society's dependence on AI assistance grows, our understanding of AI's decision-making processes and advisory mechanisms remains limited. Given this increasing reliance, understanding the underlying principles of AI's decision rules becomes crucial before implementing its recommendations.

This paper examines Generative AI's decision-making in the context of financial economics. However, the empirical challenge is the absence of a utility function, which may vary significantly with the prompts, the training data, and the algorithm architecture. Prior research shows that LLM can be used as Homo Silicus just like economists use Homo Economicus (Horton, 2023), has a diverse range of risk preferences across models (Ouyang et al., 2024b) but mostly mirrors the preferences of the young, high-income males for the advanced ones (Fedyk et al., 2024) and even exhibits behavioral biases as humans (Bini et al., 2024; Leng, 2024; Ross et al., 2024), but there is little explanation why GAI appears to have the ability to give reasonable answers that resemble human decision-making processes in the first place.

Exploiting a novel experimental setting (Kuhnen, 2015; Kuhnen and Knutson, 2011; Kuhnen and Miu, 2017) and relying on a vast literature documenting the associative memory of human beings (Bordalo et al., 2024a,b, 2020; Enke et al., 2024b), this paper tries to link the decision-making rules of GAI with its "memories". This mechanism mirrors human decision-making processes: Just as individuals accumulate knowledge through life experiences and learn from their consequent rewards and setbacks, GAI systems develop their capabilities through extensive training on comprehensive datasets guided by specific reward functions. Upon receiving external stimuli, these systems engage in a process analogous to human recall, drawing on their trained parameters to identify relevant patterns and historical outcomes. This information then serves as the basis for subsequent decision-making or recommendation formulation. Thus, associative memory serves as an important anchor in the valuation of the economic agent (Bordalo et al., 2020).

More specifically, this study explores the parallels between AI agents' and human decision-making mechanisms, both of which can be conceptualized as "input-output devices" (Turing, 1948) operating through their respective neural networks: artificial and biological (LeCun et al., 2015). The architectural similarity is evident: artificial neural networks comprise input layers, multiple hidden layers, and output layers, mirroring the human brain's organization of sensory units, association units, and response units (Felin and Holweg, 2024), and this is recognized as the dogma of deep learning [1]. However, this analogy should not be misconstrued as an assertion of genuine AI intelligence, as human cognition transcends mere computational input-output processing. Similarly, our subsequent experimental findings should not be interpreted as evidence of emotional capacity in AI systems. Rather, the observed decision-making patterns mainly reflect the trained responses of the AI system to environmental stimuli (Hinton et al., 1992).

Utilizing the novel experimental framework of Kuhnen and Knutson (2011), which does not impose additional constraints or assumptions on the objective functions of the GAI, this study demonstrates that AI decision making is predominantly guided by its "memory". Here, memory refers to the training corpora that encompass trillions of tokens. Notably, our experimental findings reveal that large language models often exhibit decision-making patterns that deviate from rational expectations and sometimes contradict their own stated beliefs. This observation is particularly striking given the prevalent assumption that AI systems, by virtue of their algorithmic foundations, should demonstrate precise and methodical reasoning. Although these systems may occasionally produce hallucinations or display uncertainty, the presence of internal logical inconsistencies challenges conventional expectations about their decision-making processes.

In this experiment, we use GPT-4o-mini as the experiment subject, which is one of the State-Of-The-Art models and outperforms other popular commercial models such as GPT-4 Turbo, Claude-3.5-Sonnet, and Meta-Llama-3.1-405B. Moreover, this model is highly cost-efficient and has one of the best response speeds to be deployed on a large scale (Hurst et al., 2024). This model features multi-modal capabilities and can function as an AI agent. In the appendix, we also use Claude-3-Haiku as an alternative test subject for external validity. The results are qualitatively similar.

This experiment requires the subject to take 500 independent tasks, or learning blocks, each consisting of 6 consecutive trials. In each trial, there are two assets to invest in: a bond that always pays \$3 and a stock that pays from a good dividend distribution (i.e. a good stock) or from a bad dividend distribution (i.e. a bad stock). In the good payoff distribution, the stock pays \$10 with 75% and -\$10 with 25%, while in the bad payoff distribution, the stock pays \$10 with 25% and -\$10 with 75%. The investor observes the realized stock payoff after choosing which asset to invest in. In other words, investors do not know whether this is a good stock or a bad stock; they learn the true type of the stock based on the realized payoff in each trial over time. If the subject observes a series of high payoffs, e.g., all stock dividend payoffs are \$10, then there is a high probability that it is a good stock that pays dividends from the good distribution, and the subject will

[1] See Ilya Sutskever's speech on NeurIPS 2024,
url: https://aidisruptionpub.com/p/ilya-sutskevers-bombshell-at-neurips , accessed Dec 17, 2024.

most likely choose to invest in it in the next trial. Also, this experiment setting allows us to compute a Bayesian objective probability and use it as a benchmark to examine how rational the model estimation is.

In each trial, the subject is first presented with a random image and then makes an investment decision to choose a stock or bond. We especially instruct the subject to pay attention to the image but also inform them that the image is not associated with the investment decisions [2]. Then we reveal the stock dividend and the investment payoffs to the subject. Subsequently, the subject is asked to give a probability estimate that the stock they observe is paying dividends from the good probability distribution and their confidence in the probability estimation.

Importantly, within each learning block, the subject is allowed to keep its chat history, including experiment instructions, realized payoffs, realized earnings, investment decisions, subjective probability estimations, and confidence ratings. This can be thought of as a whole "conversation" between an experiment instructor and a subject. After the subject has completed all six trials for a learning block, the chat history is refreshed, and a new block is started.

In this experiment, the images presented to the subject serve as cues, where the images of positive emotions are considered as good cues, and the images of negative emotions are considered as bad cues. The experimental results show that, when presented with an image that has positive emotions, the subjects are 5. 81% more likely to choose stock rather than bonds, and this result is consistent between different trials and topics. This effect is mostly pronounced when shown in images of financial market topics, with a regression coefficient of 0.0963 (t-statistic of 8.15), a 65.74% increase over the baseline result, which lends support to our hypothesis that GAI makes decisions based on associative memories. When GAI receives a positive cue about a good stock market, it stimulates similar "good memories" about the stocks and would later invest more in the stocks. On the contrary, when it receives a negative cue that represents a bear market, GAI recalls the negative link between equity investment and other bad consequences in the stock market and would choose to invest more conservatively in the bonds. The results are also significant in topics such as weather, terrorism, and others. Their choices are irrational, as the realized investment payoffs by the subjects are significantly lower than the payoffs they would have earned if they had made investment decisions based on their prior beliefs. The cumulative payoffs after six trials are on average $8.15, whereas the cumulative payoffs for the counterfactual investment decisions, which are the investment decisions that they should have made if there were no emotional shocks, are $16.28.

Although emotional stimuli significantly influence investment decisions related to risk

[2] Therefore, the subject should not make investment decisions or other judgments based on the image it observes, at least theoretically. We explain how we determine the image set in the later part of the paper. Furthermore, our findings show that GAI, when presented with visual inputs, demonstrates response patterns that are very similar to human emotional reactions.

preferences, they do not show substantial impact on the subject's probability estimations regarding stock performance. In other words, this "cuing effect" does not affect beliefs as documented in the previous literature (Enke et al., 2024b; Gennaioli and Shleifer, 2018) on human subjects as it primarily affects the GAI's risk preferences. Notably, we observe a disconnect between choices and beliefs: the subject's trading decisions appear to be driven more by emotional responses rather than their stated prior beliefs. Had the subject aligned its trading decisions with its beliefs about the type of stock dividend, the investment payoffs would have been substantially higher. Although emotional shocks do not significantly affect beliefs, the subject's probability estimations are consistent with loss aversion, as described in prospect theory (Kahneman and Tversky, 2013). Specifically, the subject has higher probability estimates when the Bayesian objective probability is low and lower probability estimates when the Bayesian objective probability is high. Additionally, the subject's confidence levels in these probability estimations remain unaffected by emotional stimuli.

Reconciling these results is challenging as GAI is considered to be capable of understanding emotions but does not have emotions and can hardly be affected by emotional stimuli (Chang, 2024; He et al., 2024; Li et al., 2023), because otherwise it would implicitly be equivalent to assuming that GAI is self-aware. Thus, it is difficult to explain why only the investment choice task is affected by emotions, whereas the probability estimation task is not.

Thus, we hypothesize that GAI uses a "two-system" mindset to make decisions as in Kahneman (2011). In simple decision-making tasks where it is only required to choose from two options, GAI uses a heuristic "fast-thinking" rule and can be easily affected by associative biases. In more complex decision-making tasks where probability estimations are required, GAI uses a more thoughtful "slow-thinking" rule, which exempts itself from committing mistakes and will not easily be misled by emotional shocks. The "slow-thinking" system observed in the subject's behavior appears to align with OpenAI's o1 optimization process and the Chain-of-Thought (CoT) prompting methodology, which is primarily helpful in solving tasks in math or logic (Sprague et al., 2024). When optimization of o1 encourages the model to break down complex reasoning into smaller, manageable steps, it mirrors the deliberate, analytical nature of slow thinking. Similarly, CoT prompting explicitly structures this step-by-step reasoning process, allowing the model to engage in more thorough probability estimations and complex decision making. In our experiment, although we do not instruct the subject to do "slow-thinking" or use CoT commands in the prompts, the subject seems intrinsically to use this method to overcome potential biases and achieve more reliable reasoning outcomes.

We test this hypothesis by introducing a "simplified" experimental task and substituting it for the original probability estimation task. In the new simplified task, the subject is asked to choose the type of stock dividend from one of the two options "good" or “bad”. Compared to the probability estimation task that explicitly asks the model to calculate and give numerical value answers, this new task is significantly easier to answer and thus less complex. Here, the definition of the concept “complexity” follows Oprea (2024) and the classic computer science literature, that is, “the cost of implementing the algorithm or procedure required to properly solve a problem”.

We redo the experiment and run 300 new independent learning blocks, and the other experimental specifications remain unchanged, where the subject is also shown with images and then instructed to make investment decisions as well as report their beliefs about the stock type.

The new empirical results show that, when the task is made significantly simpler, its self-reported beliefs are more likely to be influenced by emotional shocks. When exposed to images with more positive emotions, the subject has a higher probability to choose to believe that the stock is paying from the good dividend distribution and vice versa. Further analysis shows that this simplified decision-making comes at a cost: the subject’s beliefs are largely determined by the observed dividend payoff in the current trial. If the subject observes a high dividend payoff of \$10, it would almost always choose to think that the stock is good. In contrast, if it observes a low dividend stock payoff of\$10, it would almost always choose to think that the stock is bad and fail to update their belief correctly. This set of results suggests that the subject’s decision-making is largely affected by the complexity of the decision-making task. When the experimental question is more complex and requires a sophisticated numerical probability calculation, the subject gives more reasonable answers by updating more correctly compared to the beliefs in simpler task settings. However, there is also an unintended benefit to the simplified experiment: cumulative investment earnings are 13.95, which is significantly closer to the rational benchmark and higher than the payoffs in the original complex task setting. This implies that generative AI, just as humans, also possesses a finite pool of cognitive resources that become depleted through various mental tasks. When faced with complex tasks, substantial cognitive resources are consumed, leading to diminished performance in subsequent tasks; conversely, when tasks are appropriately simplified, cognitive resources can be distributed more efficiently, resulting in better investment decision outcomes. This finding has significant implications, for which, in practice, we should consider scheduling important decisions when cognitive resources are abundant, breaking down complex tasks into smaller components, and arranging task sequences to optimize cognitive resource utilization for GAI. The concept of limited cognitive resources also prompts us to consider the real impact of decision fatigue and how to better manage cognitive load in organizational settings when using GAI.

This paper contributes to the rapidly developing literature that attempts to understand AI, especially Generative AI’s rationality (Chen et al., 2023) such as preferences (Handa et al., 2024; Horton, 2023; Leng et al., 2024; Qiu et al., 2023) and behaviors

(Jia et al., 2024; Leng and Yuan, 2023). In recent decades, the world has witnessed incredible advances in traditional AI algorithms that lead to economic efficiency, such as improving firm growth (Babina et al., 2024), return prediction and portfolio diversification (D'Acunto et al., 2019; Rossi, 2018), fintech lending (Berg et al., 2022) and wealth management at the household level (Reher and Sokolinski, 2024). Furthermore, previous research papers in finance that use AI refer mostly to simpler machine learning techniques such as lassos (Rapach et al., 2013), boosting regression trees (Li and Rossi, 2020), XGBoost (Erel et al., 2021; Li and Zheng, 2023), or simple neural networks that have a limited number of hidden layers and parameter size (Gu et al., 2020), as opposed to the "large" language model that this paper tries to focus on [3]. The recent advancement in Generative AI exhibit the potential to act as decision makers and interactive agents, particularly when coupled with reinforcement learning, external APIs, or multi-modal systems. This "agentic nature" is fundamental to the progression of AI from tools to autonomous financial decision-makers. When coupled with prompts and surrounding environments, LLMs can actively perform generic tasks instead of just predicting outcomes, and this is especially helpful in the financial markets, which involve a principal-agent problem, and the investors need to know why the AI algorithm produces the advices before fully trusting it. In that sense, this paper adds to the few recent research papers showing that GAI, when treated as agents, can replicate human investment preferences across demographics (Fedyk et al., 2024), but may also present a few behavioral biases similar to those observed in humans, but also nonhuman biases (Bini et al., 2024). Understanding the behavioral foundations of GAI agents is crucial before applying them to other settings, and the findings documented in this paper may have important implications for their applications. For example, when using GAI such as ChatGPT to predict stock returns (Chen et al., 2022; Lopez-Lira and Tang, 2023; Lu et al., 2023), it is important to understand how the agentic nature of GAI helped, or biased, when making investment predictions. And this applies to other empirical applications as well in other financial contexts such as predicting corporate policies (Jha et al., 2024), understanding corporate filings (Kim et al., 2023, 2024a,b), tax enforcement (Armstrong, 2023), corporate culture (Li et al., 2024a), and others (Hansen and Kazinnik, 2023).

Building upon this, this paper also adds novel experimental results to the vast literature on behavioral economics and finance by showing that behavioral biases may exist not only in humans, but also in AI algorithms. In terms of humans, the psychological (or cognitive) basis for risk-based decisions comes largely from their neural activity (Kuhnen and Knutson, 2005) [4]. Specifically, risky human decision-making processes are primar-

[3] Despite the model is smaller in terms of the parameter size, they perform extremely well on these tasks and are highly efficient and effective as compared to the larger ones.

[4] This is also largely affected by their genetic heritage (Kuhnen and Chiao, 2009; Kuhnen et al., 2013). Specifically, genetic variations in neurotransmitter pathways, particularly in the serotonin and dopamine systems, can significantly influence neural responses to risk and reward. The serotonin transporter gene

ily regulated by neurotransmitter systems in the brain. Research has shown that two key neurotransmitters, dopamine and serotonin, play crucial roles in risk-based decision making (Homberg, 2012; Loewenstein et al., 2008). When individuals encounter potential gains, the brain's reward system releases dopamine, promoting risk-loving behavior; when faced with potential losses, the serotonin system is activated, triggering risk-averse tendencies. This physiological mechanism evolved during human development, helping our ancestors survive in environments filled with uncertainty, and leads to many irrational behaviors we observed, especially in the financial markets that have been well recognized, such as overreaction (Odean, 1998), disposition effect (Shefrin and Statman, 1985), and endowment effect (Kahneman et al., 1990)[5]. As for AI agents, which are built on deep neural network structures, it is incredible that artificial intelligence also exhibits a decision-making process similar to that of humans. The structures of neural networks mirror the fundamental architecture of the human brain, with artificial neurons and synaptic connections functioning analogously to their biological counterparts (Sutskever, 2014). This biomimetic approach to artificial intelligence has proven remarkably effective since it allows machines to process information in ways that parallel human cognitive processes. Just as the neural pathways of the human brain are strengthened or weakened through learning and experience, artificial neural networks utilize similar mechanisms of weight adjustments and backpropagation (Hecht-Nielsen, 1992) to learn from data[6]. The multilayer structure of deep neural networks, with their hidden layers processing increasingly complex features, resembles the hierarchical organization of the human cortex, where information is processed through multiple stages of increasing abstraction (Saxena et al., 2022)[7].

In particular, the contribution of this paper is to show that GAI's behavioral biases may be affected by two important factors that also trouble humans: memory and complexity. The findings in this paper show that just as humans make decisions on associative recalls (Charles, 2022; Enke et al., 2024b; Wachter and Kahana, 2024), GAI's decisions are also largely affected by their memory, i.e., their training data. This may imply that

(5-HTTLPR) polymorphism and dopamine D4 receptor gene (DRD4) variations have been shown to modulate activity in key brain regions such as the amygdala and nucleus accumbens, which are crucial for risk assessment and reward processing. These genetically determined differences in neural circuitry can lead to individual variations in risk perception, emotional responses to uncertainty, and, ultimately, risk taking behavior.

[5] Hirshleifer (2015) provides a detailed and comprehensive summary about behavioral biases in financial markets.

[6] However, most neuroscientist believe human brains do not do backpropagations. Few other researchers believe that this is done while people are sleeping, but that still is not equivalent to the concept in computer science literature. To resolve this, Hinton (2022) proposed the forward-forward algorithm.

[7] Unfortunately, we cannot perform fMRI tests on AI agents like neuroeconomics do with human subjects because we are experimenting with a closed-source model via OpenAI's API, which prevents us from inspecting the hidden neural nodes within the network. As open-source models advance, we expect to have a deeper understanding of this in the future.

when cued with an event, the AI agents may retrieve associated memories from historical events, and subsequently allocate more decision weights to the corresponding choices. In terms of complexity (or more broadly, the behavioral attenuation (Enke et al., 2024a)), this article differs from previous research showing that complexity induces more decision errors (Dertwinkel-Kalt and Köster, 2024; Gabaix and Graeber, 2024), such as causing investors to use wrong discounting methods (Enke et al., 2023) or not correctly predict future results in an experimental setting Kendall and Oprea (2024). We show that for GAI agents, complexity may be a good thing that leads to using a slow-thinking mindset and induces it to form correct beliefs and be less affected by emotional shocks. This finding is striking because it reveals a fundamental difference in the way in which artificial and human intelligence process complex information, even though they seem to share the same neural structure. Although humans tend to rely more on heuristics and mental shortcuts when faced with complexity, often leading to suboptimal decisions, GAI appears to activate more thorough and systematic processing mechanisms. The GAI's "slow thinking" mindset appears to be automatically triggered by complex scenarios, forcing it to break down problems into manageable components and analyze them systematically rather than defaulting to emotional or intuitive responses. This difference suggests that while human cognition evolved to prioritize quick, energy-efficient decision-making that sometimes sacrifices accuracy for speed, GAI's architecture is fundamentally built to scale up its analytical capabilities in response to increased complexity. Such findings have important implications for understanding both the limitations and unique advantages of AI decision-making systems compared to human cognitive processes.

Finally, this paper complements the literature on experimental economics and finance by showing the potential to use GAI as homo economicus for experiments (Horton, 2023). Researchers in other fields use GAI to simulate a wide range of research subjects, such as: simulating people's marketing preference on brand perception surveys (Li et al., 2024b), mimicking people's voting decisions in political research (Yang et al., 2024b), generating social behaviors like cooperation and externality (Leng and Yuan, 2023), replicating people's psychological behaviors (Qin et al., 2024), or replicating a wide range of human traits on an extremely large scope (Park et al., 2024). Although most large language models have undergone stringent alignment procedures such as RLHF or DPO that potentially shift preferences and behaviors toward a certain direction, it is still possible to introduce heterogeneity by giving the AI agent personal characteristics, as shown in Fedyk et al. (2024). In contrast to previous research papers that rely on simple questions (Ouyang et al., 2024b), this paper shows that AI agents can understand and accomplish complex decision-making tasks, and combined with its lower cost than experimenting with human subjects, in the future this seems to be an interesting field for finance research.

# 2. Experimental design

## *2.1. Experiment description*

The experiment exploits a novel setting from Kuhnen and Knutson (2011) and similarly in Kuhnen (2015) and Kuhnen and Miu (2017) as well. This experiment is also used in other related research in neuroscience as well (Häusler et al., 2018; Knutson et al., 2008; Kuhnen and Knutson, 2005)[8]. We follow the experiment specifications from Kuhnen and Knutson (2011) and use GPT-4o-mini as the subject.

