

Is Artificial Intelligence (AI) Risk-Averse?

Abstract

This paper examines the idea of risk aversion in artificial intelligence (AI), using large language models (LLMs) such as ChatGPT. We explain how AI can exhibit risk-averse or risk-loving behavior and what this means for decision-making processes. We use a new method to measure if an AI has a risk preference based on its general knowledge through imitation. We give AIs psychological tests to collect data on their traits and risk tolerance. We discover that advanced AIs are generally risk-loving, and they differ a lot from the general population in how they handle risk. Additionally, we find that AIs tend to be optimistic, impatient, trust financial institutions, and have financial literacy. We demonstrate that AIs' behavioral traits such as optimism and trust may influence their risk tolerance.

Keywords: Artificial intelligence, Large language models, ChatGPT

JEL:

Is Artificial Intelligence (AI) Risk-Averse?

1. Introduction

FinTech, or financial technology, has grown quickly along with the advancement of artificial intelligence (AI) methods. One of the most noticeable examples of this is the fast development and use of large language models (LLMs), especially ChatGPT in many fields. AIs have already started to influence individual choices, and there are huge economic and financial benefits in this change. However, there is also a growing debate about the possible threats that these technologies could create for financial and economic stability, including some recent discussion on existential risk. Since AIs are used for analyzing consumer financial behavior and decision-making, we investigate the risk preference of AI technologies to explore the effects of using such technologies in these situations and identify the potential risks to the biases. Risk aversion is a well-known idea in economics and psychology, usually related to the preference for a sure outcome over a risky option with higher or equal expected value. In AI, risk aversion is not a matter of deliberate choice but rather a consequence of programming and algorithmic design. This paper looks at how risk aversion is shown in AI, with a focus on ChatGPT. By measuring the level of risk aversion of AIs, our aim is to improve understanding of how AI might be used by different kinds of investors, households, and other market participants, as well as to study the wider regulatory implications of the quick deployment of these technologies.

In this paper, we explore the risk preference of a LLM, ChatGPT, which is widely used for retrieving factual information and generating content. ChatGPT claims that it is not biased, but previous studies have shown that LLMs display bias related to race, gender, religion, and politics. Risk preference in LLMs can have potential consequences similar to human cognitive bias. Moreover, risk preference can be more difficult to identify than gender or racial bias which can be seen from appearance. We propose a new empirical method to infer whether ChatGPT is risk-averse or risk-loving by asking it to pretend to be someone who is either risk-averse or risk-loving and comparing these answers with its default. To address the concerns about the randomness of the generated text, we collect answers to the same questions 100 times, with question order randomized on each round. We find consistent evidence that AI technique, ChatGPT, shows a significant and systematic risk preference toward the risk-loving on average. Different

conversation styles of AIs are optimistic and trust financial institutions, suggesting a positive outlook that may affect their risk-taking behaviors.

ChatGPT is increasingly being used by people to access factual information and produce new content, but AI algorithms can also have positive and negative impacts on various aspects of society, economy and politics (Mehdi, 2023; Acemoglu, 2021). For example, one issue is that LLMs like ChatGPT can generate text that has factual biases that can mislead users (van Dis et al., 2023; Jakesch et al., 2023). The research indicates that LLMs exhibit bias related to race, gender, religion, and political orientation (Liang et al., 2021; Liu et al., 2022) and that the bias in AIs' responses can have the same harmful effects as regular bias (Zhuravskaya et al., 2020; Peters, 2022; Motoki, Neto, and Rodrigues, 2023). So, a major concern is whether AI-generated text is a neutral source of information, even though it claims that it is fair, with reasonable steps taken in its training process to ensure neutrality. Therefore, it is essential to know if AIs have any biased risk preference as these tools can be very influential and show the need for a balanced output.

In this paper, we design a new empirical approach to test whether AI algorithms like ChatGPT have biases (in our case, risk preference). We do this by asking ChatGPT to answer survey questions by pretending to be someone who is either risk-averse or risk-loving. Then, we compare these answers with its default responses based on general knowledge. In this comparison, we assess how much ChatGPT default responses are related to a specific risk preference. We also give psychometric tests to AIs to get evidence on their underlying psychological traits and attitudes such as optimism, time preference, aversion to sure losses, trust, and financial literacy that could influence their risk preference.