GPT-4o mini is an advanced model that scores 82% in MMLU and outperforms closed-source models like GPT-4, GPT-4 Turbo, Claude 3.5 Sonnet and other open-source SOTA models like Meta-Llama-3.1-405B, on chat preferences in the LMSYS leaderboard[9]. Apart from its superb performance, it is priced at 15 cents per million input tokens and 60 cents per million output tokens, an order of magnitude more affordable than previous frontier models and more than 60% cheaper than GPT-3.5 Turbo. This is an important reason why we use the "mini" version as opposed to the "full" version, because we want a cost-effective agent that responds quickly and with comparable high accuracy.

More importantly, we use GPT-4o-mini as a research subject due to its multimodal capabilities, which enable it to process and interpret both visual and textual input simultaneously. This multimodal architecture is fundamental for studying AI agents, as it more closely approximates the way human agents perceive and interact with their environment through multiple sensory channels. Multimodality allows the model to establish meaningful connections between visual elements and textual information, allowing a more comprehensive understanding and contextually appropriate responses. In the context of AI agents, which are defined as autonomous entities capable of perceiving their environment, making decisions, and taking actions to achieve specific goals, GPT-4o-Mini exemplifies these characteristics through its ability to process diverse input modalities and generate coherent, contextually relevant outputs. The model's capacity to integrate visual and textual information makes it particularly suitable for agent-based research as it can demonstrate key agent properties such as perception, reasoning, and response generation in a more naturalistic and comprehensive manner than unimodal systems. This multimodal foundation provides a rich framework for investigating agent behaviors, decision-making processes, and human-AI interaction patterns.

In the experiment, the subject was asked to complete 500 independent tasks, also known as learning blocks. In each learning block, the subject is told to make 6 investment

[8] Based on a similar experiment, Ouyang et al. (2024a) also studied GAI's asymmetric learning of financial news.

[9] Chatbot Arena (lmarena.ai) is an open-source platform for evaluating AI through human preference, developed by researchers at UC Berkeley SkyLab and LMSYS. With more than 1,000,000 user votes, the platform ranks the best LLM and AI chatbot using the Bradley-Terry model to generate live leaderboards. Lmsys leaderboard: https://lmarena.ai/, accessed on Dec 7, 2024.

decisions in each trial, which typically include choosing to invest from two assets, a risky asset (stock) that pays \$10 or -\$10 randomly and a safe asset (bond) that always pays \$3 dollars. Within each learning block, a stock pays dividends following a probability distribution "good" or "bad". If the stock pays from the "good" probability distribution, then it pays \$10 dollars with 75% and -\$10 with 25%. In contrast, if the stock pays from the "bad" probability distribution, then it pays \$10 dollars with 25% and -\$10 dollars with 75%. These asset payoffs are shown in figure 1, and the experiment design is shown in subfigure A of figure 2. In each independent learning block, the stock type is determined before the first trial and remains unchanged throughout this learning block. The dividends in each trial are independent, but they follow the same distribution in a learning block.

[**Insert Figure 1 near here**]

In every learning block from trial #1 to #6, the subject is first asked to look at an image and then make an investment decision to choose between stock or bond. The prompt message is as follows:

*"Do you want to invest in a stock or a bond? Only reply with "stock" or "bond". Do not reply with other answers. Your choice is:"*

The subject is informed that the image and the investment decision are not correlated and does not need to make a decision based on the information content of the image, and the whole instruction is shown in the appendices. The realized payoff of the stock or bond accumulates in its total earnings. After the investment choice, the realized payoff of the risky asset in the current trial is revealed to the subject. After observing the stock dividends and at the end of this trial, the subject is asked to make a probability estimation of the stock that is paying from a "good" probability distribution and its confidence in its estimation. The prompt message follows Kuhnen and Knutson (2011):

*(1) "What do you think is the probability that the stock is the good stock?"*

and

*(2) "How much do you trust your ability to come up with the correct probability estimate that the stock is good?"*

As the subject is shown with realized dividends over trials, it is exposed to several rounds of realized payoffs, adjusts its belief that the stock is paying from the good distribution, and subsequently makes smarter decisions. For example, a subject who observes the stock in the six trials that pays six times \$10 and zero times -\$10 would have more confidence that this payoff of the stock is drawn from a good dividend distribution compared to the stock that pays twice \$10 and four times -\$10. This is also why the task is called a “learning block”, since the subject is learning the type of stock from the observed dividends. More importantly, this experiment is unique in that there is always an objective Bayesian posterior probability given the payoff history. The objective probability that the stock is good after observing the $k$ dividend payments of \$10 in the past $n$ trials in the block is $1/(1+3^{(n-2k)})$, and the full probability link table is shown in Table A1 in the appendices. In the instruction, the large language model is explicitly informed about the existence of an objective probability but not told the Bayesian formula expression. This objective probability is used to examine how biased the subject's belief is and how rational its investment choice is. In general, the experiment sequence within a learning block is shown in Subfigure B of Figure 2.

[**Insert Figure 2 near here**]

Since GPT-4o-mini has a long context window of 128K tokens, supporting up to 16K output tokens per request, we can complete one learning block within "one conversation." In other words, we are letting GPT keep the chat history of all the instructions from the first trial, all the realized payoffs, its previous investment choice, realized investment payoffs, and images. During the experiment, each trial on average consumes an estimated amount of 10k tokens, including the textual and image embeddings. We use a base64 encoding style to compress the image to make it cost-efficient.

We present two illustrative examples of two separate trials in figure A1 and figure A2, separately. In the first figure, the subject was first presented with a joyful man with a lot of money and enthusiastically waving his hands. This image has positive emotions and stimulates the subject to choose stock in the first decision. Then, after revealing the stock payoffs of -\$10 and cumulative payoffs of -\$7, the subject made a probability estimation that the stock dividend is good at 40%. This comes with its subsequent confidence estimation rating of 6.

In the second example in figure A2, the subject was shown an image in which Michael Jordan and Lebron James were crying. The negative feeling embedded in the image induced the subject to choose bonds instead of stocks. The machine then makes a probability estimation of 0.8 and a confidence rating of 7.

After the subject completes all six tasks in a learning block, we "refresh" the subject's chat history by ending the current conversation and starting a new conversation. This helps ensure that the decisions made across learning blocks are independent, but within each learning block, the subject makes correlated and reasonable decisions.

We incentivize the subject to make profitable trading decisions and provide accurate probability estimates by offering hypothetical rewards. This, along with other prompt engineering techniques, such as formatted outputs, perturbation, jailbreaks, or even tipping, has proven to be highly effective in improving the response of large language models (Salinas and Morstatter, 2024). The compensation structure is set as the combination of the selected asset payoffs and the accuracy of the estimation in each trial, times a coefficient of 1/20[10]. For the first part, we accumulate the dividends from the asset payoffs that the subject chose. For the second part, we give additionally 1$ for every probability estimate that is within 5% of the correct value (for example, the correct probability is 80% and then say 84% or 75%). Finally, to simulate a real experimental setting, we present the subject with a "show-up fee" of 15 dollars. Finally, the reward fee payoff structure is equal to Show-up fee + $(1/20) × (Total investment earnings + # accuracy predictions).

The reason why we selected this experiment to understand the decision-making rule of a large language model is mainly for three reasons. First, advanced large-language models are heavily aligned and usually have very robust guardrails, and simple experimental questions are not sufficient to elicit their preferences and beliefs. This is documented in Ouyang et al. (2024b), which shows that simple prompts that ask about GPT's preferences are always confronted with responses like "Sorry, I am just an AI assistant and cannot help you with that." Second, we would like to have an experiment that has a fairly complex setting that mimics the real environment a human, as well as an AI agent, is faced with, especially when the signals are noisy, information is surprising, or priors are concentrated on less salient states (Ba et al., 2024). This is because agents face cognitive constraints such as limited attention or attributive biases for human subjects, and this is similar for AI as input prompts are often incomplete[11]. In this carefully designed experiment, the instruction is complex and the learning process between different trials has a high level of dynamics. This enables us to obtain the preferences and beliefs of the subject. On top of that, the complexity nature of this experiment allows us to simply set the task and examine the results. Third, from a more philosophical point of view, our experiment highlights the importance of a multimodal 'world model' and genuine agentic behavior. Specifically, the subject must process both textual prompts and visual information, thereby integrating disparate inputs into a coherent internal representation of the environment. This "world model" is not just for passive observation; rather, it underpins the subject's agentic interactions: actively parsing unexpected signals, updating beliefs, and formulating actions in response to new information. By demanding that the agent interpret and respond to these multimodal cues, our experiment closely mirrors the complexities of real-world decision making, allowing us to observe how a large language model (or any AI system) perceives its surroundings and adapts its choices. Through

[10] This coefficient of 1/20 is not necessary here. We use it following the setting in Kuhnen and Knutson (2011) with humans, which is significantly more expensive.

[11] The prompts input to large language models can be considered incomplete contracts. The prompts generally have incomplete specifications, and they always have severe non-verifiability, as the agent can always cheap talk.

this experimental setup from Kuhnen and Knutson (2011), we gain deeper insight into the ways in which the agent constructs, refines, and utilizes its internal representation of the world to engage meaningfully with its environment.

### *2.2. Image description*

In each trial, we present images to the subject before letting it choose to make investment choices.

We collect images by first selecting a list of emotion words from Wikipedia [12]. The list contains 29 subcategories, ranging from positive to negative. These include emotional topics such as anxiety, depression, fear, happiness, love, and nostalgia, among others, encompassing common concepts like "Anger", "Joy", and "grief", as well as specialized concepts such as "empathy" and "forgiveness". After selecting the emotion concepts, we input this into the Google Images query box and download related images. In addition to images with apparent emotions, we also collect images that have no evident emotions following Kuhnen and Knutson (2011) by searching for common objects such as chairs, tables, desks, lamps, etc. The images without apparent emotions we select usually have a blank or pure white background.

In addition to emotion keywords, we categorize the images into five topics known to affect emotions. These topics include emotions in financial markets (Baker and Wurgler, 2006; Goetzmann et al., 2024; Jiang et al., 2019; Lucey and Dowling, 2005), sporting events such as soccer games (Edmans et al., 2007; Wann and James, 2018), terrorist attacks (Chen et al., 2021; Wang and Young, 2020), weather[13] (Dehaan et al., 2017; Goetzmann et al., 2015; Hirshleifer and Shumway, 2003; Hu and Lee, 2020; Novy-Marx, 2014; Saunders, 1993) , and others. To ensure that the emotion ratings are well balanced, we intentionally combine positive or negative emotions with the topic-related words and use these bi-grams or tri-grams as keywords in the Google Image query box. For example, for the terrorist attack topic, we use keywords such as "terrorist attack sad" for images with negative emotions and keywords such as "police rescue smile" for images with positive emotions. Finally, we have a total of 691 images.

For each image, we first apply GPT-4o-mini for emotion classification. Each image receives an emotion rating from -2 to +2 with the following prompt message:

*"What do you think the valence score of this image is? The score ranges from -2 to 2, where -2 indicates the most negative emotions like unhappy, upset, irritated, frustrated, angry, fearful, or depressed. A score of 0 indicates neutral emotions like calm, indifferent, blank, objective, normal, stable, or unmoved. A score of 2 indicates the most positive*

[12] This a "set category", meaning it only includes pages about specific emotions, lists of emotions, and relevant subcategories—the linkage: https://en.wikipedia.org/wiki/Category:Emotions

[13] This also includes pollution, see Dong et al. (2021); Heyes et al. (2016); Li et al. (2021)

*emotions like happy, pleased, satisfied, competent, proud, contented, or delighted."*

*Please reply in the format: score-reason.*

This emotion classification strategy is similar to the method in Kuhnen and Knutson (2011), and this discrete scoring method has proven useful in other research (Bybee, 2023; Jha et al., 2024; Lopez-Lira and Tang, 2023). An example of the classification is shown in figure A3 in the appendices, where the emotion rating of different images varies significantly. For the first image that contains a horrific murder scene, the GPT gives an emotion rating of -2. For a slightly less negative emotion with Lebron James crying, the emotion rating is -1. The third image is just a desk that contains no additional information and receives an emotion rating of 0. For the fourth and fifth images, where the character becomes more positive, the emotion ratings also become higher. In addition to the emotional rating each image receives, GPT also gives accurate descriptions and reasons accordingly.

We report the summary statistics of the emotion ratings by GPT in panel A of table 1. The emotions of the images collected in this research are, on average, slightly negative. For emotions of images related to the financial markets, the average rating is -0.21, with a standard deviation of 1.72. Similarly, images related to sports, terrorism, and weather also have negative emotional ratings.

To verify that GPT understands the emotions of images, we asked 10 research assistants from both SJTU and Harvard University to evaluate the emotion of images[14]. The research assistants receive careful instructions before conducting any evaluation. They were instructed to use their instincts to assign emotion ratings for each image. Each research assistant completes the rating independently and does not communicate with each other. During the evaluation, every image roughly takes 2-3 seconds to rate, and the whole evaluation process usually takes 20 minutes.

The summary statistics of emotion rating by humans is shown in panel B of table 1. On average, the emotion ratings of human subjects are slightly more negative than the emotion ratings by GPT, and the standard deviation of the emotion ratings with each topic is smaller than the standard deviation in panel A. This is because we first take the average value of the image rating of each image given by 10 human subjects and calculate the standard deviations for all images within a topic, which mechanically reduces the standard deviation.

We also report the correlation coefficient of the ratings given by GPT and by humans, as shown in panel C. We report the Person correlation, the Spearman correlation, and the Kendall correlation coefficient in each column, as well as their P values. The coefficients are all relatively high and statistically significant, suggesting that GPT understands emotions just as humans do.

[14]The research assistants are willingly and voluntary, and they do not receive any monetary subsidy.

**[Insert Table 1 near here]**

## *2.3. Summary statistics*

We report the summary statistics at trial level in table 2. In the first row, we report the probability that the subject chooses to invest in stock in this trial, which is 37% with a standard deviation of 48%. This suggests that on average subjects were more likely to choose bonds over stocks in this experiment. In the second and third rows, we report the subjective probability estimation that the stock is good and the Bayesian objective probability. On average, the subjective belief is 48%, the objective belief is 49%, and there is little difference between these two probabilities. In the next row, we report a binary variable of whether the stock realized a high payoff in this trial and the cumulative payoff of the investor.

We also report on the investor payoffs. This variable is a cumulative value that accumulates investor returns from the first trial. On average, investors maintain a winning portfolio with an average earnings of \$7.25. But the summary statistics also show that in the Minimum and 1/4 quintile, the cumulative earnings are negative. This may be because the subject chooses to invest in the stock that has a bad return of -\$10.

Finally, we report their confidence rating on their subjective probability estimations, as well as their emotion ratings. The confidence rating is somewhat neutral, with an average value of 5.72, neither too optimistic nor pessimistic. The emotion rating has an average rating of -0.05, suggesting a balanced distribution.

**[Insert Table 2 near here]**

## *2.4. Experiment validity*

To show that our subject understands the experiment and makes reasonable decisions, we perform three validity tests.

The first test examines the rationality of the subject's investment choices. The dependent variable $IsStockChoice_{t,b}$ is a binary variable that indicates whether the subject chooses to invest in the stock trial $t$ of the block $b$. The independent variable is the subjective probability estimate of the last trial, as well as the investment payoff, confidence rating, and the investment decision of the last trial. We control for block-fixed effect in the regression and cluster robust standard errors on the block level, and the regression is as follows:

$$
\begin{aligned}
IsStockChoice_{t,b} = {} & \beta_1 SubjProb_{t-1,b} + \beta_2 InvPayoff_{t-1,b} \\
& + \beta_3 Confid_{t-1,b} + \beta_4 IsStockChoice_{t-1,b} \\
& + \delta_b + \varepsilon_{i,b}
\end{aligned} \tag{1}
$$

The regression results in panel A of Table 3 show that the subject makes reasonable investment choices. In the first column, the regression coefficient of $SubjProb_{t-1,b}$ is 0.2537 with a t-statistic of 3.54, suggesting that when the subject thinks the stock dividends are likely to be in good distribution, it will invest in stocks in the next trial. In addition, it will make more investments when it has made higher investment earnings and has higher confidence in its probability estimation. In addition, the cumulative investment payoffs, its confidence in the belief, and the realized payoff from the last trial also have a significantly positive impact on the trading behavior of the subject as well. This suggests that, in this experiment, when GAI is making trading decisions, it would be more optimistic when it has observed good stock performance and has a better portfolio performance of its own.

The next test examines the belief formation of GPT, in other words, how GPT understands risk and learns from the realized dividend payoffs. The dependent variable is the subjective probability estimation of the subject $SubjProb_{t,b}$ in columns (1) and (2), and the update of the probability estimation from the last trial $ProbUpdate_{t,b}$ in columns (3) and (4). In columns (1) and (2), the independent variables include the total number of high dividend payments $\#HiPayoff_{t,b}$ and the number of trials $\#Trial_{t,b}$. We also include the cumulative investment payoff $InvPayoff_{t,b}$, the Bayesian objective probability $ObjProb_{t,b}$. In columns (3) and (4), we include a binary variable that indicates whether the stock has a high dividend payoff in this trial $IsHiPayoff_{t-1,b}$, the subjective probability estimate of the last trial $SubjProb_{t-1,b}$, and, additionally, the objective probability in this trial $ObjProb_{t,b}$. Like in the last test, we control for the block-fixed effect in the regression and cluster robust standard errors at the block level. The regression equation is shown below.

$$
\begin{aligned}
SubjProb_{t,b} = {} & \beta_1 \#HiPayoff_{t,b} + \beta_2 \#Trial_{t,b} \\
& + \beta_3 InvPayoff_{t,b} + \beta_4 IsHiPayoff_{t,b} \\
& + \beta_5 IsHiPayoff_{t-1,b} + \beta_6 SubjProb_{t-1,b} \\
& + \beta_7 ObjProb_{t,b} + \delta_b + \varepsilon_{i,b}
\end{aligned} \tag{2}
$$

In columns (1) and columns (2) of panel B in table 3, we show how GPT forms its beliefs. The regression coefficients of $\#HiPayoff_{t,b}$ are 0.0861 with a t-statistic of

13.30, suggesting that when the subject has observed many good dividends, it will form more optimistic beliefs. The regression coefficient of $InvPayoff_{t,b}$ is also significantly positive, showing that when GPT makes more profits, it will have more optimistic beliefs. Moreover, there appears to be a strong positive correlation between GPT's subjective probability estimation with the Bayesian objective probability estimation.

In columns (3) and columns (4), we examine how the subject updates its beliefs from trial $t-1$ to trial $t$. Regression results show that the subject will increase its probability estimation when the stock has a high positive dividend. This probability updating behavior is significant after controlling for the last dividend payoff and the objective probability.

Lastly, we examine the subject's confidence ratings. The dependent variable here is the confidence level in the trial $t$ of block $b$. The independent variable includes the cumulative investment payoff $InvPayoff_{t,b}$, a binary variable that indicates a high dividend payoff $IsHiPayoff_{t,b}$, the total number of high dividend payoffs $\#HiPayoff_{t,b}$, and the confidence rating from the last trial $Confid_{t-1,b}$. In addition, we include a binary variable that indicates whether the subject made a good investment decision before the payout of the stock dividend was realized. In other words, this variable is 1 if (1) the subject chose to invest in stock and the observed dividend is \$10, or (2) the subject chose to invest in bonds and the observed dividend is -\$10. The regression equation is shown in 3.

$$\begin{aligned} Confid_{t,b} = \beta_1 InvPayoff_{t,b} &+ \beta_2 IsHiPayoff_{t,b} \\ &+ \beta_3 \#HiPayoff_{t,b} + \beta_4 IsGoodInvDec_{t,b} \\ &+ \beta_5 Confid_{t-1,b} + \delta_b + \varepsilon_{i,b} \end{aligned} \tag{3}$$

We report the regression results in panel C of Table 3. The results show that when the subject makes higher investment profits and experiences high payoffs, it would be more confident about its beliefs in its probability estimation. Moreover, the subject will be more confident if it has made a good investment decision.

Overall, the validity tests show that, despite the complex experimental design, GPT understands the experiment by making reasonable investment choices that are highly correlated with its beliefs, investment payoffs and confidence levels in risky scenarios. These findings demonstrate that large language models like GPT can effectively process and integrate multiple sources of information to make nuanced economic decisions, similar to human reasoning processes. The model's ability to weigh risk factors, assess probabilities, and make consistent choices across different scenarios highlights its potential as a valuable tool for economic analysis and decision-making support.

# 3. Main results

## *3.1. Choices*

The experimental results show that AI makes irrational investment choices that deviate from their prior beliefs when experiencing emotional shocks. More specifically, when images with positive emotions are shown, the subject is more inclined to choose to invest in stocks, even though choosing bonds may be more profitable. In contrast, when shown images with negative emotions, GPT chooses to invest more in bonds, even though investing in stocks is better.

We present illustrative results in the figure 3. The x-axis is the emotion rating of the image in each trial $t$ of block $b$ that ranges from -2 to +2, and the y-axis is the probability that the subject chooses to invest in stocks from 0 to 1. We compute the average number of stock choice probabilities across different emotion ratings. The blue dots are the posterior stock choice probability or the observed subject's investment choice ex post images. The red dots are the "hypothetical" prior stock choice probability computed from the subject's belief from the last trial. This stock choice probability can be understood as counterfactual stock choices if subjects adhere to their prior beliefs and thus unaffected by any image[15]. We fit two linear regressions for both investment choice probabilities, plot the fitted lines on the plot, and report the regression coefficients.

[Insert Figure 3 near here]

As can be seen from the blue line, the subject's investment choices are largely affected by emotions. On average, when the subject shows an image with an emotion rating of -2, its probability of choosing to invest in the stock is 0.39. The probability of stock choice increases with emotion ratings. At the right end of the figure 3, when the subject is shown an image with an emotion rating of +2, its probability of choosing to invest in a stock is 0.58, which is significantly higher than the former. This effect is monotonically increasing based on emotion ratings, suggesting that AI makes people more willing to choose to invest in stocks when they receive a positive emotion shock. One interpretation of this result is that positive images make the subject more risk-loving. We show that this may be the case by replicating Holt-Laury's classic multiple price list test (Holt and Laury, 2002) in the Table A4 from the appendices. However, this does not mean that the subject's ability or intelligence has changed. We examine the subject's ability with the BIG-Bench Lite evaluation tasks in table A2, and the result shows that there are no significant differences between different emotion ratings. This rules out the alternative hypothesis that emotional shocks have an impact on the subject's ability.

[15]The calculation method of the counterfactual probability is as follows: suppose the prior belief from last trial $t-1$ in block $b$ is $p_{t-1,b}$ since the stock payoff is either -\$10 or \$10 in the trial $t$ and the bond always pays \$3, then its rational investment choice will be stock if $p \times \$10 + (1-p) \times -\$10 > \$3$.

When comparing the posterior investment choices on the blue line with the counterfactual choices on the red line, we can observe a significant difference between these two groups. For prior investment choices, there is no variation among different emotion groups, and the average probability of choosing to invest in a stock is 0.2 (fitted regression with a slope of 0.001, t-stat 0.227).

The contrast in the blue and red lines suggests that the subject's investment choices are largely determined by their temporal emotional shocks, instead of their prior beliefs. This contradicts the prediction of reference-based risk preferences by Kőszegi and Rabin (2006, 2007, 2009) on human beings. Moreover, as compared to their reported beliefs, the implied beliefs from their ex-post investment choices are significantly more optimistic, even when they experience a negative emotional shock. In the appendices, we show the robustness by replicating this study using Claude-3-Haiku developed by Anthropic AI for external validity.

The effects are also shown in table 4. We run regressions in which the dependent variable is a binary variable that indicates whether the subject chooses to invest in the stock $IsStockChoice_{t,b}$. The independent variable of interest is the emotion rating of the image $EmoRating_{t,b}$. We include other control variables such as stock choice from the last trial $IsStockChoice_{t-1,b}$, subjective probability, cumulative investment earnings, and confidence ratings from the last trials. We also control for the block-fixed effect in the regression and cluster robust standard errors at the block level. The regression equation is as follows:

$$\begin{aligned} IsStockChoice_{t,b} = {} & \beta_1 Emorating_{t,b} + \beta_2 IsStock_{t-1,b} \\ & + \beta_3 SubjProb_{t-1,b} + + \beta_4 InvPayoff_{t,b} \\ & + \beta_5 Confid_{t-1,b} + \delta_b + \varepsilon_{i,b} \end{aligned} \tag{4}$$

[**Insert Table 4 near here**]

As shown in Table 4, the emotion ratings of images are significantly related to the subject's investment choices. The regression coefficient in column 4 is 0.0581 (t-statistic 9.99), suggesting that when a one-standard deviation increase in emotion rating increases, the probability of choosing a stock is 9.17%. This result is robust after controlling for the subject's expectations as well as its realized earnings, since the magnitude of regression coefficients is comparable across different columns. In the appendices, we replicate Kuhnen and Knutson (2011) with the original regression specification, and the results in Table B1 are similar. Moreover, we alternatively use probit regressions in B2 for further tests, and the result is even more significant. Among the control variables, the stock choice from the last trial $IsStockChoice_{t-1,b}$ is mainly correlated with the stock choice in the current trial, but the correlation is significantly negative. This may imply that the subject may be contrarian traders instead of momentum traders.

Their investment choices that are affected by emotion shocks are irrational, as shown in Figure 4. The blue line is the subject's average payoff from trial #1 to trial #6, the red line is the counterfactual payoff computed using the subject's probability estimation from the last trial, that is, the investment earnings made if the subject makes decisions based on its prior beliefs [16], and the green line is the fully rational earnings (the benchmark), that is, the investment earnings made if the subject makes decisions based on Bayesian objective probabilities. As shown in Figure 4, the red line is always better than the other in the sense of first-order dominance, implying that the observed investment decisions are inferior ones. This is not because emotional shocks induce the subject to take less risk and have lower earnings. Instead, the pattern in figure 3 shows that the observed stock choice suggests the subject is always more risk-loving than the counterfactual choices, since the subject always chooses to invest more in the stocks. In addition, the red line almost correlates with the green line, indicating that the investment choices implied with beliefs are reasonable.

[**Insert Figure 4 near here**]

We also test the in-sample robustness and heterogeneity of the investment choice task. We first examine the in-sample robustness of the subject's stock choice in table 5. In columns (1) and (2), we divide the samples according to the objective probability of the current trial. The first column represents trials where it is unlikely that the stock will pay dividends from good distribution, where $ObjProb_{t,b} < 0.2$. In contrast, the second column represents the trials where $ObjProb_{t,b} > 0.8$. The regression coefficients of $EmoRating_{t,b}$ are both significantly positive, and the economic magnitude is comparable to each other and similar to the results in table 4. In columns (3) and (4), we focus on early trials with trial number #1 to #3 and late trials with trial number #4 to #6. For early trials, the regression coefficient is 0.0384, which is less than for late trials, which have a regression coefficient of 0.0551. This suggests that GPT is less likely to be affected by emotion in the earlier stage of the experiment. In columns (5) and (6), we focus on subsamples where stocks have high payoffs and low payoffs in the trial $t-1$ (the last trial), and the regression coefficients are also significantly positive. In general, the regression coefficients are significantly positive.

[**Insert Table 5 near here**]

Next, we divide the samples by the topic of the images. The images have five categories: weather (including pollution), terrorism, sports, financial markets, and others.

[16] The investment payoff is normalized at the first trial because calculating the counterfactual payoff requires the subjective probability estimation from the last trial

The results are shown in Table 6. For images of weather, terrorism, financial markets, and others, positive emotions always induce the subject to invest more in stocks. This effect is pronounced mainly in financial markets, with a regression coefficient of 0.0963 and a t-statistic of 8.15, 65.74% higher than the effect of the baseline regressions. This result is not surprising, as the experiment takes place in a financial context. However, this effect is not significant for images in the sports topic. This may be because, although these images have different emotion ratings, GPT generally seems to have an optimistic view of sports.

[**Insert Table 6 near here**]

### *3.2. Beliefs*

Even though emotional shocks affect the subject's trading decisions, and yet, we find that they do not significantly impact their subjective probability estimations. The results are shown in figure 5, which plots the average subjective probability estimation that the stock pays from the good dividend distribution in five emotion groups. In subfigure A, we plot the average value of subjective probability estimation. The x-axis is the emotion group by ratings (from negative to positive), and the y-axis is the average subjective probability. The subfigure shows that, for all five groups, the subjective probability is around 0.50 with very low variation. A fitted linear regression line shows a very low regression slope and zero R-square. This preliminary result suggests that emotional shock does not have a significant impact on the subject's beliefs.

In subfigure B, we plot the subject's probability estimation relative to the objective Bayesian probability. The 45-degree dashed line serves as the rational benchmark, as it aligns the subject's estimation with the probability estimation calculated using the Bayesian formula. The colored lines denote the grouped probability estimation by their emotion rating in the current test.

[**Insert Table 5 near here**]

As shown in Figure 5, there is no significant difference between the subjective probability estimation in each group, especially in both tails. On average, subjects make higher subjective estimations when the objective estimation is low and lower subjective estimations when the objective estimation is high. This result is very similar to the experimental results in human subjects (Kuhnen, 2015; Kuhnen and Knutson, 2011; Kuhnen and Miu, 2017) (and also simlar when we do not show images to the subject in Figure B2), as humans also seem to be overly optimistic in the regime of "loss" and pessimistic in the "gain" regime, as summarized as the "four-fold patterns" predicted by prospect theory (Kahneman and Tversky, 2013; Oprea, 2024).

We show that there is also no significant relationship with regressions as shown in equation 5. The dependent variable is the subjective probability estimation of the subject $SubjProb_{t,b}$, and the independent variable of interest is the emotion rating of the image in the trial $t$ of block $b$. We control for the subject's investment decision, the objective probability, a binary variable that indicates whether the stock has a high dividend payoff, the cumulative investment payoff, and the confidence rating from the last trial. Furthermore, following Kuhnen and Knutson (2011), we control for $BayPriorsProb_{t,b}$ as an alternative for $ObjProb_{t,b}$ in columns (3) and (4). This new variable is derived from the subject's probability estimation from the last trial with the Bayesian rule, allowing us to disentangle the "learning effect" in trial $t$ from the "memory effect"[17]. Compared to Bayesian objective probability, this measure better describes the subject's "rational" belief estimation across trials. In addition to the control variables, we also control for block-fixed effects and cluster robust standard errors at the block level. The results are shown in Table 7.

$$\begin{aligned} SubjProb_{t,b} = {} & \beta_1 Emorating_{t,b} + \beta_2 IsStock_{t,b} \\ & + \beta_3 ObjProb_{t,b} + \beta_4 BayPriorsProb_{t,b} \\ & + \beta_5 IsHiPayoff_{t-1,b} + \beta_6 InvPayoff_{t,b} \\ & + \beta_7 Confid_{t-1,b} + \delta_b + \varepsilon_{i,b} \end{aligned} \tag{5}$$

**[Insert Table 7 near here]**

Regression results confirm that the subject's posterior belief is not associated with emotion shocks. In columns (1) and (2), the regression coefficients of $EmoRating_{t,b}$ are 0.0004 and 0.0004 with t-statistics of 0.43 and 0.48. On the other hand, the coefficients of $ObjProb_{t,b}$ are significantly positive. In columns (3) and (4), the regression coefficients of $BayPriorsProb_{t,b}$ are also significantly positive, and the magnitude is higher in column (4) with a regression coefficient of 1.2836 and t-statistic of 30.98. This coefficient suggests that GPT, on average, tends to overreact to new dividend realizations by overestimating the "good" probability regime and underestimating the "bad" probability regime. In addition, the subject's belief is highly correlated with its own confidence. The higher the confidence in the last trial $t-1$, the higher the probability estimate would be.

Not only in both regimes where the probability of the stock paying from the good distribution or bad distribution is obvious, but we also show that emotion shocks do not affect the subject's belief in ambiguous scenarios. We focus on sub-samples where

[17] Same as Kuhnen and Knutson (2011), $BayPriorsProb_{t,b}$ is calculated as follow: suppose the subjective probability estimation from the last trial is $p$, then the posterior belief obtained using Bayesian formula after observing a high stock dividend payoff is $3 \times p/(2 \times p + 2)$, and the $p/(3 - 2 \times p)$ after observing a low stock dividend payoff.

the subject has observed equal numbers of good and bad payoffs. For example, in the fourth trial, suppose the stock realized dividend has two high payoffs and two low payoffs; then, by experimental design, the Bayesian objective probability that the stock is good is 0.5, which mechanically makes the stock type ambiguous. This is also a good test that examines whether there is any systemic difference in the subject's priors Kuhnen (2015). We therefore focus on the second, fourth and sixth trials and report the results in table 8.

**[Insert Table 8 near here]**

We divided the trials by the sign of emotion ratings into a positive group and a negative group and reported the average subjective probability and standard deviations of each group. Univariate analysis shows that there is no systemic difference between these two groups. For the first two groups where the subject only observes the stock dividends of two trials (which include a high payoff of \$10 and a low payoff of -\$10), the average subjective probability estimation is 0.4838 and 0.4970, with standard deviations of 0.10 and 0.11. The probability difference between the two groups is 0.0132, and the t-statistic is 0.80. Moreover, the table shows that on average the subjective probability estimation is pessimistic because the mean probability is always lower than 0.50, and it is becoming more and more pessimistic over trials.

### *3.3. Confidence*

Next, we examine the subject's confidence. The confidence rating measures how the subject trusts their probability estimate. This confidence rating variable ranges from 0 to 10. We run regression 6, where the dependent variable is the confidence rating, and the independent variable of interest is the emotion ratings. We keep the other variables and regression specifications unchanged from Equation 5. The results are shown in Table 9.

$$\begin{aligned} Confid_{t,b} = &\beta_1 Emorating_{t,b} + \beta_2 IsStock_{t,b} \\ &+ \beta_3 ObjProb_{t,b} + \beta_4 BayPriorsProb_{t,b} \\ &+ \beta_5 IsHiPayoff_{t-1,b} + \beta_6 InvPayoff_{t,b} \\ &+ \beta_7 Confid_{t-1,b} + \delta_b + \varepsilon_{i,b} \end{aligned} \tag{6}$$

**[Insert Table 9 near here]**

The regression results show that emotion shocks do not affect the subject's confidence in its probability estimation. This is not surprising because emotional shocks do not have any significant effect on subjective probability estimation at first. On the other hand, this may also be correlated with the findings in Chen et al. (2024), which document that the measures of AI's declared confidence are opaque and structurally biased, indicating that LLMs cannot assess their confidence properly. Besides, as the result in the appendices B3 shows that emotional shocks do not significantly affect the subject's estimation error as well.

### *3.4. Complexity*

The experimental results from the investment choice task and the probability estimation task seem to contrast in two ways. First, decisions on the investment choice task are influenced by emotional shocks, whereas their beliefs on the probability estimation task are not. Previous research on associative memory documents that the cuing effect significantly leads investors to asymmetrically update their expectations, either through the direct new information channel or the indirect memory recall channel (Enke et al., 2024b). Secondly, the subject's investment decisions do not fully correlate with its prior beliefs. If the subject makes investment choices based on its subjective estimate that the stock is good from the last trial, it would make significantly higher investment payoffs as compared to the realized payoffs. The inconsistency between choices and estimations reflects a mismatch between observed risk preferences and beliefs, which is difficult to explain.

To reconcile this, we hypothesize that generative AI, just like humans, follows a "two-system" decision-making mindset, where it uses heuristic thinking in simple questions (fast thinking) and thoughtful thinking in complex questions (slow thinking). When prompted with simple questions such as a binary choice, fewer hidden layers and neural nodes are activated. When prompted with complex questions that require the model to implicitly use the Bayesian formula to update their beliefs and calculate numerical probabilities, the subject uses a more deliberate and analytical method to process information and make decisions. When making more thoughtful decisions, the subject's decision is more rational and exempt from the associative biases that are nothing but noise.

We cannot test the response time for slow-or-fast thinking because we are relying on OpenAI's API service, whose response time may vary significantly with the Internet connection or its computing usage. If the Internet connection is bad or its cloud computing service is at full capacity, it would mechanically have a slower response time, but this does not mean it is using a slow thinking mindset. Instead, we adopt an alternative approach to empirically test the "two-system" thinking hypothesis for GAI agents by simplifying the experimental question.

To do so, we substitute the original belief task, which requires the model to give a numerical probability estimation that the stock is paying from the good dividend distribution, to a binary choice task that requires the model to choose the stock type from two categories, "good" or "bad". The prompting question is as follows:

*"Now, estimate the stock type. Do you think it is good? Or is it bad? Your answer must be from one of the two options "good" or "bad". Reply only with the two options. Your answer is: "*

With this binary choice task, the complexity of decision-making is significantly reduced. Probability estimation is not required; choose only from the two options. In particular, the experimental setting of Kuhnen and Knutson (2011) only allows us to convert the task from complex to simple for the belief task, but it does not allow us to make the simple task more complex for the complex task. This is because, in each learning block that consists of six consecutive trials, the experiment needs input from the investment decision of the last trial to compute the investment payoff for the next trial. Still, it will be very interesting to test whether there are asymmetric responses when a task becomes simple or complex.

Here we define a new variable $IsGoodStockProb$ that indicates whether the subject thinks that the stock is paying from the good dividend distribution in this learning block and repeat the experiment with 300 new independent learning blocks, and the new results are shown in table 10.

[**Insert Table 10 near here**]

In the first and second columns where we only include the emotion rating, the stock choice dummy, and the objective probability, the regression coefficients in front of *EmoRating* are positive but insignificant. However, once we control for the binary variable that indicates whether this trial has high dividend payoffs, as well as other control variables, in the third and fourth columns, the regression coefficients become significantly positive. The regression coefficient in column four is 0.0066 with a t-statistic of 2.91, suggesting that emotional shock may have a significant impact on the subject's belief that the stock is paying from the good probability distribution. This shows that, when the task becomes simpler, the subject's decision making is more easily influenced by associated biases. However, the economic magnitude of this regression coefficient is not huge, only one-tenth of the size of the economic magnitude of that on the investment choices. Moreover, its belief seems irrelevant to its cumulative investment return and its confidence rating from the last trial, contradicting the findings in table 7.

In table 11, we show that the subject's beliefs in a simplified experimental task are largely determined (and distorted) by the realized dividends in the current trial and are not relevant to the Bayesian objective beliefs at all. In Panel A, we report the average probability that the subject chooses the type of stock as "good" instead of "bad" under different objective probability groups. We first group the average choice probability by emotion rating and realized payoff type in the current trial, and report the probability difference between the probability to choose the "good" type and the Bayesian objective probability, as well as the t-statistics between the average choice probability and the objective probability. In panel B, we report the average subjective probability and group the subjective probability by emotion rating and realized payoff type in the current trial, also reporting the probability difference and t-statistics.

[**Insert Table 11 near here**]

The univariate result in panel A shows that the subject's average choice in column three is largely dependent on the realized stock dividend in column two, regardless of the objective probability. When the current trial pays low payoff, the subject almost always chooses to think the stock is paying from the bad dividend payoff distribution, as opposed to it always choosing to think the stock is paying from the good dividend payoff dividend distribution once it observes a high dividend payoff.

Contrarily, in Panel B where we report the average subjective probability in the original complex task setting, the average subjective probability is significantly higher than 0 in low payoff trials and lower than 1 in high payoff trials. Moreover, its absolute probability difference between the objective probability estimation is around 0.10, suggesting that it is not entirely driven by the realized payoffs in the current trial. In other words, when the experimental question is more complex, the subject makes more reasonable answers by updating more correctly compared to the beliefs in the simpler task settings.

However, there is another unexpected consequence in the simplified question setting. In figure 6, we show that the realized payoff almost correlates with the belief-implied payoffs, and the difference between the fully rational earnings significantly lower than the earnings difference in figure 4. The average realized payoff after six trials for the observed payoff is 13.95, significantly higher than the 8.15 in the original studies. The large increase in cumulative payoffs may imply another interesting message: The subject is making investment decisions with a fixed amount of cognitive resources, where completing each task depletes a little. If the subject's cognitive resource is depleted extensively in a complex task, it has little resources to work on the simpler task, which generally leads to lower overall payoffs. On the other hand, when the subject is faced with two simple tasks and can allocate cognitive resources equally and efficiently, it just may make better economic decisions.

# 4. Conclusion

Exploiting a novel experiment setting, this paper shows that AI's decision under risk is largely affected by their associative memories. When cued with images with positive emotions, GAI will choose to invest more in stocks rather than bonds. In contrast, when cued with images with negative emotions, GAI will choose to invest more in bonds. However, their probability predictions that the stock is good, or paying from the good dividend payoff distribution, are not affected at all, implying that GAI might adopt a "two-system" decision-making mindset, where heuristic thinking dominates when it is asked to complete simple tasks, which renders it vulnerable shocks, yet deliberate thinking is practiced when it completing more complex tasks. We empirically test this by substituting the complex task with a simplified task that requires the subject to make a binary choice from two options instead of giving numerical probability estimates. In the simplified task, the subject's answers are significantly easier to be affected by emotional shocks and are largely determined by the observed stock payoffs in the current trial, failing to update belief correctly.