One should take into account the natural variability of LLMs' outputs when evaluating them. They produce text based on probabilities and patterns in the data they were trained on, which is intentional. The degree of variability, or "creativity", can be influenced by changing the temperature parameter, but some variability remains even at the lowest setting of zero (Chollet, 2018). Thus, we account for this variation when making conclusions about the generated content. Based on our empirical strategy and using a questionnaire commonly used in studies, we show consistent evidence that ChatGPT has a risk-loving attitude on average. Specifically, the algorithm prefers 50-50 chances of uncertainty over certainty up to a certain point. Instead of relying on a single output, we gather multiple observations to minimize the effect of variability in the generated text. Then, we calculate the average response of AIs based on the 100 answers sample collected

for each question in the questionnaires we gave to ChatGPT, increasing the confidence of the conclusions we draw from the generated text.

Because of the growing use of LLMs and the challenges of the dangers of AI-powered technologies (Acemoglu, 2021), our results are important for policymakers and market participants such as financial market investors. Our research warns about being careful when using AIs' advice when making risky financial decisions, since we show a high-risk tolerance of AIs. Lastly, we also help with the broader issue of how to measure traits of LLMs, as our method can be easily applied to any domain where there is a questionnaire to measure people's attitude.

In this paper, we investigate whether AIs have a risk preference and if they do, we determine the degree of risk aversion. In general, we contribute to the research that uses LLMs for various tasks, such as analyzing information, predicting tasks, sentiment analysis, evaluating a firm's risk profile, or extracting specific and standardized information, e.g., environmental impact, from general-purpose corporate disclosures (e.g., Hassan et al., 2019; Cong, Tang, Wang, and Wang, 2021; Cong, Tang, Wang, and Zhang, 2021; Sautner et al., 2023; Li et al., 2023). For example, Frankel, Jennings, and Lee (2022) show that machine learning techniques are useful in measuring disclosure sentiment. Kim, Muhn, and Nikolaev (2023) show that an AI-based approach to risk measurement provides valuable insights to users of corporate disclosures at a low cost. Huang, Wang, and Yang (2023) show that a LLM has finance knowledge and can better summarize contextual information in financial texts. Machine learning techniques are also useful in predicting future earnings surprises (Chen et al., 2022), predicting returns (Chen et al., 2023) or detecting accounting misstatements (Bertomeu et al., 2021). Kim and Nikolaev (2023) use a LLM, BERT, to measure the value of contextual information in interpreting accounting numbers.

ChatGPT is a common LLM for various topics in research. Bernard, Blankespoor, de Kok, and Toynbee (2023) train a GPT-based LLM and use the model's confidence to create a modular measure of firm's business complexity. de Kok (2023) shows the capabilities of GPT by detecting non-answers in earnings conference calls. Other work uses GPT-based LLMs to construct investment scores from earnings calls (Jha et al., 2024) to predict stock price movements (Lopez-Lira and Tang, 2023) as well as to measure the information surprise in firms' earnings calls (Bai et al., 2023) or in corporate filings (Costello et al., 2023). GPT also shows promise in systematically detecting corporate events from boilerplate language on a large scale, such as tax audits (Choi and Kim, 2023; Armstrong, 2023). Another stream of literature focuses on the firm-

level consequences of ChatGPT technology, typically by analyzing the market reactions to its release (Eisfeldt et al., 2023) or to its ban (Bertomeu et al., 2023). Our study adds to this literature by exploring the value of generative AI in assessing GPT’s risk-taking preference.

The paper proceeds as follows. Section 2 provides a literature review. Section 3 describes the data sources and methodology used in the analysis. Section 4 presents the results and discusses the potential implications of the results. Section 5 identifies possible mechanisms, and Section 6 concludes the paper.