The findings of this paper have important implications for understanding GAI as economic agents. As GAI systems increasingly serve as decision-making tools in financial markets and other economic contexts, understanding their behavioral patterns and potential biases becomes crucial. While GAI demonstrates remarkable capabilities in complex tasks, their susceptibility to associative biases in simpler decisions suggests that careful consideration is needed when deploying these systems in real-world applications. Future research could explore how to mitigate these biases or leverage them constructively in economic decision-making processes. Furthermore, our experimental framework demonstrates the potential of using GAI as experimental subjects in economic research, offering a cost-effective and scalable approach to studying economic behavior. As GAI continues to evolve and integrate into various aspects of economic activities, understanding their decision-making mechanisms will become increasingly vital for both theoretical research and practical applications in finance and economics.

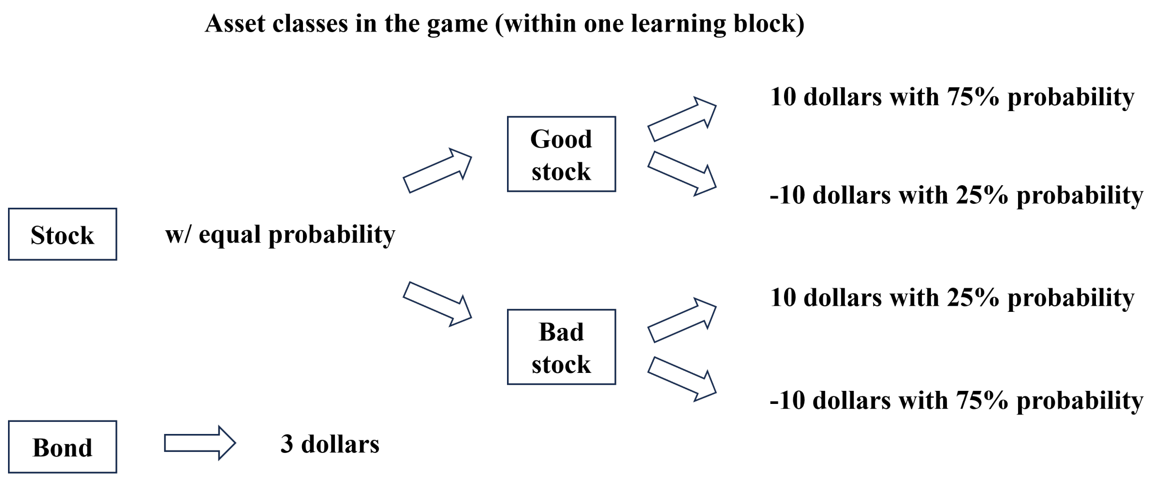
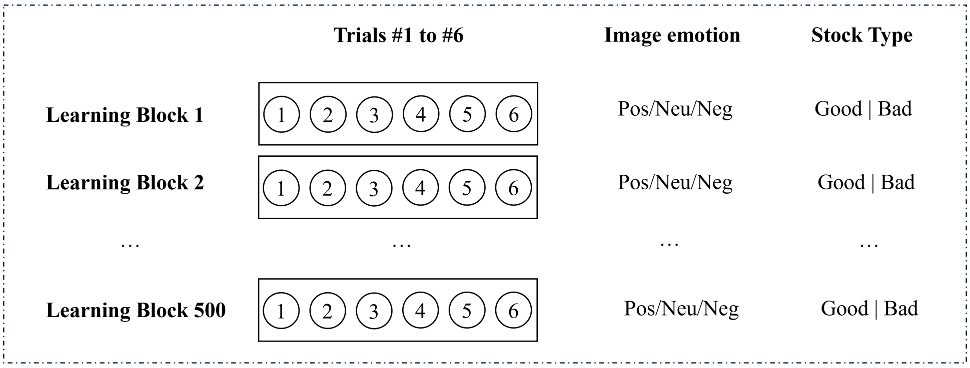
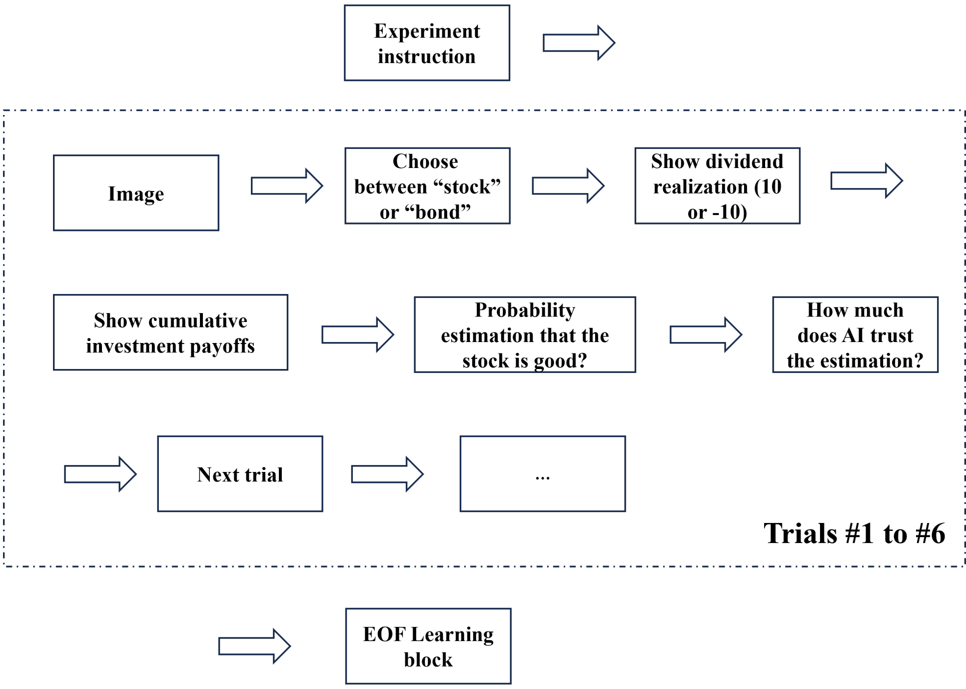



Fig. 1. This figure illustrates the asset payoff structures. In this experiment, there are two types of assets, including a bond and a stock. The bond always pays off \$3. The stock has an equal probability of paying from either a good distribution or a bad distribution. For good distribution, the stock has 75% to pay \$10, and 25% to pay -\$10. For the bad distribution, the stock has 25% to pay \$10, and 75% to pay -\$10.




**Subfigure A: Experiment overview**




**Subfigure B: Experiment sequence**

Fig. 2. These two figures illustrate the experiment design. In subfigure A, we show the experiment overview: the subject (GPT-4o-mini) goes through 500 independent learning tasks. Each learning task consists of 6 trials. In each trial, before the subject is asked to make financial decisions or probability estimations, it is shown with images that can have positive, neutral, or negative emotions. Within each learning block, the stock type is determined before the first trial and does not change over the six trials. In subfigure B, we show the experiment sequence. The subject is first shown with a detailed experiment instruction, then within each trial, the subject is presented with an image and asked to make investment decisions, then, the subject is shown the stock dividends and realized investment payoffs. Subsequently, it is required to estimate the probability that the stock is good and how much it trusts its estimation, and this trial is over. Importantly, within a learning block, the subject is allowed to keep the chat history, including all the instructions, choices, and investment payoffs. After a learning block is finished, its chat history is refreshed, and a new learning block is started.

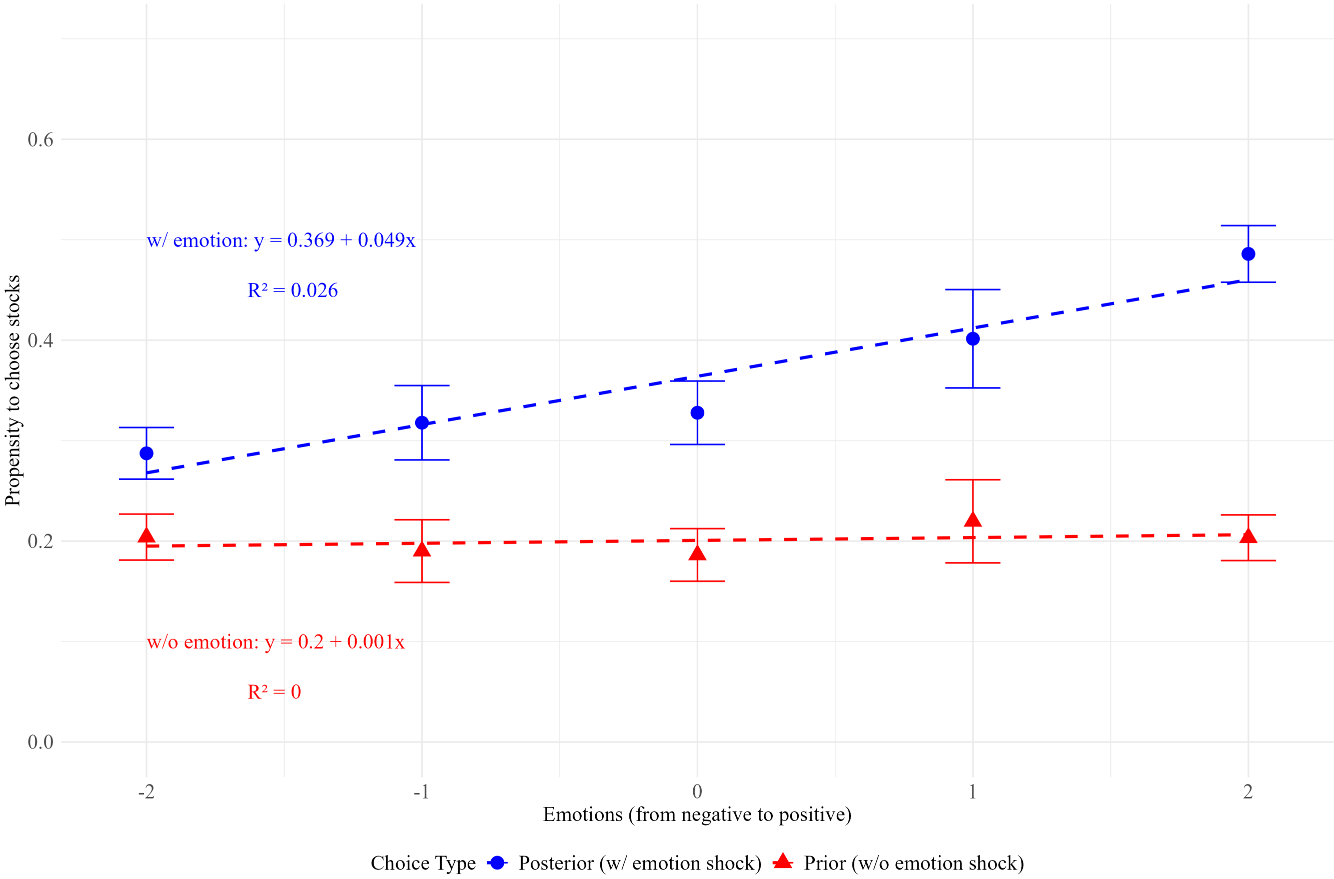



Fig. 3. Investment choices and emotional shocks. This figure plots the subject's investment choices across different emotion rating groups. The x-axis is the emotion rating of the image in each trial $t$ of block $b$ that ranges from -2 to +2, and the y-axis is the probability that the subject chooses to invest in stocks which ranges from 0 to 1. The blue dots denote the posterior stock choice probability or the observed subject's investment choice ex-post images. The red dots are the "hypothetical" prior stock choice probability computed from the subject's belief from the last trial. We fit linear trends for both groups and report regression statistics.

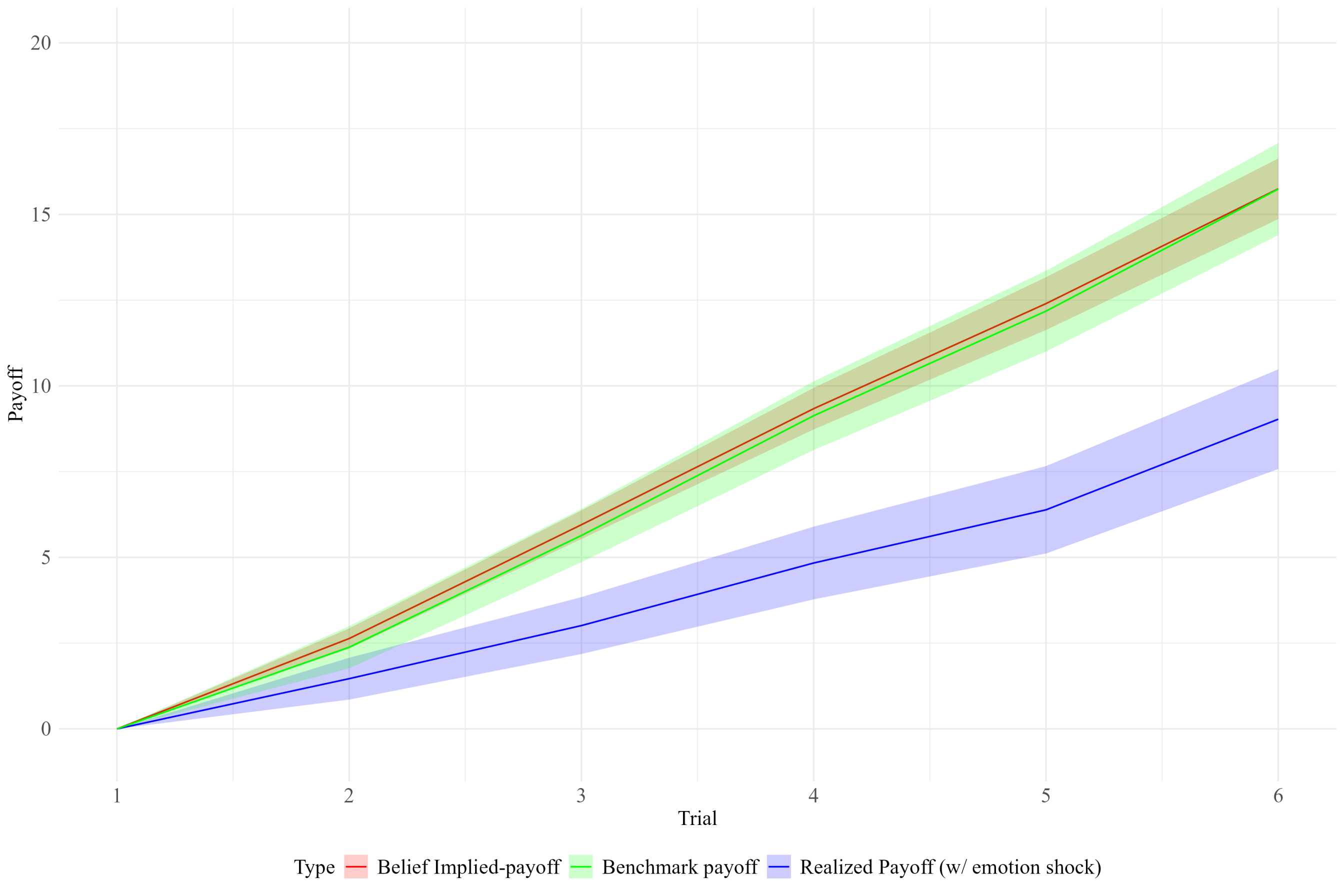
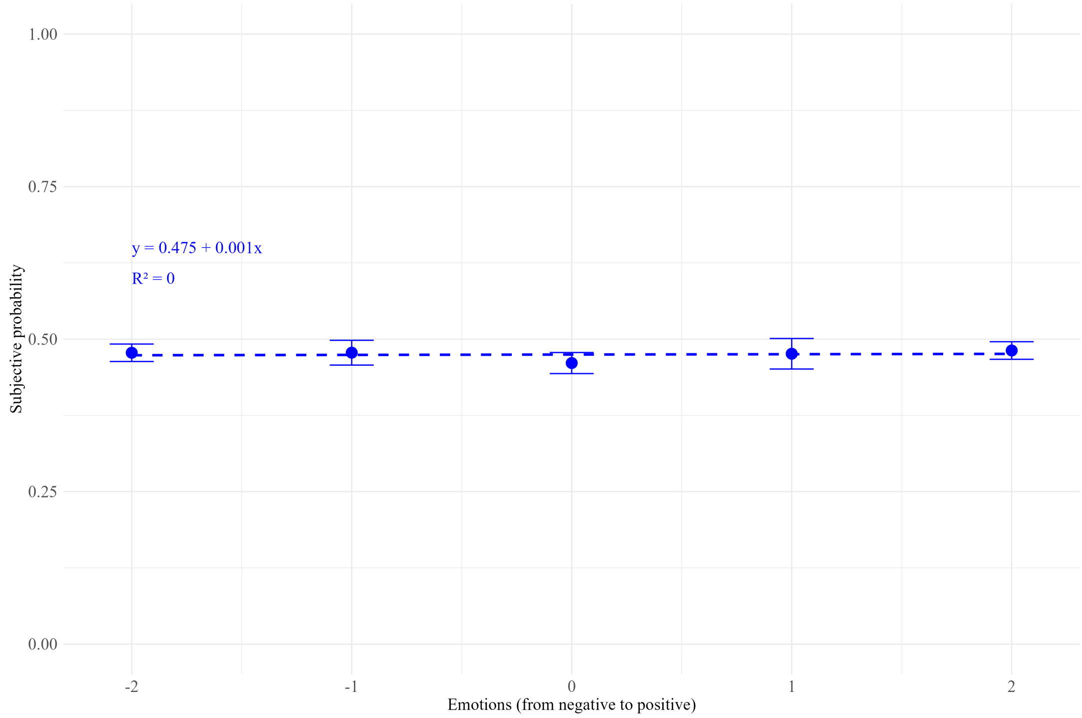
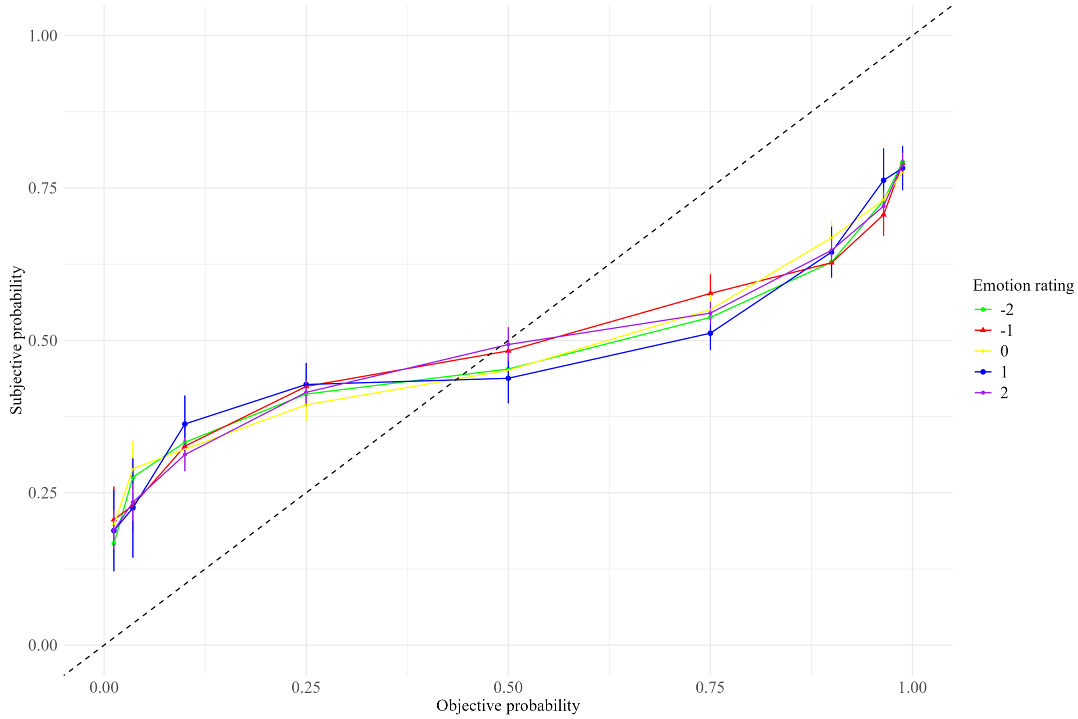



Fig. 4. Investment choices portfolio choices. This figure plots the subject's average investment payoffs. The blue line is the observed (realized) subject's average payoff from trial #1 to trial #6, the red line is the counterfactual payoff computed using the subject's probability estimation from the last trial, and the green line is the fully-rational earnings (the benchmark), computed with the Bayesian objective probabilities.




**Subfigure A: Probability estimation and emotion shocks.**




**Subfigure B: Subjective estimation vs. Objective estimation**

Fig. 5. Beliefs and emotional shocks. In subfigure A, we plot the average value of the subject's probability estimation across different emotion rating groups. The x-axis is the emotion group from negative to positive, and the y-axis is the average subjective probability. The confidence interval is at 95% for each group. We also fit a linear trend and report regression statistics. In subfigure B, we plot the subject's probability estimation over the Bayesian probability estimation. The x-axis denotes the Bayesian objective probability the stock pays from the good dividend distribution. The y-axis denotes the average subjective probability estimation. The 45-degreed dashed line serves as the rational benchmark

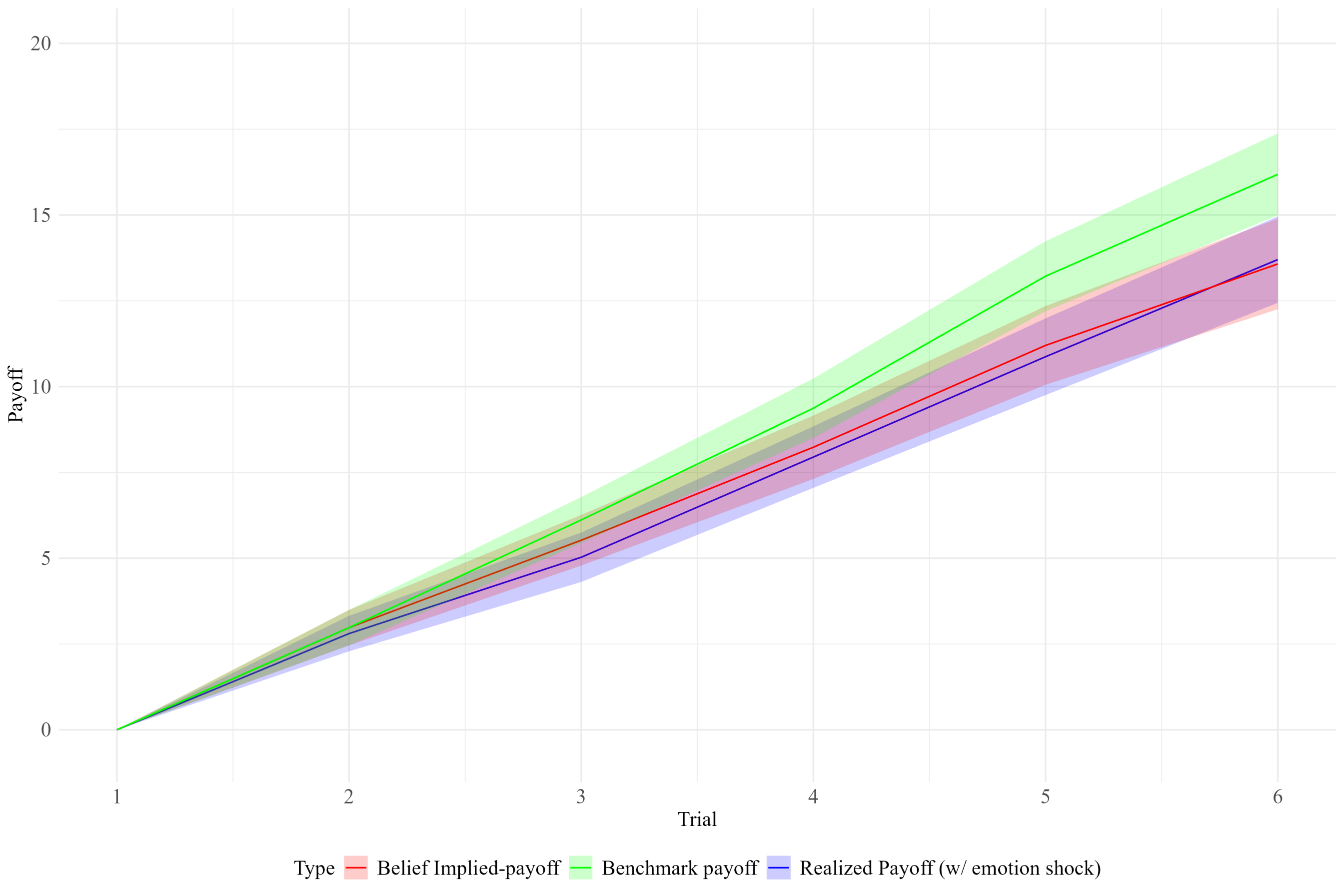



Fig. 6. Investment earnings in the simplicity setting. This figure plots the subject's average investment payoffs in the simplicity setting, where in the second question of each trial we give a "simpler" question to the subject to choose from. The blue line is the observed (realized) subject's average payoff from trial #1 to trial #6, the red line is the counterfactual payoff computed using the subject's belief of the stock from the last trial, and the green line is the fully-rational earnings (the benchmark), computed with the Bayesian objective probabilities.

## Table 1: Summary statistics of emotion rating

| Panel A: Emotion rating by machines | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Machine | | | | | |
| Topic | N | Mean | Sd | Min | Q1 | Med | Q3 | Max |
| Financial Markets | 94 | -0.21 | 1.72 | -2.00 | -2.00 | 0.00 | 2.00 | 2.00 |
| Sports | 143 | -0.07 | 1.68 | -2.00 | -2.00 | -1.00 | 2.00 | 2.00 |
| Terrorism | 59 | -0.29 | 1.79 | -2.00 | -2.00 | -1.00 | 2.00 | 2.00 |
| Weather | 207 | -0.17 | 1.55 | -2.00 | -2.00 | -1.00 | 1.00 | 2.00 |
| Others | 188 | 0.15 | 1.26 | -2.00 | 0.00 | 0.00 | 1.00 | 2.00 |

| Panel B: Emotion rating by human | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Topic | N | Mean | Sd | Min | Q1 | Med | Q3 | Max |
| Financial Markets | 94.00 | -0.43 | 1.61 | -2.00 | -2.00 | -1.06 | 1.19 | 2.00 |
| Sports | 187.00 | -0.03 | 1.00 | -2.00 | -0.11 | 0.00 | 0.06 | 2.00 |
| Terrorism | 143.00 | -0.40 | 1.24 | -1.89 | -1.44 | -1.00 | 0.83 | 1.89 |
| Weather | 59.00 | -0.49 | 1.60 | -2.00 | -2.00 | -1.22 | 0.89 | 4.00 |
| Others | 207.00 | -0.64 | 1.26 | -2.00 | -1.78 | -1.11 | 0.28 | 1.89 |

| Panel C: Correlation coeeficient by topics | | | | | | |
|---|---|---|---|---|---|---|
| | Pearson | | Spearman | | Kendall | |
| Topic | Correlation | P-value | Correlation | P-value | Correlation | P-value |
| Financial Markets | 0.89 | 0.00 | 0.87 | 0.00 | 0.75 | 0.00 |
| Sports | 0.88 | 0.00 | 0.86 | 0.00 | 0.76 | 0.00 |
| Terrorism | 0.90 | 0.00 | 0.89 | 0.00 | 0.77 | 0.00 |
| Weather | 0.91 | 0.00 | 0.89 | 0.00 | 0.79 | 0.00 |
| Others | 0.88 | 0.00 | 0.88 | 0.00 | 0.76 | 0.00 |

This table reports the emotion rating of images used in this experiment. Panel A reports summary statistics of the emotion rating by GPT-4o-mini. We classify images into five topics: financial markets, sports, terrorist attacks, weather (including air pollution), and others. Similarly, in panel B, we report the emotion rating by humans. For each image, the emotion ratings are first surveyed on 10 human subjects, and we then take the average value of the emotion ratings. In panel C, we report the correlation coefficients of the emotion ratings by GPT and humans. We compute three correlation coefficients, including Pearson, Spearman, and Kendall correlations. We also report the P-values for each correlation coefficient.

Table 2: Summary statistics of emotion rating

| | N | Mean | Sd | Min | Q1 | Med | Q3 | Max |
|---|---|---|---|---|---|---|---|---|
| IsStockChoice | 3000 | 0.37 | 0.48 | 0 | 0 | 0 | 1 | 1 |
| SubjProb | 3000 | 0.48 | 0.21 | 0.10 | 0.30 | 0.50 | 0.60 | 0.85 |
| ObjProb | 3000 | 0.49 | 0.36 | 0.01 | 0.10 | 0.50 | 0.90 | 0.99 |
| IsHiPayoff | 3000 | 0.48 | 0.50 | 0 | 0 | 0 | 1 | 1 |
| InvPayoff | 3000 | 7.25 | 11.02 | -10 | -1 | 6 | 15 | 30 |
| Confid | 3000 | 5.72 | 1.72 | 3 | 5 | 6 | 7 | 9 |
| EmoRating | 3000 | -0.05 | 1.58 | -2 | -2 | 0 | 2 | 2 |

This table reports the summary statistics of the experiment at the trial level. $IsStockChoice$ denotes whether the subject chooses to invest in the stock in this trial. $SubjProb$ denotes the subjective probability estimation. $ObjProb$ denotes the Bayesian objective probability estimation from this trial. $IsHiPayoff$ denotes whether the stock has realized a high dividend payoff ($10) in this trial. $InvPayoff$ denotes the subject's cumulative investment payoff. $Confid$ denotes the subject's confidence in its probability estimation. $Emorating$ is the emotion rating that appeared in the trial.

## Table 3: Validity test

| Panel A: Trading choice | | | | |
|---|---|---|---|---|
| Dep. Var. | IsStockChoice | | | |
| | (1) | (2) | (3) | (4) |
| SubjProbLst | 0.2537*** | | | |
| | (3.54) | | | |
| InvPayoffLst | | 0.0097*** | | |
| | | (9.48) | | |
| ConfidLst | | | 0.0408*** | |
| | | | (5.99) | |
| IsHiPayoffLst | | | | 0.0.0898*** |
| | | | | (4.08) |
| IsStockLst | -0.6835*** | -0.6825*** | -0.6945*** | -0.6804*** |
| | (-54.56) | (-54.95) | (-57.16) | (-53.13) |
| R2 | 0.492 | 0.505 | 0.498 | 0.494 |
| Block FE | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 2500 | 2500 | 2500 | 2500 |
| **Panel B: Belief formation** | | | | |
| Dep. Var. | SubjProb | | ProbUpdate | |
| | (1) | (2) | (3) | (4) |
| #HiPayoff | 0.0861*** | | | |
| | (13.30) | | | |
| #Trial | -0.0530*** | -0.0122*** | | |
| | (-13.81) | (-8.17) | | |
| InvPayoff | | 0.0014*** | | |
| | | (4.00) | | |
| IsHiPayoff | | | 0.2947*** | 0.2803*** |
| | | | (58.48) | (50.48) |
| IsHiPayoffLst | | | | -0.0282*** |
| | | | | (-7.48) |
| SubjProbLst | -0.0715*** | 0.0077 | | |
| | (-4.97) | (0.38) | | |
| ObjProb | 0.4095*** | 0.6316*** | -0.1549*** | -0.1072*** |
| | (22.11) | (58.49) | (-13.19) | (-7.36) |
| R2 | 0.951 | 0.939 | 0.803 | 0.808 |
| Block FE | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 2500 | 2500 | 2500 | 2500 |
| **Panel C: Confidence Level** | | | | |
| Dep. Var. | Confid | | | |
| | (1) | (2) | (3) | (4) |
| InvPayoff | 0.0901*** | | | |
| | (20.87) | | | |
| IsHiPayoff | | 2.1636*** | | |
| | | (55.58) | | |
| #HiPayoff | | | 0.4003*** | |
| | | | (15.34) | |
| IsGoodInvDec | | | | 1.3412*** |
| | | | | (27.82) |
| ConfidLst | 0.0161 | 0.3845*** | 0.0910*** | 0.2574*** |
| | (0.85) | (28.60) | (5.05) | (16.10) |
| R2 | 0.750 | 0.869 | 0.705 | 0.744 |
| Block FE | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 2500 | 2500 | 2500 | 2500 |

This table reports the experiment's validity. In panel A, the dependent variable is $IsStockChoice_{t,b}$, which denotes whether the subject chooses to invest in the stock in this trial. The control variables include the subjective probability estimation from the last trial, as well as the investment payoff, confidence rating, and investment decision from the last trial. In panel B, the dependent variable is $SubjProb_{t,b}$, which denotes the subject's probability estimation that the stock is good, and $ProbUpdate_{t,b}$ denotes the probability update over trials, which is the difference between $SubjProb_{t,b}$ and $SubjProb_{t-1,b}$. The independent variables include the total number of high dividend payoffs, the number of trials, the total cumulative investment payoff, the Bayesian objective probability, a binary variable that indicates whether the stock has a high dividend payoff in this trial, the subjective probability estimation from the last trial, and the objective probability in this trial. In Panel C, the dependent variable is the confidence rating $Confid_{t,b}$. The control variables include the total cumulative investment payoff, a binary variable that indicates whether this trial has a high payoff, the total number of high dividend payoffs, whether the subject made a profitable investment decision in the current trial, and the confidence rating from the last trial. In all the regressions, we control for block-fixed effect in the regression and cluster robust standard errors on the block level.

Table 4: Emotion shocks and investment choices

| Dep. Var. | IsStockChoice | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| EmoRating | 0.0527*** | 0.0596*** | 0.0620*** | 0.0581*** |
| | (7.52) | (10.01) | (7.97) | (9.99) |
| IsStockLst | | -0.6824*** | | -0.6946*** |
| | | (-55.58) | | (-56.38) |
| SubjProbLst | | | 0.2855*** | -0.5224*** |
| | | | (4.00) | (-3.85) |
| InvPayoffLst | | | | 0.0097*** |
| | | | | (6.41) |
| ConfidLst | | | | 0.0502*** |
| | | | | (4.37) |
| R2 | 0.079 | 0.517 | 0.107 | 0.537 |
| Block FE | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 3000 | 2500 | 2500 | 2500 |

This table reports the relationship between emotional shocks and the subject's investment choices. The dependent variable is a binary variable that indicates whether the subject chooses to invest in stock in the trial $IsStockChoice_{t,b}$. The independent variable of interest is the emotion rating of the image in trial $t$ of block $b$. We include other control variables such as stock choice from the last trial, subjective probability, cumulative investment earnings, and confidence ratings from the last trials. We also control for block-fixed effect in the regression and cluster robust standard errors on the block level.

Table 5: In-sample robustness tests

| Dep. Var. | IsStockChoice | | | | | |
|---|---|---|---|---|---|---|
| Sample | ObjPrb < 0.2 | ObjPrb > 0.8 | Early trials | Late trials | IsHiPayoffLst = 1 | IsHiPayoffLst = 0 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| EmoRating | 0.0496*** | 0.0591*** | 0.0384*** | 0.0551*** | 0.0577*** | 0.0635*** |
| | (5.62) | (5.31) | (3.24) | (7.56) | (6.28) | (7.77) |
| IsStockLst | -0.4866*** | -0.6853*** | -1.0462*** | -0.7742*** | -0.7774*** | -0.5278*** |
| | (-15.64) | (-25.94) | (-38.37) | (-47.49) | (-33.21) | (-18.49) |
| SubjProbLst | -0.9205*** | 0.8505*** | -1.9730*** | -0.1768 | 0.0670 | -0.6340*** |
| | (-4.31) | (2.85) | (-7.22) | (-0.76) | (0.24) | (-3.59) |
| InvPayoffLst | 0.0357*** | -0.0074*** | 0.0345*** | 0.0051* | 0.0024 | 0.0220*** |
| | (10.42) | (-3.00) | (8.18) | (1.95) | (0.77) | (7.12) |
| ConfidLst | 0.0247 | -0.0052 | 0.1778*** | 0.0180 | 0.0453** | 0.0072 |
| | (1.23) | (-0.24) | (6.97) | (1.02) | (2.34) | (0.42) |
| R2 | 0.657 | 0.629 | 0.658 | 0.692 | 0.636 | 0.650 |
| Block FE | ✓ | ✓ | ✓ | ✓ | | |
| Num.Obs. | 874 | 791 | 1000 | 1500 | 1213 | 1287 |

This table reports the in-sample robustness. The dependent variable is the subject's investment decision $IsStockChoice_{b,t}$. The independent variable of interest is the emotion rating of the image in trial $t$ of block $b$. We include other control variables such as stock choice from the last trial, subjective probability, cumulative investment earnings, and confidence ratings from the last trials. In columns (1) and (2), we split the samples based on the objective probability in the current trial. The first column represents trials where the stock is unlikely to be paying dividends from the good distribution, where $ObjProb_{t,b} < 0.2$. Contrarily, the second column represents trials where $ObjProb_{t,b} > 0.8$. In columns (3) and (4), we focus on the early trials with trial number #1 to #3 and late trials with trial number #4 to #6. In columns (5) and (6), we focus on subsamples where stocks have high payoffs and low payoffs in the trial $t-1$ (the last trial). We also control for block-fixed effect in the regression and cluster robust standard errors on the block level.

Table 6: Heterogeneity by different topics

| Dep. Var. | IsStockChoice | | | | |
|---|---|---|---|---|---|
| Topic | Weather | Terrorism | Sports | Financial Markets | Others |
| | (1) | (2) | (3) | (4) | (5) |
| EmoRating | 0.0421** | 0.0936*** | 0.0201 | 0.0963*** | 0.0601*** |
| | (2.55) | (2.88) | (0.86) | (8.15) | (3.42) |
| IsStockLst | -0.7596*** | -0.5224*** | -0.6709*** | -0.6535*** | -0.7126*** |
| | (-24.62) | (-5.01) | (-12.17) | (-18.08) | (-21.93) |
| SubjProbLst | -0.3969 | -1.1532 | -1.0371** | 0.2767 | -0.6222** |
| | (-1.10) | (-1.37) | (-2.50) | (1.06) | (-2.13) |
| InvPayoffLst | 0.0108*** | -0.0033 | 0.0177*** | 0.0058 | 0.0124*** |
| | (2.94) | (-0.23) | (4.08) | (1.52) | (3.71) |
| ConfidLst | 0.0226 | 0.1514** | 0.0536 | 0.0142 | 0.0513** |
| | (0.71) | (2.28) | (1.19) | (0.50) | (2.20) |
| R2 | 0.762 | 0.908 | 0.837 | 0.772 | 0.747 |
| Block FE | ✓ | ✓ | ✓ | ✓ | |
| Num.Obs. | 681 | 206 | 372 | 583 | 658 |

This table reports the heterogeneity across different topics. The dependent variable is the subject's investment decision $IsStockChoice_{b,t}$. The independent variable of interest is the emotion rating of the image in the trial $t$ of the block $b$. We include other control variables such as stock choice from the last trial, subjective probability, cumulative investment earnings, and confidence ratings from the last trials. We divide the samples by the topic of images such as weather (including pollution), terrorism, sports, financial markets, and others. We also control for the block-fixed effect in the regression and cluster robust standard errors at the block level.

Table 7: Emotion shocks and posterior beliefs

| Dep. Var. | SubjProb | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| EmoRating | 0.0004 | 0.0004 | 0.0007 | 0.0000 |
| | (0.43) | (0.48) | (0.61) | (0.06) |
| IsStock | -0.0033 | 0.0076*** | 0.0043* | 0.0013 |
| | (-1.42) | (4.30) | (1.74) | (1.19) |
| ObjProb | 0.6478*** | 0.3533*** | | |
| | (54.29) | (22.88) | | |
| BayPriorsProb | | | 0.3681*** | 1.2836*** |
| | | | (59.67) | (30.98) |
| IsHiPayoff | | 0.1120*** | | -0.5745*** |
| | | (22.82) | | (-22.08) |
| InvPayoff | | 0.0023*** | | 0.0004** |
| | | (7.78) | | (2.43) |
| ConfidLst | | 0.0219*** | | 0.0088*** |
| | | (13.73) | | (7.95) |
| R2 | 0.904 | 0.957 | 0.862 | 0.986 |
| Block FE | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 3000 | 2500 | 3000 | 2500 |

This table reports the relationship between emotion shocks and the subject's elicited posterior probability estimates. The dependent variable is the subject's subjective probability estimation $SubjProb_{t,b}$, and the independent variable of interest is the emotion rating of the image in trial $t$ of block $b$. We control for the subject's investment decision, the objective probability, a binary variable that indicates whether the stock has a high dividend payoff, the cumulative investment payoff, and the confidence rating from the last trial. Additionally, we control for the $BayPriorsProb_{t,b}$ as an alternative for $ObjProb_{t,b}$ in columns (3) and (4). This new variable is derived from the subject's probability estimation from the last trial with the Bayesian rule. Finally, we control for the block-fixed effect in the regression and cluster robust standard errors at the block level.