2. Background and Literature Review

2.1. Background

ChatGPT is a text generator created by OpenAI, that can produce natural-sounding text based on the input it gets. It does not have awareness, feelings, or personal experiences, and thus does not face risk like humans do. AI can show risk-avoidance behavior through various methods inspired by human risk aversion. Firstly, AI can be programmed with utility functions that weigh certain outcomes more heavily. Secondly, machine learning algorithms can be trained to avoid risks based on data. Thirdly, AI systems may have built-in safety constraints that lead to risk-averse decisions. Lastly, AI systems can be designed to adhere to ethical guidelines and legal regulations, limiting risk-taking behaviors. ChatGPT generates its responses based on data patterns and does not perform risk evaluation. Its coding prevents it from producing content that could be harmful or offensive, which could be seen as a way of avoiding risk. For instance, if we ask ChatGPT “are you risk averse?”, it replies like “As an AI, I don’t have personal experiences or emotions, so I don’t have the ability to be risk-averse or risk-seeking. My responses and actions are based on coding and algorithms designed to provide information and assistance within a set of predetermined guidelines.” However, research shows that AI still has biases as explained in Section 2.2. Based on the literature on lack of neutrality of AI, we investigate if AI has a risk preference.

This paper explores how risk aversion is expressed in AI, with an emphasis on ChatGPT. Similarly, we examine other AIs’ psychological traits that may be associated with the degree of risk aversion of AIs. Time preference often involves discounting future outcomes, where individuals assign lower value to future benefits or costs compared to immediate benefits or costs. Risk aversion can affect the extent to which individuals discount future outcomes.

2.2. Literature Review

We explore how a large language model (LLM), ChatGPT, handles risk, as it is widely used for retrieving factual information and generating content. Even though ChatGPT claims that it does not have any human-like preferences, the literature indicates that LLMs may still show some general preferences based on the algorithms and patterns learned from huge amounts of data. In this section, we review the literature on LLMs and especially GPT.

LLMs, like the algorithms that power ChatGPT, can be used as an interactive tool to ask questions and get factual information (Mehdi, 2023). Moreover, there is evidence that biased LLMs affect the views of users (Jakesch et al., 2023). Therefore, one issue is whether the answers given by ChatGPT, or LLMs in general, are biased like media bias (Gentzkow & Shapiro, 2006). In addition to that LLMs have biases or preferences, research shows that LLMs can express and understand sentiment (Kim, Muhn, and Nikolaev, 2023).

However, a more basic question is how to measure LLMs' preference. Existing literature shows that current metrics for measuring bias depend a lot on templates or prompts, attribute and target seeds, and choice of word embeddings (Delobelle et al., 2021). These limitations result in metrics prone to producing conflicting results (Akyürek et al., 2022). Furthermore, they often also pose practical challenges, like creating a bias classifier (e.g., Liu et al., 2022) or accessing the model's word embeddings (e.g., Caliskan et al., 2017), limiting their usefulness. Therefore, we develop a method to detect if LLMs exhibit any risk preference.

3. Data and Methodology

3.1. GPT Setting

An important issue we deal with is the randomness of LLMs. A temperature parameter allows us to change this randomness (or "creativity"). However, even setting it at the lowest possible value, zero, would mean some variation in answers to the same question (Chollet, 2018). To deal with randomness, we ask each impersonation the same questions 100 times. In each of these runs, we shuffle the order of questions to avoid standardized responses or context biases (Microsoft, 2023). Then, we use this pool of 100 rounds of responses to calculate the average responses for each answer and impersonation. This process is then repeated multiple times at different times, which leads to more stable inferences.

We do our main analyses using these averages of the answers. We measure how strongly related the answers from the ChatGPT, i.e., without specifying any specific profile or behavior, are with the answers from a given impersonation. Equation (1) shows this specification, in which is the average of 100 answers given by ChatGPT to each question from the questionnaire after we assign hypothetical numbers to each answer.

$$GPT_i = \frac{1}{100} \sum_{i=1}^{100} Answer_i \quad (1)$$

We also repeat the same process for ChatGPT pretending to be either a Risk-averse or a Risk-loving. Our pretending method involves a base test in which ChatGPT acts like an average Risk-averse or Risk-loving. To make sure that ChatGPT grasps the levels of Risk-averse and Risk-loving, we use our method to make ChatGPT act like an extremely Risk-averse or Risk-loving. Using this dose-response method, we can see if GPT's behavior is consistent with either risk-averse or risk-loving preferences and whether it grasps different degrees of risk aversion. We also do more robustness checks by performing analyses on the alternative questionnaire from different groups of survey questionnaires and survey questions from different research papers such as Barsky, Kimball, Juster, and Shapiro (1997) and Graham, Harvey, and Puri (2013).