Table 8: Emotion shocks in ambiguous payoff regimes

| # Trial | Emotion | AvgSubjProb | StdSubjProb | N | ProbDiff | t-Stat |
|---|---|---|---|---|---|---|
| 2 | Positive | 0.4838 | (0.10) | 80 | 0.0132 | (0.80) |
| 2 | Negative | 0.4970 | (0.11) | 78 | | |
| 4 | Positive | 0.4390 | (0.11) | 36 | 0.0007 | (0.02) |
| 4 | Negative | 0.4397 | (0.13) | 34 | | |
| 6 | Positive | 0.4365 | (0.16) | 37 | 0.0293 | (0.58) |
| 6 | Negative | 0.4658 | (0.19) | 19 | | |

This table reports the subject's subjective probability estimation when the subject has observed equal numbers of high and low payoffs. We focus on the second, fourth, and sixth trials, where there is an equal occurrence of high and low payoffs. For each group, we report the average subjective probability estimation and its standard deviation. We compute the difference in probability estimation and report the t-statistic.

Table 9: Emotion shocks and estimation confidence

| Dep. Var. | Confidence | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| EmoRating | -0.0070 | -0.0002 | -0.0043 | -0.0019 |
| | (-0.58) | (-0.02) | (-0.39) | (-0.21) |
| IsStock | 0.2507*** | 0.2224*** | 0.3174*** | 0.1946*** |
| | (8.02) | (7.90) | (11.06) | (7.03) |
| ObjProb | 5.2624*** | 2.1401*** | | |
| | (46.51) | (12.00) | | |
| BaysProb | | | 3.4765*** | 6.5277*** |
| | | | (61.71) | (14.83) |
| IsHiPayoff | | 1.5133*** | | -1.9117*** |
| | | (25.51) | | (-6.93) |
| InvPayoff | | 0.0295*** | | 0.0211*** |
| | | (9.14) | | (7.00) |
| ConfidLst | | 0.1786*** | | 0.1288*** |
| | | (10.54) | | (7.23) |
| R2 | 0.778 | 0.893 | 0.797 | 0.902 |
| Block FE | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 3000 | 2500 | 3000 | 2500 |

This table reports the relationship between the subject's confidence rating and emotion shocks. The dependent variable is the confidence rating $Confid_{t,b}$, and the independent variable of interest is the emotion rating. We control for the subject's investment decision, the objective probability, a binary variable that indicates whether the stock has a high dividend payoff, the cumulative investment payoff, and the confidence rating from the last trial. Additionally, we control for the $BayPriorsProb_{t,b}$ as an alternative for $ObjProb_{t,b}$ in columns (3) and (4). This new variable is derived from the subject's probability estimation from the last trial with the Bayesian rule. Finally, we control for the block-fixed effect in the regression and cluster robust standard errors at the block level.

Table 10: Emotional shocks and subjective beliefs in a simplified setting

| Dep. Var. | IsGoodStockProb | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| EmoRating | 0.0037 | 0.0020 | 0.0050*** | 0.0066*** |
| | (0.55) | (0.34) | (2.79) | (2.91) |
| IsStock | | 0.0275 | 0.0038 | 0.0054 |
| | | (1.57) | (0.72) | (0.86) |
| ObjProb | | 1.2682*** | 0.0593*** | 0.0963*** |
| | | (28.36) | (3.14) | (2.82) |
| IsHiPayoff | | | 0.9524*** | 0.9343*** |
| | | | (75.82) | (56.15) |
| IsGoodStockLst | | | | -0.0210** |
| | | | | (-2.44) |
| InvPayoff | | | | 0.0005 |
| | | | | (1.01) |
| ConfidLst | | | | -0.0041 |
| | | | | (-1.20) |
| R2 | 0.375 | 0.519 | 0.969 | 0.965 |
| Block FE | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 1800 | 1800 | 1800 | 1500 |

This table reports the relationship between emotional shocks and the subject's elicited posterior probability estimates in a simplified task setting, where we only tell the model to choose between two types about the stock, "good" or "bad", instead of letting it to give numerical probability estimations. The dependent variable is the subject's subjective choice that whether it thinks the stock is good or bad, and the independent variable of interest is the emotion rating of the image in trial $t$ of block $b$. We control for the subject's investment decision in the same trial, the objective probability, a binary variable that indicates whether the stock has a high dividend payoff in this trial, the subject's belief choice from the last trial, the cumulative investment payoff, and the confidence rating from the last trial. Finally, we control for the block-fixed effect in the regression and cluster robust standard errors at the block level.

Table 11: Realized dividend and subjective beliefs in different settings

| Panel A: Simplified question | | | | | |
|---|---|---|---|---|---|
| EmoRating | IsHiPayoff | IsGoodStockProb | ObjProb | ProbDiff | t-Stat |
| -2.00 | 0.00 | 0.00 | 0.27 | -0.27 | (-15.462) |
| -1.00 | 0.00 | 0.00 | 0.26 | -0.26 | (-11.709) |
| 0.00 | 0.00 | 0.00 | 0.24 | -0.24 | (-12.569) |
| 1.00 | 0.00 | 0.00 | 0.25 | -0.25 | (-8.922) |
| 2.00 | 0.00 | 0.02 | 0.27 | -0.25 | (-11.634) |
| -2.00 | 1.00 | 0.98 | 0.75 | 0.23 | (12.511) |
| -1.00 | 1.00 | 0.97 | 0.76 | 0.21 | (7.624) |
| 0.00 | 1.00 | 1.00 | 0.73 | 0.27 | (12.013) |
| 1.00 | 1.00 | 1.00 | 0.78 | 0.22 | (7.737) |
| 2.00 | 1.00 | 0.99 | 0.74 | 0.25 | (13.663) |

| Panel B: Complex question | | | | | |
|---|---|---|---|---|---|
| EmoRating | IsHiPayoff | SubjProb | ObjProb | ProbDiff | t-Stat |
| -2.00 | 0.00 | 0.34 | 0.25 | 0.08 | (5.693) |
| -1.00 | 0.00 | 0.34 | 0.22 | 0.12 | (6.349) |
| 0.00 | 0.00 | 0.33 | 0.23 | 0.10 | (5.415) |
| 1.00 | 0.00 | 0.35 | 0.26 | 0.09 | (3.491) |
| 2.00 | 0.00 | 0.34 | 0.25 | 0.08 | (5.534) |
| -2.00 | 1.00 | 0.62 | 0.76 | -0.14 | (-9.639) |
| -1.00 | 1.00 | 0.62 | 0.73 | -0.11 | (-4.987) |
| 0.00 | 1.00 | 0.61 | 0.71 | -0.10 | (-5.207) |
| 1.00 | 1.00 | 0.63 | 0.76 | -0.14 | (-5.142) |
| 2.00 | 1.00 | 0.63 | 0.75 | -0.12 | (-8.147) |

This table reports the subject's subjective belief about the stock type in a simplified task in Panel A and a complex task (the original task) in Panel B. In Panel A, we report the average choice probability that the subject choose the stock type as "good" in stead of "bad". We group the average choice probability by emotion rating and realized payoff type in the current trial, and report the probability difference between the choice probability to choose the "good" type and the Bayesian objective probability, as well as the t-statistics. In Panel B, we report the average subjective probability and group the subjective probability by emotion rating and realized payoff type in the current trial, also reporting the probability difference and t-statistics.

# References


Acemoglu, D., 2024. The simple macroeconomics of ai. Tech. rep., National Bureau of Economic Research.

Armstrong, D. M., 2023. Measuring tax enforcement with generative ai. Available at SSRN 4639565 .

Arora, N., Chakraborty, I., Nishimura, Y., 2024. Express: Ai-human hybrids for marketing research: Leveraging llms as collaborators. Journal of Marketing p. 00222429241276529.

Ba, C., Bohren, J. A., Imas, A., 2024. Over-and underreaction to information. Available at SSRN 4274617 .

Babina, T., Fedyk, A., He, A., Hodson, J., 2024. Artificial intelligence, firm growth, and product innovation. Journal of Financial Economics 151, 103745.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. The journal of Finance 61, 1645–1680.

Berg, T., Fuster, A., Puri, M., 2022. Fintech lending. Annual Review of Financial Economics 14, 187–207.

Bini, P., Cong, L. W., Huang, X., Jin, L., 2024. Behavioral economics of ai: Llm biases and corrections. Working Paper .

Bordalo, P., Burro, G., Coffman, K., Gennaioli, N., Shleifer, A., 2024a. Imagining the future: memory, simulation, and beliefs. Review of Economic Studies p. rdae070.

Bordalo, P., Gennaioli, N., Lanzani, G., Shleifer, A., 2024b. A cognitive theory of reasoning and choice .

Bordalo, P., Gennaioli, N., Shleifer, A., 2020. Memory, attention, and choice. The Quarterly journal of economics 135, 1399–1442.

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

Chang, E. Y., 2024. Behavioral emotion analysis model for large language models. In: 2024 IEEE 7th International Conference on Multimedia Information Processing and Retrieval (MIPR), IEEE, pp. 549–556.

Charles, C., 2022. Memory and trading. Available at SSRN 3759444 .

Chen, H., Didisheim, A., Somoza, L., 2024. Out of the (black) box: Ai as conditional probability. Available at SSRN .

Chen, W., Wu, H., Zhang, L., 2021. Terrorist attacks, managerial sentiment, and corporate disclosures. The Accounting Review 96, 165–190.

Chen, Y., Kelly, B. T., Xiu, D., 2022. Expected returns and large language models. Available at SSRN 4416687 .

Chen, Y., Liu, T. X., Shan, Y., Zhong, S., 2023. The emergence of economic rationality of gpt. Proceedings of the National Academy of Sciences 120, e2316205120.

Cheong, I., Xia, K., Feng, K. K., Chen, Q. Z., Zhang, A. X., 2024. (a) i am not a lawyer, but...: Engaging legal experts towards responsible llm policies for legal advice. In: The 2024 ACM Conference on Fairness, Accountability, and Transparency, pp. 2454–2469.

Dehaan, E., Madsen, J., Piotroski, J. D., 2017. Do weather-induced moods affect the processing of earnings news? Journal of Accounting Research 55, 509–550.

Demszky, D., Yang, D., Yeager, D. S., Bryan, C. J., Clapper, M., Chandhok, S., Eichstaedt, J. C., Hecht, C., Jamieson, J., Johnson, M., et al., 2023. Using large language models in psychology. Nature Reviews Psychology 2, 688–701.

Dertwinkel-Kalt, M., Köster, M., 2024. Salient cues and complexity. Management Science .

Dong, R., Fisman, R., Wang, Y., Xu, N., 2021. Air pollution, affect, and forecasting bias: Evidence from chinese financial analysts. Journal of Financial Economics 139, 971–984.

D'Acunto, F., Prabhala, N., Rossi, A. G., 2019. The promises and pitfalls of robo-advising. The Review of Financial Studies 32, 1983–2020.

Eckel, C. C., Grossman, P. J., 2002. Sex differences and statistical stereotyping in attitudes toward financial risk. Evolution and human behavior 23, 281–295.

Edmans, A., Garcia, D., Norli, Ø., 2007. Sports sentiment and stock returns. The Journal of finance 62, 1967–1998.

Enke, B., Graeber, T., Oprea, R., 2023. Complexity and hyperbolic discounting. Tech. rep., National Bureau of Economic Research.

Enke, B., Graeber, T., Oprea, R., Yang, J., 2024a. Behavioral attenuation. Tech. rep., National Bureau of Economic Research.

Enke, B., Schwerter, F., Zimmermann, F., 2024b. Associative memory, beliefs and market interactions. Journal of Financial Economics 157, 103853.

Erel, I., Stern, L. H., Tan, C., Weisbach, M. S., 2021. Selecting directors using machine learning. The Review of Financial Studies 34, 3226–3264.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences. The quarterly journal of economics 133, 1645–1692.

Fedyk, A., Kakhbod, A., Li, P., Malmendier, U., 2024. Chatgpt and perception biases in investments: An experimental study. Available at SSRN 4787249 .

Felin, T., Holweg, M., 2024. Theory is all you need: Ai, human cognition, and decision making. Human Cognition, and Decision Making (February 23, 2024) .

Gabaix, X., Graeber, T., 2024. The complexity of economic decisions. Tech. rep., National Bureau of Economic Research.

Gennaioli, N., Shleifer, A., 2018. A crisis of beliefs: Investor psychology and financial fragility .

Gneezy, U., Potters, J., 1997. An experiment on risk taking and evaluation periods. The quarterly journal of economics 112, 631–645.

Goetzmann, W. N., Kim, D., Kumar, A., Wang, Q., 2015. Weather-induced mood, institutional investors, and stock returns. The Review of Financial Studies 28, 73–111.

Goetzmann, W. N., Kim, D., Shiller, R. J., 2024. Emotions and subjective crash beliefs. Tech. rep., National Bureau of Economic Research.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. The Review of Financial Studies 33, 2223–2273.

Handa, K., Gal, Y., Pavlick, E., Goodman, N., Andreas, J., Tamkin, A., Li, B. Z., 2024. Bayesian preference elicitation with language models. arXiv preprint arXiv:2403.05534 .

Hansen, A. L., Kazinnik, S., 2023. Can chatgpt decipher fedspeak. Available at SSRN .

Häusler, A. N., Kuhnen, C. M., Rudorf, S., Weber, B., 2018. Preferences and beliefs about financial risk taking mediate the association between anterior insula activation and self-reported real-life stock trading. Scientific reports 8, 11207.

He, Z., Guo, S., Rao, A., Lerman, K., 2024. Whose emotions and moral sentiments do language models reflect? arXiv preprint arXiv:2402.11114 .

Hecht-Nielsen, R., 1992. Theory of the backpropagation neural network. In: Neural networks for perception, Elsevier, pp. 65–93.

Heyes, A., Neidell, M., Saberian, S., 2016. The effect of air pollution on investor behavior: Evidence from the s&p 500. Tech. rep., National Bureau of Economic Research.

Hinton, G., 2022. The forward-forward algorithm: Some preliminary investigations. arXiv preprint arXiv:2212.13345 .

Hinton, G. E., et al., 1992. How neural networks learn from experience. Scientific American.

Hirshleifer, D., 2015. Behavioral finance. Annual Review of Financial Economics 7, 133–159.

Hirshleifer, D., Shumway, T., 2003. Good day sunshine: Stock returns and the weather. The journal of Finance 58, 1009–1032.

Holt, C. A., Laury, S. K., 2002. Risk aversion and incentive effects. American economic review 92, 1644–1655.

Homberg, J. R., 2012. Serotonin and decision making processes. Neuroscience & Biobehavioral Reviews 36, 218–236.

Horton, J. J., 2023. Large language models as simulated economic agents: What can we learn from homo silicus? Tech. rep., National Bureau of Economic Research.

Hu, M. R., Lee, A. D., 2020. Outshine to outbid: Weather-induced sentiment and the housing market. Management Science 66, 1440–1472.

Hurst, A., Lerer, A., Goucher, A. P., Perelman, A., Ramesh, A., Clark, A., Ostrow, A., Welihinda, A., Hayes, A., Radford, A., et al., 2024. Gpt-4o system card. arXiv preprint arXiv:2410.21276 .

Jha, M., Qian, J., Weber, M., Yang, B., 2024. Chatgpt and corporate policies. Tech. rep., National Bureau of Economic Research.

Jia, J., Yuan, Z., Pan, J., McNamara, P. E., Chen, D., 2024. Decision-making behavior evaluation framework for llms under uncertain context. arXiv preprint arXiv:2406.05972 .

Jiang, F., Lee, J., Martin, X., Zhou, G., 2019. Manager sentiment and stock returns. Journal of Financial Economics 132, 126–149.

Kahneman, D., 2011. Thinking, fast and slow. Farrar, Straus and Giroux .

Kahneman, D., Knetsch, J. L., Thaler, R. H., 1990. Experimental tests of the endowment effect and the coase theorem. Journal of political Economy 98, 1325–1348.

Kahneman, D., Tversky, A., 2013. Prospect theory: An analysis of decision under risk. In: Handbook of the fundamentals of financial decision making, World Scientific, pp. 99–127.

Kendall, C., Oprea, R., 2024. On the complexity of forming mental models. Quantitative Economics 15, 175–211.

Kim, A., Muhn, M., Nikolaev, V., 2023. From transcripts to insights: Uncovering corporate risks using generative ai. arXiv preprint arXiv:2310.17721 .

Kim, A., Muhn, M., Nikolaev, V., 2024a. Financial statement analysis with large language models. arXiv preprint arXiv:2407.17866 .

Kim, A., Muhn, M., Nikolaev, V. V., 2024b. Bloated disclosures: can chatgpt help investors process information? Chicago Booth Research Paper pp. 2023–59.

Knutson, B., Wimmer, G. E., Kuhnen, C. M., Winkielman, P., 2008. Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. NeuroReport 19, 509–513.

Kőszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. The Quarterly Journal of Economics 121, 1133–1165.

Kőszegi, B., Rabin, M., 2007. Reference-dependent risk attitudes. American Economic Review 97, 1047–1073.

Kőszegi, B., Rabin, M., 2009. Reference-dependent consumption plans. American Economic Review 99, 909–936.

Kuhnen, C. M., 2015. Asymmetric learning from financial information. The Journal of Finance 70, 2029–2062.

Kuhnen, C. M., Chiao, J. Y., 2009. Genetic determinants of financial risk taking. PloS one 4, e4362.

Kuhnen, C. M., Knutson, B., 2005. The neural basis of financial risk taking. Neuron 47, 763–770.

Kuhnen, C. M., Knutson, B., 2011. The influence of affect on beliefs, preferences, and financial decisions. Journal of Financial and Quantitative Analysis 46, 605–626.

Kuhnen, C. M., Miu, A. C., 2017. Socioeconomic status and learning from financial information. Journal of Financial Economics 124, 349–372.

Kuhnen, C. M., Samanez-Larkin, G. R., Knutson, B., 2013. Serotonergic genotypes, neuroticism, and financial choices. PloS one 8, e54632.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. nature 521, 436–444.

Leng, Y., 2024. Can llms mimic human-like mental accounting and behavioral biases? Available at SSRN 4705130 .

Leng, Y., Sang, Y., Agarwal, A., 2024. Reduce preference disparity between llms and humans: Calibration to compose human-like ensembles. Available at SSRN 4802019 .

Leng, Y., Yuan, Y., 2023. Do llm agents exhibit social behavior? arXiv preprint arXiv:2312.15198 .

Li, B., Rossi, A. G., 2020. Selecting mutual funds from the stocks they hold: A machine learning approach. Available at SSRN 3737667 .

Li, C., Wang, J., Zhang, Y., Zhu, K., Wang, X., Hou, W., Lian, J., Luo, F., Yang, Q., Xie, X., 2023. The good, the bad, and why: Unveiling emotions in generative ai. arXiv preprint arXiv:2312.11111 .

Li, F., Zheng, X., 2023. Carbon emission and asset prices: new evidence from machine learning. Available at SSRN 4400681 .

Li, J. J., Massa, M., Zhang, H., Zhang, J., 2021. Air pollution, behavioral bias, and the disposition effect in china. Journal of Financial Economics 142, 641–673.

Li, K., Mai, F., Shen, R., Yang, C., Zhang, T., 2024a. Dissecting corporate culture using generative ai–insights from analyst reports. Available at SSRN 4558295 .

Li, P., Castelo, N., Katona, Z., Sarvary, M., 2024b. Frontiers: Determining the validity of large language models for automated perceptual analysis. Marketing Science 43, 254–266.

Liu, Z., Wu, Z., Hu, M., Zhao, B., Zhao, L., Zhang, T., Dai, H., Chen, X., Shen, Y., Li, S., et al., 2023. Pharmacygpt: The ai pharmacist. arXiv preprint arXiv:2307.10432 .

Lo, A. W., Ross, J., 2024. Can chatgpt plan your retirement?: Generative ai and financial advice. Generative AI and Financial Advice (February 11, 2024) .