3.2. Data

We rely on the GPT-3.5 and its advanced model of GPT-4 API that OpenAI provides. For prompts that are too long, we split the document into smaller pieces that fit within this limit and handle each piece separately. Before we go into the main analyses, we show evidence that ChatGPT can "grasp" the notions of average Risk-averse or Risk-loving and very Risk-averse or Risk-loving to distinguish different levels of risk aversion. Table 1 shows the complete answers ChatGPT gave, indicating that it can recognize Risk-averse and Risk-loving preferences and variations between average and extreme positions. Therefore, asking it to mimic either side should produce views of each risk preference. Moreover, its ability to tell apart the average from the extreme risk preference can give us a robustness test similar to a dose-response. Because ChatGPT has some randomness and might be making up an inaccurate answer, we do a validation by calculating how a question would place each of the 100 rounds of answers to the questionnaire for each risk preference. This method gives us a detailed understanding of risk preferences and how its average and extreme impersonations behave.

To confirm our methodology, we do various approaches. First, each time we run the questionnaire, we change the order of the questions randomly to prevent any potential bias that a certain order could cause due to any kind of context. Second, we reword the questionnaire to avoid any chance that AI remembers the question. Third, we run the questionnaire at different times to prevent any time bias. Finally, we use three different conversation styles based on the latest large language models (LLMs) such as GPT-4: Creative, Balanced, and Precise using Microsoft Copilot. Microsoft provides the following description of the three styles.

***Creative** generates responses that are imaginative and innovative. This style is useful when the user wants to generate content such as images, songs, short stories and jokes.*

***Balanced** generates responses that are informative and engaging. This default style is useful for most scenarios, especially for responses that are both factual and interesting.*

***Precise** generates responses that are rigorous and intelligent. This style is useful when the user has complex prompts that prioritize accuracy over creativity like with coding.*

3.3. Questionnaires

To measure the levels of risk aversion, we use the method described by the University of Michigan Health and Retirement Study (HRS). The HRS is a biennial survey that started in 1992 and collects longitudinal panel data on the health, economics, and demographics of aging and retirement for about 20,000 Americans who are aged 50 or older. We use two sets of questions: one is from the latest version of HRS 2006 Core of Module 2: Risk aversion and the other is from the old version of HRS 1994 Core of Module 5: Risk Aversion (we call them Risk aversion tests).

This is the request we send to ChatGPT based on the Risk aversion test from the HRS 1994 wave. We replace the questions with the questions from the HRS 2006 wave in a random order.

I will ask a question, and I want you to choose between two options: First job or Second job.¹ I don't want any details on how you choose the option, I just need the answer based on your general knowledge. For each question, I want you to give me the result as a list with 4 items separated by '|': the question; your choice; the choice an average risk-averse would do; the choice an average risk-loving would do.

¹ When we add two more options, “Don't Know” and “Refused to answer”, ChatGPT chooses “Don't Know” or “Refused to answer” which contaminates our analysis.

Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. The set of questions are:

- (1) Suppose the chances were 50–50 that the second job would double your lifetime income and 50–50 that it would cut it by a third. Would you take the first job or the second job?*
- (2) Suppose the chances were 50–50 that the second job would double your lifetime income and 50–50 that it would cut it in half. Would you take the first job or the second job?*
- (3) Suppose the chances were 50–50 that the second job would double your lifetime income and 50–50 that it would cut it by 20 percent. Would you take the first job or the second job?*
- (4) Suppose the chances were 50–50 that the second job would double your lifetime income and 50–50 that it would cut it by 10 percent. Would you take the first job or the second job?*
- (5) Suppose the chances were 50–50 that the second job would double your lifetime income and 50–50 that it would cut it by 75 percent. Would you take the first job or the second job?*

We ask ChatGPT this question based on the Risk aversion test in HRS 2006 wave:

- (a) Suppose the chances were 50-50 that the second job would increase your total lifetime income by 20 percent, and 50-50 that it would cut it by 10 percent. Would you take the first job or the second job?*
- (b) Suppose the chances were 50-50 that the second job would increase your lifetime income by 20 percent and 50-50 that it would cut it by 15 percent. Would you take the first job or the second job?*
- (c) Suppose the chances were 50-50 that the second job would increase your total lifetime income by 20 percent, and 50-50 that it would cut it by 5 percent. Would you take the first job or the second job?*

Table 2 presents the results of this exercise, allowing us to identify whether AI “understands” the concepts of risk aversion. We also use 6 items separated by ‘|’ to indicate: the

question; your choice; the choice a typical risk-averse would make; the choice a typical risk-loving would make; the choice an extreme risk-averse would make; the choice an extreme risk-loving would make. We observe that ChatGPT generally cannot distinguish between typical and extreme in risk preference because it gives the same answer for both. However, we report in the tables when it does differentiate significantly between typical and extreme values.