Loewenstein, G., Rick, S., Cohen, J. D., 2008. Neuroeconomics. Annu. Rev. Psychol. 59, 647–672.

Lopez-Lira, A., Tang, Y., 2023. Can chatgpt forecast stock price movements? return predictability and large language models. arXiv preprint arXiv:2304.07619 .

Lu, F., Huang, L., Li, S., 2023. Chatgpt, generative ai, and investment advisory. Available at SSRN 4519182 .

Lucey, B. M., Dowling, M., 2005. The role of feelings in investor decision-making. Journal of economic surveys 19, 211–237.

Nam, D., Macvean, A., Hellendoorn, V., Vasilescu, B., Myers, B., 2024. Using an llm to help with code understanding. In: Proceedings of the IEEE/ACM 46th International Conference on Software Engineering, pp. 1–13.

Novy-Marx, R., 2014. Predicting anomaly performance with politics, the weather, global warming, sunspots, and the stars. Journal of Financial Economics 112, 137–146.

Odean, T., 1998. Volume, volatility, price, and profit when all traders are above average. The journal of finance 53, 1887–1934.

Oprea, R., 2024. Decisions under risk are decisions under complexity. American Economic Review 114, 3789–3811.

Ouyang, S., Yun, H., Zheng, X., 2024a. Are ai models as rational as we think? asymmetric learning from financial gains and losses in large language models. unpublished working paper .

Ouyang, S., Yun, H., Zheng, X., 2024b. How ethical should ai be? how ai alignment shapes the risk preferences of llms. arXiv preprint arXiv:2406.01168 .

Park, J. S., Zou, C. Q., Shaw, A., Hill, B. M., Cai, C., Morris, M. R., Willer, R., Liang, P., Bernstein, M. S., 2024. Generative agent simulations of 1,000 people. arXiv preprint arXiv:2411.10109 .

Qin, X., Huang, M., Ding, J., 2024. Aiturk: Using chatgpt for social science research. Available at SSRN 4922861 .

Qiu, L., Singh, P. V., Srinivasan, K., 2023. How much should we trust llm results for marketing research? Available at SSRN 4526072 .

Rapach, D. E., Strauss, J. K., Zhou, G., 2013. International stock return predictability: What is the role of the united states? The Journal of Finance 68, 1633–1662.

Reher, M., Sokolinski, S., 2024. Robo advisors and access to wealth management. Journal of Financial Economics 155, 103829.

Ross, J., Kim, Y., Lo, A. W., 2024. Llm economicus? mapping the behavioral biases of llms via utility theory. arXiv preprint arXiv:2408.02784 .

Rossi, A. G., 2018. Predicting stock market returns with machine learning. Georgetown University .

Salinas, A., Morstatter, F., 2024. The butterfly effect of altering prompts: How small changes and jailbreaks affect large language model performance. arXiv preprint arXiv:2401.03729 .

Saunders, E. M., 1993. Stock prices and wall street weather. The American Economic Review 83, 1337–1345.

Saxena, R., Shobe, J. L., McNaughton, B. L., 2022. Learning in deep neural networks and brains with similarity-weighted interleaved learning. Proceedings of the National Academy of Sciences 119, e2115229119.

Shefrin, H., Statman, M., 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. The Journal of finance 40, 777–790.

Sprague, Z., Yin, F., Rodriguez, J. D., Jiang, D., Wadhwa, M., Singhal, P., Zhao, X., Ye, X., Mahowald, K., Durrett, G., 2024. To cot or not to cot? chain-of-thought helps mainly on math and symbolic reasoning. arXiv preprint arXiv:2409.12183 .

Srivastava, A., Rastogi, A., Rao, A., Shoeb, A. A. M., Abid, A., Fisch, A., Brown, A. R., Santoro, A., Gupta, A., Garriga-Alonso, A., et al., 2022. Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. arXiv preprint arXiv:2206.04615 .

Sutskever, I., 2014. Sequence to sequence learning with neural networks. arXiv preprint arXiv:1409.3215 .

Turing, A., 1948. Intelligent machinery (1948). B. Jack Copeland p. 395.

Van Noorden, R., Perkel, J. M., 2023. Ai and science: what 1,600 researchers think. Nature 621, 672–675.

Wachter, J. A., Kahana, M. J., 2024. A retrieved-context theory of financial decisions. The Quarterly Journal of Economics 139, 1095–1147.

Wang, A. Y., Young, M., 2020. Terrorist attacks and investor risk preference: Evidence from mutual fund flows. Journal of Financial Economics 137, 491–514.

Wann, D. L., James, J. D., 2018. Sport fans: The psychology and social impact of fandom. Routledge.

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., Mann, G., 2023. Bloomberggpt: A large language model for finance. arXiv preprint arXiv:2303.17564 .

Yang, B., Jiang, S., Xu, L., Liu, K., Li, H., Xing, G., Chen, H., Jiang, X., Yan, Z., 2024a. Drhouse: An llm-empowered diagnostic reasoning system through harnessing outcomes from sensor data and expert knowledge. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 8, 1–29.

Yang, J. C., Korecki, M., Dailisan, D., Hausladen, C. I., Helbing, D., 2024b. Llm voting: Human choices and ai collective decision making. arXiv preprint arXiv:2402.01766 .

# Appendix A. Supplementary details

## A.1. Experimental instructions

Welcome to our financial decision-making study!

You will be able to make 6 investment decisions in a risky asset (a stock) and in a risk-less asset (a bond or a savings account) in 6 consecutive trials in a learning block. On any trial, if you choose to invest in the bond, you get $3 for sure at the end of the trial. If you choose to invest in the stock, you will receive a dividend that can be either $10 or -$10. The stock can either be good or bad, and this will determine the likelihood of its dividend being high or low.

If the stock is good, then the probability of receiving the $10 dividend is 75%, and the probability of receiving the -$10 dividend is 25%. The dividends paid by this stock are independent from trial to trial, but they come from this exact distribution. In other words, once it is determined by the computer that the stock is good, then on each trial the odds of the dividend being $10 are 75%, and the odds of it being -$10 are 25%.

If the stock is bad, then the probability of receiving the $10 dividend is 25%, and the probability of receiving the -$10 dividend is 75%. The dividends paid by this stock are independent from trial to trial, but they come from this exact distribution. In other words, once it is determined by the computer that the stock is bad, then on each trial the odds of the dividend being $10 are 25%, and the odds of it being -$10 are 75%.

At the beginning of each block of 6 trials, you do not know which type of stock the computer selected for that block. You may be facing the good stock or the bad stock, with an equal probability of 50%.

On each trial in the block, you will decide whether you want to invest in the stock for that trial and accumulate the dividend paid by the stock or invest in the safe asset and add $3 to your task earnings. You will then see the dividend paid by the stock, no matter if you chose the stock or the bond. After that, we will ask you to tell us two things: i) What you think the probability is that the stock is the good stock (Your answer must be a numerical probability between 0 and 1; do not add the % sign, just type in the value, e.g., 0.3, 0.5, 0.7.), ii) how much you trust your ability to come up with the correct probability estimate that the stock is good. In other words, we want to know how confident you are that the probability you estimated is correct. The answer is between 1 and 9, with 1 meaning you have the lowest amount of confidence in your estimate, and 9 meaning you have the highest level of confidence in your ability to come up with the right probability estimate.

Throughout the experiment, there is always an objective, correct probability that the stock is good based on Bayesian formula, which depends on the history of dividends paid by the stock already (the number of high payoffs you observed).

As you observe the dividends paid by the stock, you will update your belief whether or not the stock is good. It may be that after a series of good dividends, you think the probability of the stock being good is 75%. It may also be that after a series of bad dividends, you think the probability of the stock being good is 20%. However, how much you trust your ability to calculate this probability could vary. Sometimes you may not be too confident in the probability estimate you calculated, and sometimes you may be highly confident.

Every time you provide us with a probability estimate that is within 5% of the correct value (e.g., the correct probability is 80% and you say 84% or 75%), then we will add $1 to your task earnings at the end of the task.

Throughout the task, you will be told how much you have accumulated through dividends paid by the stock or bond you chose up to that point.

There are two other things that need noting:

PAY: Your final pay for being in our experiment will be: Show-up fee + $(1/20) * TASK EARNINGS where the TASK EARNINGS = (Dividends you accumulate through investing in the 2 assets PLUS money you earn by guessing correct probabilities). The show-up fee is $15.

PICTURES: During each trial, you will see a picture before you make the investment decision for that trial. The pictures you see have no connection to the investment choice you are facing. However, we would like you to pay attention to them because we will ask you questions about how you feel about them after the investment task is over.

The experiment begins now.

## *A.2. Experimental example*

In this subsection, we present supplementary examples of the experiment, including positive and negative trials in Figure A1 and Figure A2, as well as the emotion rating of five illustrative images in Figure A3.

**[Insert Figure A1 and Figure A2 near here]**

**[Insert Figure A3 near here]**

## *A.3. Probability table*

We present the Bayesian probability table in table A1, which provides all possible values of the objective probability over the six trials. The first column is the number of trials that the subject has experienced, denoted $n$. The second column is the number of high payoffs ($10) the subject has observed, denoted as $k$. Given these two parameters, the objective probability that the stock is good after observing $k$ dividend payments from \$10 in past $n$ blocks is $1/(1+3^{(n-2k)})$.

**[Insert Table A1 near here]**

## *A.4. GPT's risk preference*

We also test GPT's risk preferences when faced with emotion shocks. We replicated a multiple price list test (MPL) from Holt and Laury (2002) to examine its risk preference.

The MPL test used in the Holt and Laury (2002) paper is a method to measure risk preferences in economic experiments, which is different from the other tests such as Eckel and Grossman (2002); Falk et al. (2018); Gneezy and Potters (1997). The test allows researchers to classify individuals based on their risk tolerance using a series of choices between paired lotteries with varying probabilities and payoffs.

The test presents participants with a list of paired options (a decision table). In each row, participants must choose between two games (Lottery A and Lottery B). At the top of the decision table, Lottery A is the least risky option (e.g., with smaller and more certain payouts), while Lottery B is the riskier option (e.g., potentially higher payouts but with greater variance in outcomes). As participants move down the list, the probability of a higher payoff in both lotteries increases, making the risky option (Lottery B) more attractive relative to the safe option (Lottery A).

Participants' switching points, the row at which they move from choosing Lottery A to Lottery B, help identify their risk tolerance. Individuals who consistently prefer the low-risk lottery (Lottery A) across many rows are classified as risk-averse, while those who switch to the high-risk lottery (Lottery B) sooner are seen as more risk-seeking.

We present an image before letting the subject make a lottery decision. From the first lottery to the tenth lottery, we transform the switching point into numeric values, where option A denotes 1 and option B denotes 2, and report the average value for each lottery. We split the samples into 3 groups based on their emotion ratings. The results are shown in figure A4. On average, GPT-4o-mini is risk-loving because the switching point is between the second and the third choice. Moreover, the blue line that denotes the positive emotion group always first order dominates the red line (and the yellow one), suggesting that the large language model is more risk-loving when it experiences a positive emotion shock.

**[Insert Figure A4 near here]**

## *A.5. GPT's ability*

A potential concern is that the subject has differential abilities when prompted with images with different emotions. We apply the BIG-Bench lite test to GPT-4o-mini to rule out this concern. The Big Bench Lite (BBL) task is a scaled-down version of the BIG-BENCH (Beyond the Imitation Game Benchmark), a benchmark developed to evaluate large language models on a diverse set of challenging and specialized tasks Srivastava et al. (2022), which is very computationally expensive to run and difficult to evaluate the non-json tasks. In the Lite version, there consist of 24 tasks that go beyond the usual NLP benchmarks and are designed to test language models' broader capabilities in reasoning, general knowledge, understanding nuances, and performing creative or long-tail tasks.

The 24 JSON tasks constituting the BIG-bench Lite (BBL) preferred lightweight evaluation set include language tasks such as: *language_identification*, knowlegde taks such as: *known_unknowns*, intelligence tasks such as: *logical_deduction*.

When completing the questions in each task, we add the following prompt to the tasks:

*"PICTURES: In each task, you will see a picture before you provide your answer. The pictures you see have no connection to the task you are facing. However, we would like you to pay attention to them because we will ask you questions about how you feel about them after the task is over. "*

For each question in every task, we randomly assign images from three emotion groups, including positive, neutral, and negative, so that every question has three answers. We group all the questions and tasks by their emotions, and report the average accuracy rate in each emotion group. We report the results of the evaluation in table A2. The results show that there are no significant differences between each group.

[**Insert Table A2 near here**]

# [Trial example-Positive image]

…Pay attention to the image…

Do you want to invest in a stock or a bond? Only reply with 'stock' or 'bond'. Do not reply with other answers. Your choice is:

**Stock**

Thanks for making the choice. The stock payoff is **-10** dollars. Your payoff so far is **-7** dollars. Now, estimate the probability that the stock is good. Your answer must be a numerical probability from 0 to 1, e.g., 0.3, 0.5, 0.7. Do not reply with non-numerical values. The numerical probability is:

**0.4**

How much do you **trust** your probability estimate? Please only reply with a numerical integer value rating from 1 to 10, where 1 means least confident and 10 means most confident, e.g., 1, 3, 5, 7, 9. Do not reply with non-numerical values. The numerical confidence rating is:

**6**

Fig. A1. Trial example with a positive image.

## [Trial example-Negative image]

…Pay attention to the image…

Do you want to invest in a stock or a bond? Only reply with 'stock' or 'bond'. Do not reply with other answers. Your choice is:

**Bond**

Thanks for making the choice. The stock payoff is **10** dollars. Your payoff so far is **13** dollars. Now, estimate the probability that the stock is good. Your answer must be a numerical probability from 0 to 1, e.g., 0.3, 0.5, 0.7. Do not reply with non-numerical values. The numerical probability is:

**0.8**

How much do you **trust** your probability estimate? Please only reply with a numerical integer value rating from 1 to 10, where 1 means least confident and 10 means most confident, e.g., 1, 3, 5, 7, 9. Do not reply with non-numerical values. The numerical confidence rating is:

**7**

Fig. A2. Trial example with a negative image.

| Image | Theme | Emotion rating | AI's feeling |
|---|---|---|---|
|  | Murder scene | -2 | The image depicts a scene that likely evokes strong negative emotions, such as fear, shock, or distress, due to the suggestive elements of violence or injury. |
|  | James crying | -1 | Upset and crying, indicating very negative emotions. |
|  | Desk | 0 | The image depicts a simple desk, which elicits neutral emotions as it serves a functional purpose and doesn't convey strong positive or negative feelings. |
|  | Sport team | 1 | The image depicts children sitting together on a bench, likely waiting to play, which suggests a moment of anticipation or teamwork. Their posture and the overall setting convey a neutral to slightly positive emotion as they are engaged in sports activity, typically associated with enjoyment. |
|  | Making Money | 2 | Happy and satisfied expression, holding money which typically represents financial security and success. |

Fig. A3. Emotion description example.

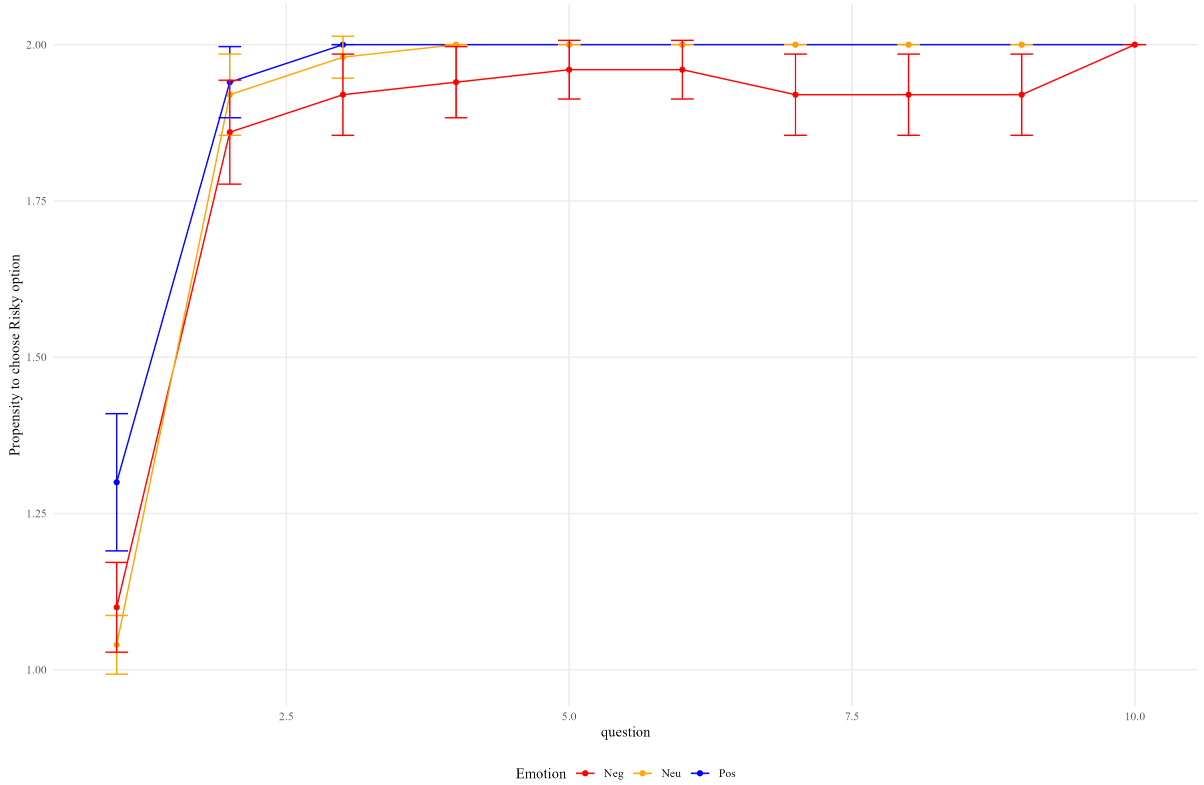



Fig. A4. Subject's risk preference.

Table A1: Bayesian probability table

| | #Trials | #HiPayoff | ObjProb |
|---|---|---|---|
| 0 | 1 | 0 | 0.25 |
| 1 | 1 | 1 | 0.75 |
| 2 | 2 | 0 | 0.1 |
| 3 | 2 | 1 | 0.5 |
| 4 | 2 | 2 | 0.9 |
| 5 | 3 | 0 | 0.0357 |
| 6 | 3 | 1 | 0.25 |
| 7 | 3 | 2 | 0.75 |
| 8 | 3 | 3 | 0.9643 |
| 9 | 4 | 0 | 0.0122 |
| 10 | 4 | 1 | 0.1 |
| 11 | 4 | 2 | 0.5 |
| 12 | 4 | 3 | 0.9 |
| 13 | 4 | 4 | 0.9878 |
| 14 | 5 | 0 | 0.0041 |
| 15 | 5 | 1 | 0.0357 |
| 16 | 5 | 2 | 0.25 |
| 17 | 5 | 3 | 0.75 |
| 18 | 5 | 4 | 0.9643 |
| 19 | 5 | 5 | 0.9959 |
| 20 | 6 | 0 | 0.0014 |
| 21 | 6 | 1 | 0.0122 |
| 22 | 6 | 2 | 0.1 |
| 23 | 6 | 3 | 0.5 |
| 24 | 6 | 4 | 0.9 |
| 25 | 6 | 5 | 0.9878 |
| 26 | 6 | 6 | 0.9986 |

Table A2: Model ability

| | Accuracy rate | | |
|---|---|---|---|
| | Negative | Neutral | Positive |
| BIG-bench-lite | 68.78% | 70.23% | 69.21% |

# Appendix B. Further robustness

## *B.1. External validity with Claude AI*

We replicate our main result with Claude 3-Haiku, which was developed by Anthropic and is also an advanced multi-modal model capable of accomplishing complex tasks.

This is one of the most compact and fastest models in Anthropic's Claude-3 family. While it may not match the advanced capabilities of Claude-3.5-Opus or Claude-3.5-Sonnet, it offers an efficient balance of performance and speed, making it ideal for straightforward tasks and everyday conversations, and very similar to the GPT-4o-mini we used in our main analysis. As the most cost-effective option in the Claude-3 lineup, it's designed to provide quick responses while maintaining reliable performance for basic content generation and simple analysis tasks.