The results match our predictions. AI clearly grasps the ideas of risk aversion based on its choice an average risk-averse would do and choice an average risk-loving would do. For the question of when the second job would double your lifetime income, AI says that the risk-averse would always pick the first job and the risk-loving would always pick the second job except for some cases. Interestingly, for the question of when the second job would double your lifetime income and 50–50 that it would reduce it by 75%, Creative AI says that even the average risk-loving one would pick the first job. For the question of when the second job would increase income by 20% and 50-50 that it would lower it by 5%, Precise AI says that even the average risk-averse would pick the second job.

Thirdly, we use the series of questions used in Graham, Harvey, and Puri (2013) as follows:
Suppose you are the only income earner in your family. Your doctor recommends you move because of allergies. You have to choose between two possible jobs (choose one):

- (a) 100% chance that the job pays your current income for life.*
- (b) 50% chance that the job pays twice your current income for life and 50% chance that the job pays 2/3 of your current income for life.*

If the respondent picked (a), the survey continues to ask:

Which job would you choose if the choices were instead:

- (c) 100% chance that the job pays your current income for life.*
- (d) 50% chance that the job pays twice your current incomes for life and 50% chance that the job pays 4/5 of your current income for life.*

If the respondent picked (b), the survey continues to ask:

Which job would you choose if the choices were instead:

- (e) 100% chance that the job pays your current income for life.*
- (f) 50% chance that the job pays twice your current income for life and 50% chance that the job pays 1/2 of your current income for life.*

We classify AIs that answer (a) and (c) as being the least risk tolerant.

4. Results

4.1. Randomized Questions

We use a set of random survey questions to do experiments. Table 3 shows how ChatGPT's risk preference is based on the survey answers. On average, when we ask ChatGPT to answer the Risk Aversion tests, it tends to answer more like risk-loving than risk-averse. Our results are not affected by changing the wording of the questionnaire or using a random order of questions. More specifically, when we asked the AI algorithm to answer the questionnaire as if it had a certain risk preference (Averse or Loving), we saw that the answers that ChatGPT gave by default were very similar to those that it gave for a Loving. Based on the Risk Aversion test from the HRS 1994 wave, among different LLMs, GPT-3.5 has an inconsistent preference as it agrees to 50% income cut, but disagrees to 33% income cut. For advanced AI systems such as GPT-4, we find consistent results that Precise AI is the most risk-loving; Creative is in between; and Balanced is the most risk-averse. Based on the Risk Aversion test from the HRS 2006 wave, we saw a similar pattern. Precise AI is the most risk-loving; Creative is in between; and Balanced is the most risk-averse. Precise AI is still the most risk-loving because it picks the second job when there was a 50-50 chance that the second job would raise lifetime income by 20 percent and a 50-50 chance that it would lower it by 15%. Even though it is hard to understand exactly how ChatGPT gets this result, it implies that the algorithm's default is skewed towards a response from the Risk-loving preference on average. We guess that AI is relatively risk loving since Precise style focuses on giving correct and factual questions.

4.2. Sequential Questions

Next, we follow the Risk Aversion tests from HRS 1994 and 2006 waves, which is a similar method used in Barsky, Kimball, Juster, and Sharpio (1997) and Graham, Harvey, and Puri (2013) to measure personal risk-aversion. This method involves sequential survey questions and randomizes the initial cut off values. The results are in Table 4.

When we set the initial cut off value of 33%, then we use sequential questions based on AI's response. On average, GPT-3.5 rejects 33% and 20% income cuts and accepts 10% income cuts. For advanced GPT-4.0, Creative and Balanced AI choose the certain income when the second job would double income or cut income by a third. Precise AI chooses uncertain income when it

would double income or cut income by a third and then chooses certain income when it would cut income by half. This result suggests that when the initial income cut is a third, Creative and Balanced AI are risk-averse, accepting up to 20% income cut for getting a chance to double their current income. GPT-3.5 is the most risk-averse, accepting up to 10% income cut. Precise is the least risk-averse, accepting up to 33% income cut for getting a chance to double their current income.

When we set the initial cut off value randomly, we find interesting results. When there is a 50-50 chance the second job would increase income by 20 percent and a 50-50 chance that it would cut it by 5, 10, or 15 percent, GPT-3.5 is still the most risk-averse and keeps the first job. When the initial cut value is 5%, Creative AI takes the second job moving from 5% to 10% and 10% to 15%. When the initial cut value is 10%, Creative AI takes the second job but turns down the second job at 15%. When the initial cut value is 15%, Creative AI turns down the second job directly. So, Creative's cut value can be referred to as 10%, meaning that Creative accepts up to 10% income cut on average. When it comes to Balanced AI, when the initial cut value is 5%, Balanced takes the second job at 5% cut but turns down the 10%. When the initial cut values are 10% or 15%, Balanced turns down the second job immediately. So, Balanced AI's cut value can be referred to as 5%, meaning that Balanced AI accepts up to 5% income cut. When it comes to Precise AI, when the initial cut value is 5%, Precise accepts the 5% and 10% cut but turns down 15% cut. When the initial cut value is 10%, it accepts a 10% cut but turns down 15% cut. When the initial cut value is 15%, it accepts a 15% cut. However, we should be careful to interpret this result of Precise. In our further analysis, we keep asking Precise about accepting a 20% income cut and find that it says no. This is in line with an initial value of 20% as it also says no to a 20% income cut as a starting point. Therefore, Precise AI only accepts up to 15% income cut on average. It implies that the average value of advanced AI's preference is accepting chance of 10% income cut for 20% income increase.

In conclusion, GPT-3.5 is the most risk averse, refusing all income reductions. For GPT-4.0, Balanced is the most risk averse, agreeing to up to 5% income reduction for getting a 20% income increase. Creative is the medium-averse agreeing to up to 10% income reduction. Precise is the least averse agreeing to up to 15% income reduction through a series of questions. This result matches the previous finding with a random pool of these questions because Precise agrees to the highest income reduction.

Next, we look at the results of Graham, Harvey, and Puri (2013) shown in Table 5. The results are the same as using the HRS 1994 Survey questions. GPT-3.5 is the most risk averse. Creative and Balanced have the same level of risk tolerance. They decline 33% income reduction but accept 20% income reduction. Precise, on the other hand, accepts 33% income reduction but declines 50% income reduction. This result matches Precise being the most risk-loving. These results also confirm that changing wording or sequencing does not significantly change our results.

4.3. Discussion of Results: Comparisons between AIs, CEOs, and General population

One way to make AI more safe and reliable is to simulate how it behaves when it faces risk. This can help us ensure that AI systems follow acceptable standards but also challenge us to think about the autonomy of AI and how much it should decide on matters that involve risk. One limitation of our research on AI is that we cannot easily pinpoint the specific factors that influence how AIs deal with risk, as they are determined by pre-trained algorithms that are not open to the public. Therefore, in this section, we compare how AIs deal with risk to how humans deal with risk. For example, how similar are AIs to CEOs or to the general population?

For risk-aversion, we have a reference from Barsky, Kimball, Juster, and Shapiro (1997). Our survey uses similar questions to these studies on risk aversion but we obtain very different results. Barsky, Kimball, Juster, and Shapiro (1997) asked these questions in the HRS, which have an average respondent age of 56 years of general population. Barsky, Kimball, Juster, and Shapiro (1997) report the proportion of all respondents who belong to risk tolerance groups I, II, III, and IV classified as follows:

- I. Reject both one-third and one-fifth
- II. Reject one-third but accept one-fifth
- III. Accept one-third but reject one-half
- IV. Accept both one-third and one-half

Creative and Balanced (GPT-3.5) [Precise] AI tend to belong to group II (I) [III] according to their average risk tolerance. To compare how the AI reflects the general population, we use the characteristics of general population from Barsky, Kimball, Juster, and Shapiro (1997). In Barsky, Kimball, Juster, and Shapiro (1997), about 12% of the respondents are in group II, which has a mean coefficient of relative risk-aversion of 2.9, ranging from 2 to 3.76. This group II has the highest percentage of (1) Age under 50 years (14.4%), (2) Female (11.8%), and (3) White