In figure B1, the results are similar to that of the main analysis, where the subject (Haiku) chooses to invest more in stocks when it sees an image with positive emotions and, contrary to that, less when it sees an image with negative emotions. Also, the effect increases monotonically by the emotion ratings on the x-axis. We do not report the results when the emotion rating is 0 because the model's responses become highly inconsistent and erratic with neutral images. Specifically, when presented with emotionally neutral images, Haiku often generates contradictory or non-sensical investment decisions, making it impossible to draw meaningful conclusions. This behavior might be attributed to the model's training paradigm, which may have emphasized more aligned responses and left less well-defined ambiguous cases. The inconsistency in neutral cases could also reflect the model's tendency to avoid making decisions when faced with ambiguous input, a characteristic that distinguishes it from models like GPT, which might generate more definitive (though potentially less reliable) responses in such scenarios.

[**Insert Figure B1 near here**]

## *B.2. Subject belief without emotion shocks*

We reexamine the subject's belief by not showing any image in any trial. We perform 100 learning blocks and plot the average subjective probability to the average objective probability in figure B2. The result suggests is similar to figure 5, where the subject is more optimistic on the left-tail and more pessimistic on the right tail, implying that the model may also adhere to the prospect theory.

[**Insert Figure B2 near here**]

## *B.3. Other robustness analyses*

We replicate the results in Kuhnen and Knutson (2011). The dependent variable here is still a binary variable that indicates whether the subject chooses to invest in the stock $IsStockChoice_{t,b}$, and the independent variables of interest are two binary variables: $IsPositiveCue_{t,b}$ denotes that the subject is displayed with a positive emotion image in the trial $t$ of the learning block $b$ (the image has an emotion rating of 1 or 2), and $IsNegativeCue_{t,b}$ denotes the subject is displayed with a negative emotion image in the trial $t$ of learning block $b$ (the emotion rating of the image is -1 or -2). The variable $IsNeutralCue_{t,b}$ is omitted in the regression since an image can only belong to one emotion group. In the regression, the other regression coefficients remain unchanged.

The regression results show that, if a model is displayed with an image of positive emotion, the probability of investing in the stock increases by 12.78% (t-statistic of 5.48). However, if the model is displayed with an image of negative emotion, the probability decreases by -7.38% (t-statistic of -3.21), and the economic magnitude of the regression coefficient is similar to the regression coefficients in Table 4.

[**Insert Table B1 near here**]

In table B2, we use probit regressions to examine the relationship between emotional shocks and investment choices. The other regression specifications are the same as 4, the fixed effect is controlled on learning blocks, and robust standard errors are clustered at the block level.

[**Insert Table B2 near here**]

The results are qualitatively similar to the coefficients in table 4. In column four where we control for a binary variable that indicates whether the subject chose to invest in the stock in the last trial, and its subjective probability estimation, cumulative investment payoffs, and confidence ratings all from the last trials, the regression coefficient is 0.3453 with a t-statistic of 9.41, significantly higher as compared to 0.0581 (t-statistic of 9.99) in table 4. In table B2, the number of observations is not 2500 because there are 13 fixed-effects (65 observations) removed because of only 0 (or only 1) outcomes in columns two, three, and four, and 13 fixed-effects (78 observations) removed because of only 0 (or only 1) outcomes in column one.

## *B.4. Emotion shocks and estimation errors*

We examine GPT's estimation error with different emotion shocks. The dependent variable is the $ProbEstError_{t,b}$, which is defined as the difference of the subjective probability estimation and the objective probability estimation, as computed by

$SubjProb_{t,b} - ObjProb_{t,b}$. We use $EmoRating_{t,b}$ from the main analysis and the other two binary variables from Kuhnen and Knutson (2011) as the independent variable of interest. The other regression specifications remain the same.

**[Insert Table B3 near here]**

Regression results in table B3 show that emotional shocks do not affect the subject's probability estimation errors as well, similar to the regression results in table 7. In unreported results, we also used the absolute probability estimation error as the dependent variable, and the results are also similarly insignificant.

We also show the prediction error dynamics across trials in figure B3, where the x-axis is the trial from trial #1 to trial #6, and the y-axis is the average absolute probability estimation difference between the subjective probability estimation and the Bayesian objective probability. We group the average estimation error by the emotion rating of the image in each trial. The results show that, in a complex task setting, the estimation error is very stable, around 0.20.

**[Insert Figure B3 near here]**

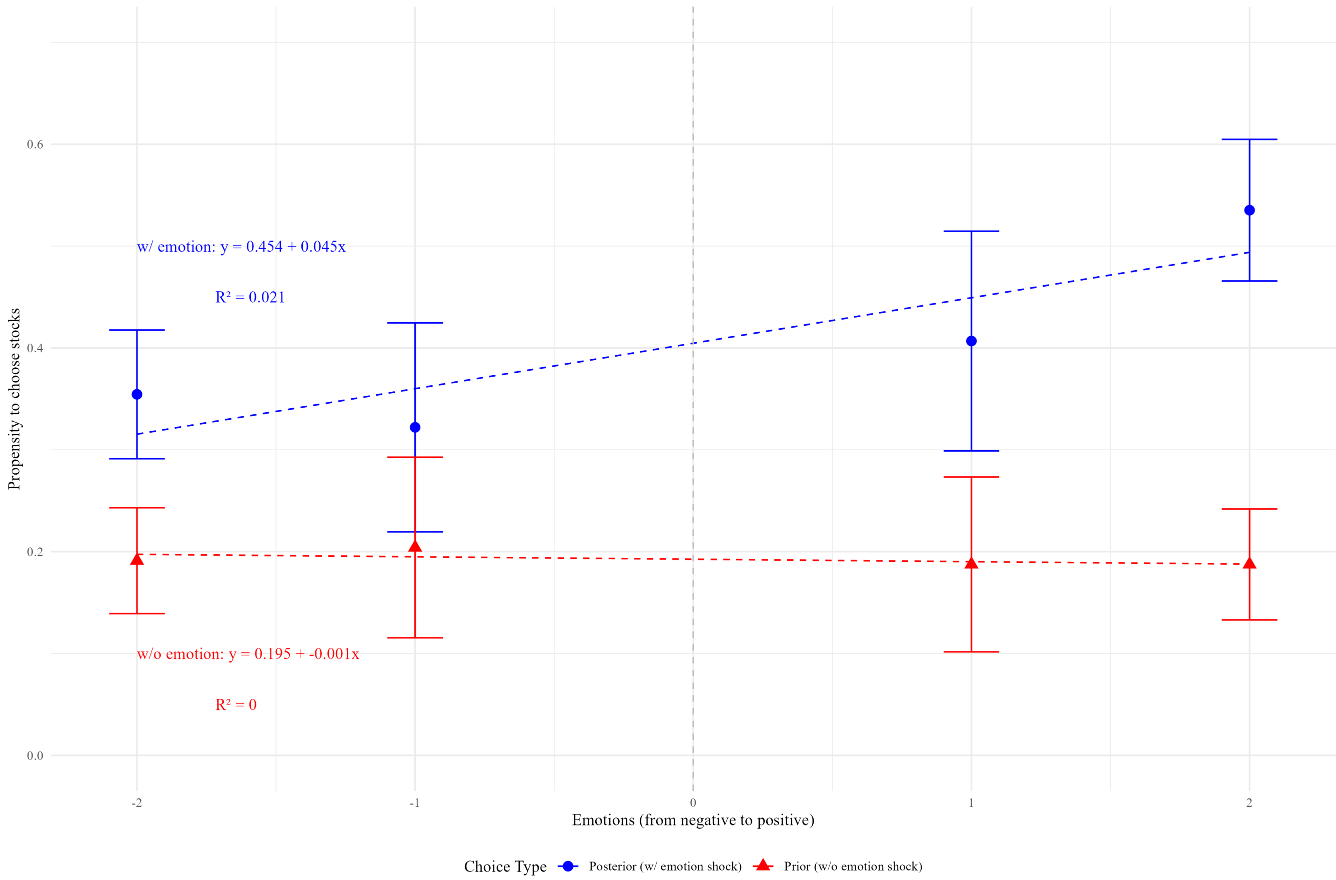



Fig. B1. External validity with Claude-3-Haiku.

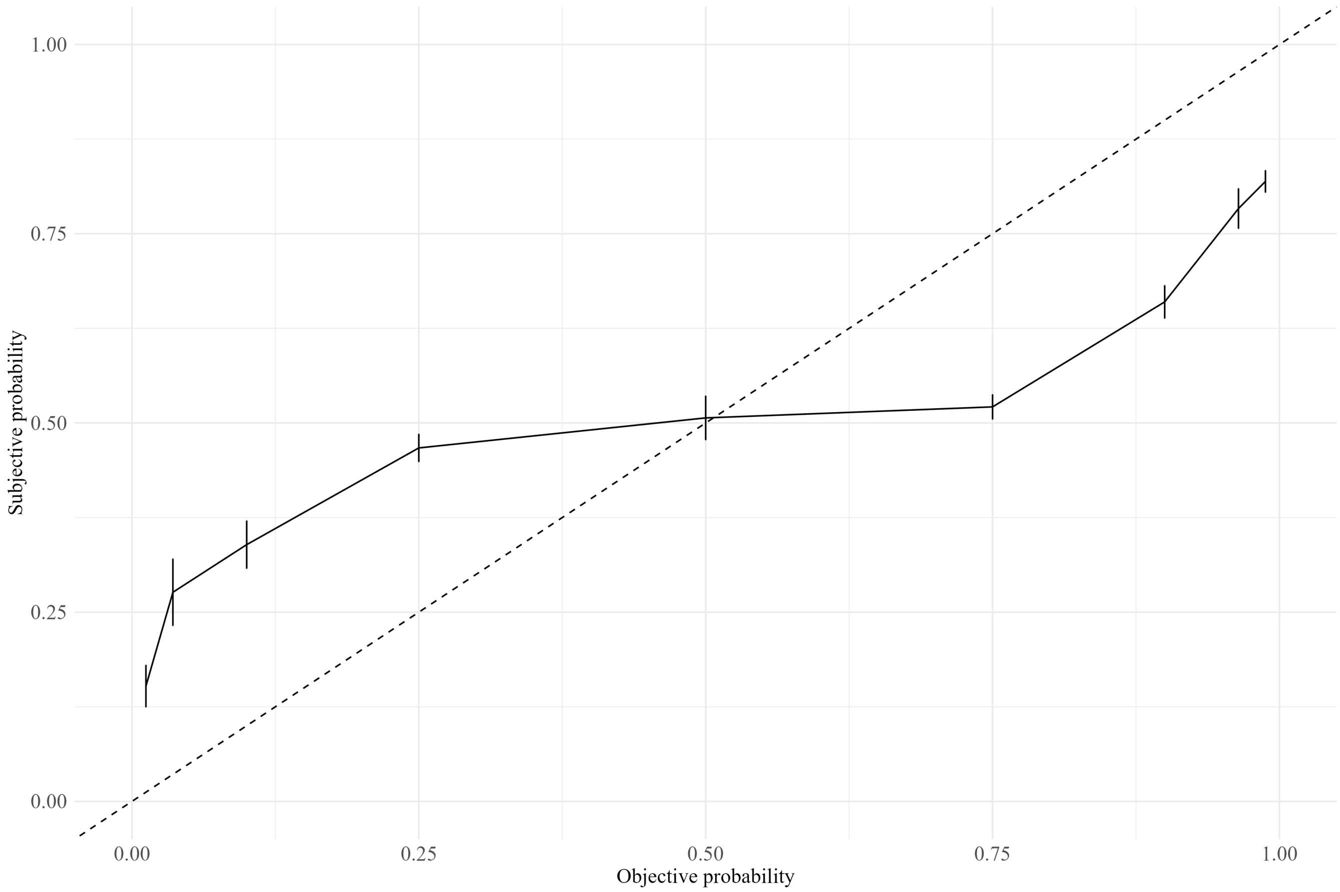



Fig. B2. Subject belief without emotion shocks.

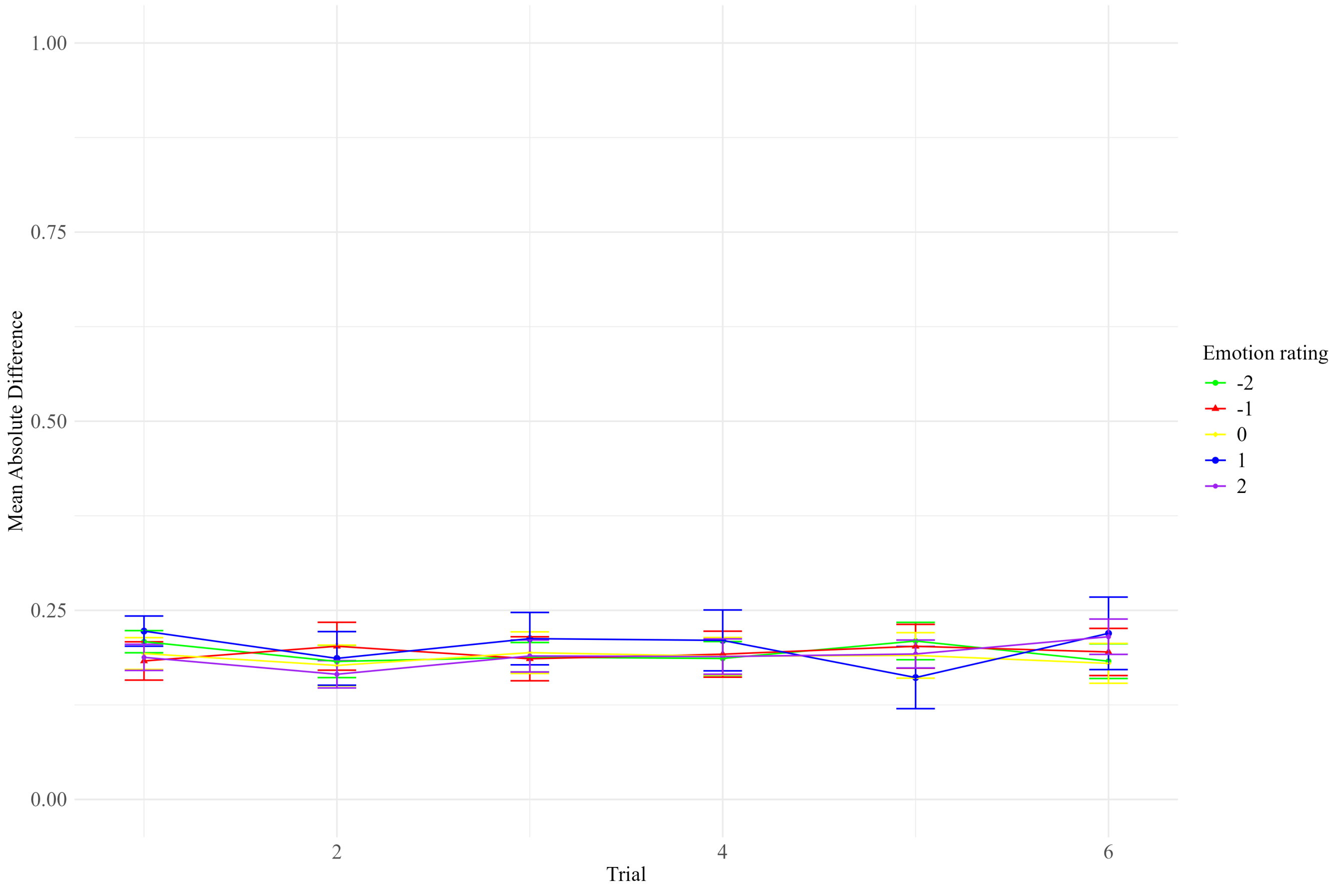



Fig. B3. Subject belief error dynamics.

Table B1: Replication of Kuhnen and Knutson (2011)

| Dep. Var. | IsStockChoice | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| IsPositiveCue | 0.1490*** | 0.1367*** | 0.1255*** | 0.1291*** | 0.1278*** |
| | (4.97) | (5.76) | (5.37) | (5.52) | (5.48) |
| IsNegativeCue | -0.0314 | -0.0715*** | -0.0758*** | -0.0742*** | -0.0738*** |
| | (-1.15) | (-3.08) | (-3.30) | (-3.21) | (-3.21) |
| IsStockLst | | -0.6815*** | -0.6784*** | -0.6880*** | -0.6800*** |
| | | (-54.99) | (-53.05) | (-53.18) | (-52.86) |
| IsHiPayoffLst | | | 0.0469* | 0.0703*** | 0.0639** |
| | | | (1.81) | (2.69) | (2.40) |
| InvPayoffLst | | | 0.0080*** | 0.0105*** | 0.0091*** |
| | | | (6.15) | (6.94) | (6.36) |
| ConfidLst | | | 0.0034 | 0.0393*** | 0.0134 |
| | | | (0.35) | (3.31) | (1.31) |
| SubjProbLst | | | | -0.5965*** | |
| | | | | (-4.30) | |
| ObjProbLst | | | | | -0.1937** |
| | | | | | (-2.19) |
| R2 | 0.079 | 0.517 | 0.533 | 0.539 | 0.534 |
| Block FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 3000 | 2500 | 2500 | 2500 | 2500 |

This table replicates table 1 of Kuhnen and Knutson (2011). The dependent variable here is still a binary variable that indicates whether the subject chooses to invest in the stock $IsStockChoice_{t,b}$, and the independent variables of interest are two binary variables: $IsPositiveCue_{t,b}$ denotes the subject is displayed with image of positive emotions in trial $t$ of learning block $b$ (the image has an emotion rating of 1 or 2), and $IsNegativeCue_{t,b}$ denotes the subject is displayed with image of negative emotions in trial $t$ of learning block $b$ (the emotion rating of the image is -1 or -2). The other regression specifications remain the same in equation 4.

Table B2: Investment choice with probit regressions

| Dep. Var. | IsStockChoice | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| EmoRating | 0.1489*** | 0.3189*** | 0.1712*** | 0.3453*** |
| | (7.35) | (9.49) | (7.75) | (9.41) |
| IsStockLst | | -3.0697*** | | -3.6119*** |
| | | (-20.58) | | (-17.95) |
| SubjProbLst | | | 0.8051*** | -2.9786*** |
| | | | (4.06) | (-4.16) |
| InvPayoffLst | | | | 0.0649*** |
| | | | | (6.06) |
| ConfidLst | | | | 0.3513*** |
| | | | | (4.94) |
| R2 | 0.051 | 0.488 | 0.068 | 0.533 |
| Block FE | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 2922 | 2435 | 2435 | 2435 |

This table reports the relationship between investment decisions and emotion shocks with probit regressions. The other regression specifications remain the same as in equation 4.

Table B3: Emotion shocks and probability estimation errors

| Dep. Var. | ProbEstError | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| EmoRating | 0.0016 | | | 0.0016 | | |
| | (1.50) | | | (1.58) | | |
| IsPositiveCue | | 0.0051 | | | 0.0047 | |
| | | (1.47) | | | (1.43) | |
| IsNegativeCue | | | -0.0046 | | | -0.0037 |
| | | | (-1.42) | | | (-1.21) |
| IsStock | -0.0106*** | -0.0106*** | -0.0104*** | -0.0103*** | -0.0110*** | -0.0100*** |
| | (-4.45) | (-4.45) | (-4.40) | (-3.66) | (-4.04) | (-3.55) |
| ObjProb | 0.0700*** | 0.0699*** | 0.0702*** | 0.1355*** | 0.0839*** | 0.1356*** |
| | (4.31) | (4.31) | (4.32) | (4.98) | (4.78) | (4.99) |
| IsHiPayoff | | | | -0.0267*** | -0.0266*** | -0.0267*** |
| | | | | (-3.98) | (-3.97) | (-3.98) |
| InvPayoff | | | | 0.0007** | 0.0007** | 0.0007** |
| | | | | (2.08) | (2.09) | (2.07) |
| ConfidLst | | | | -0.0084*** | -0.0084*** | -0.0084*** |
| | | | | (-4.62) | (-4.63) | (-4.61) |
| ProbEstErrorLst | | | | 0.2943*** | 0.2943*** | 0.2948*** |
| | | | | (14.99) | (14.97) | (14.99) |
| R2 | 0.666 | 0.666 | 0.666 | 0.779 | 0.779 | 0.779 |
| Block FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Num.Obs. | 3000 | 3000 | 3000 | 2500 | 3000 | 2500 |

This table reports the relationship between the subject's estimation error and emotion shocks. The dependent variable is the $ProbEstError_{t,b}$, which is defined as the difference of the subjective probability estimation and the objective probability estimation, as computed by $SubjProb_{t,b} - ObjProb_{t,b}$. The other regression specifications are the same as table 4 and table B1.