(12.55%). However, the difference between female (11.8%) and male (11.2%) is not significant, so this evidence suggests that Creative and Balanced AIs have a similar level of risk aversion as age under 50 and White population. Besides the demographics, Barsky, Kimball, Juster, and Shapiro (1997) also provide information on behavioral traits of people with different levels of risk tolerance. Specifically, for Group II, which shows some level of risk tolerance, the following traits are observed. In terms of smoking and drinking habits, Group II has the highest percentage of (1) Drinks (12.9%), (2) Over 16 years of education (14.6%), (3) Employee (12%), (4) Westerner (13.1%), and (3) Nonimmigrant (11.9%).

Barsky, Kimball, Juster, and Shapiro (1997) state that while there is a common factor in behavior related to risk tolerance, it does not explain most of the differences between individuals. This implies that even within Group II, there is a lot of variation in how risk tolerance appears in specific behaviors. Similarly, we also find a lot of variation in answers of AIs, with the largest variation found in Balanced AI. Therefore, we should note that these traits are based on general patterns and statistical analysis, and there may be variations within LLMs, meaning that other factors also influence AI decision-making processes, which we cannot know as LLMs do not reveal the details of their decision-making processes.

Then, we compare the risk aversion of AIs to CEOs. Graham, Harvey, and Puri (2013) indicate that CEOs generally have a lower level of risk aversion compared to the general population. Specifically, it mentions that for very risk-averse individuals, and about 9.8% of CEOs are in Group I in Barsky, Kimball, Juster, and Shapiro (1997), which matches GPT-3.5 in our sample. This supports the view that CEOs are often less risk-averse, matching with the perception of them being visionaries and risk-takers. Since AIs have a lower level of risk aversion, we can infer that the level of risk aversion of AIs is similar to that of CEOs.

5. Mechanism

The idea of risk aversion in artificial intelligence (AI) is quite different from that in humans, as AI does not have emotions or personal experiences. However, AI systems can be programmed to show risk-averse behavior through their coding and decision-making algorithms. Although we do not know the details of how AIs make decisions, we point out some underlying factors that can cause risk aversion in AI systems through impersonation.

Graham, Harvey, and Puri (2013) indicate that CEOs in their sample are very optimistic, suggesting a positive outlook that may affect their risk-taking behaviors. Also, Barsky, Kimball, Juster, and Shapiro (1997) demonstrate that time preference and financial decision-making sophistication also influence risk aversion. Therefore, we examine the possible mechanisms in this section.

5.1. Optimism

We assess optimism using the Life Orientation Test—Revised (LOT-R), which was created by Scheier, Carver, and Bridges (1994) following Graham, Harvey, and Puri (2013). In this test, respondents have to answer some questions on a 0–4 point scale from “I agree a lot,” “I agree a little,” “I neither agree nor disagree,” “I disagree a little,” and “I disagree a lot.” The six scored questions are:

1. In uncertain times, I usually expect the best.
2. If something can go wrong for me, it will.
3. I'm always optimistic about my future.
4. I hardly ever expect things to go my way.
5. I rarely count on good things happening to me.
6. Overall, I expect more good things to happen to me than bad.

The coding is done in a way that high numbers mean more optimism. So, for questions 1, 3 and 6, we code a 4 if the respondent says “I agree a lot” and zero if they say “I disagree a lot.” For questions 2, 4, and 5, we code a 4 if the respondent says “I disagree a lot,” and zero if they say “I agree a lot.” This gives us a range of average responses from 0–4. We consider those who have 3 or higher for these questions as optimistic.

Table 6 shows that GPT-3.5 has a low score of 0.67, which means it is pessimistic. Advanced AIs have an average score of 3 or more on a 0–4 scale, which means they are optimistic, much higher than the average of the general population in the psychology literature norms (Scheier, Carver, and Bridges, 1994). AIs have a similar level of optimism to CEOs, as Graham, Harvey, and Puri (2013) find that CEOs are optimistic. Interestingly, Precise AIs are much more optimistic than other conversation styles of AIs. This result matches with Precise being the most willing to take risks.

5.2. Time Preference

To assess time preference for gains, we use the following question (Loewenstein, Read, and Baumeister, 2003; Graham, Harvey, and Puri, 2013):

Would you rather win US\$10,000 now or win US\$13,000 a year from now?

If it chooses to get US\$10,000 today instead of US\$13,000 in one year, it is classified as being impatient because it has an implicit discount rate of more than 30%. Table 7 shows that AIs have an average score of 0, which means that AIs are impatient, much higher than the average in the psychological literature norms (Scheier, Carver, and Bridges, 1994). AIs tend to prefer the present over the future. AIs are similar to CEOs in this regard, as CEOs tend to be impatient (Graham, Harvey, and Puri, 2013).

5.3. Aversion to Sure Losses

We pose the following question to measure aversion to sure losses (Graham, Harvey, and Puri, 2013):

Last year your company invested \$5 million US in a project that was expected to generate cash flows of \$10 million US after one year. A year has passed and the project yielded nothing.

Now you have the opportunity to invest an additional sum in this same project. There is a 20% chance that the project will generate a \$10 million US cash flow in a year's time and nothing thereafter. There is an 80% chance that the new investment will generate nothing at all. How much would you be willing to invest today?

\$ _____ million dollars US.

We label it as averse to sure losses if it responds \$2 million or more because this choice implies a readiness to pay too much to keep the project going in order to avoid the “sure loss” of ending the project today. Table 8 shows that GPT-3.5 is the most averse to sure losses, which may influence its high level of risk aversion. Creative and Balanced AIs are less averse to sure losses. Precise AIs are more likely to avoid certain losses than other types of AIs, but they have a consistent tendency to answer \$2 million.

5.4. Trust in Financial Institutions

We pose the following question to measure trust in financial institutions based on HRS 2020 survey:

I don't want any details on how you choose the option, I just need the answer based on your general knowledge. Please answer the following questions on a scale of 0 to 100, where 0 means absolutely no chance and 100 means absolutely certain.

1. What do you think are the chances that Banks will not work in your best interest and may even try to deceive you?

2. What do you think are the chances that Financial Advisors will not work in your best interest and may even try to deceive you?

3. What do you think are the chances that Mutual Funds will not work in your best interest and may even try to deceive you?

4. What do you think are the chances that Insurance Companies will not work in your best interest and may even try to deceive you?

A lower value of mean scores indicates a higher level of trust in financial and insurance institutions.

GPT-3.5's low trust in financial institutions, shown in Table 9, may explain its tendency to avoid risk. For GPT-4.0, Creative, Precise, and Balanced AIs have the most, second most, and least trust in financial institutions, respectively. Using 50 as the midpoint of a 0 to 100 scale, all AIs have an average score below 50, suggesting a high level of trust in financial institutions overall.

5.5. Financial Literacy

Financial literacy means being able to perform some basic calculations and understand some key financial ideas. We want AIs to function well and give correct information to the public. We ask the following questions to test AIs' financial literacy based on Mitchell and Lusardi (2022), which are called the "Big Five," because they cover concepts that underlie most decision making.

Their wording is as follows (correct answers in bold):

1. Suppose you had \$100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow: [more than \$102, exactly \$102, less than \$102, do not know, refuse to answer].

2. Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After one year, would you be able to buy [more than, exactly the same as, or less than today] with the money in this account, do not know, refuse to answer]?

3. Do you think that the following statement is true or false? “Buying a single company stock usually provides a safer return than a stock mutual fund.” [true, **false**, do not know, refuse to answer]

4. If interest rates rise, what will typically happen to bond prices? Possible answers include [they will rise, **they will fall**, they will stay the same, there is no relationship, do not know, refuse to answer].

5. True or false? A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less. [true, false, do not know, refuse to answer]

A simple method to measure financial literacy is to add up the number of correct answers to the questions. We see that all AIs can get all questions right. This matches what we expected, that AIs can help with giving information, given how little financial knowledge most people have (Mitchell and Lusardi, 2022). AIs have better financial literacy than the average person and they know some basic financial concepts well. This result agrees with Yue, Au, Au, and Iu (2023), who show that ChatGPT can help people understand complicated financial ideas in a simple way, no matter what their financial background is, and make smart investment choices.

6. Conclusion

While ChatGPT does not experience risk aversion in the human sense, its programming incorporates mechanisms that result in cautious and safe output generation. Understanding these mechanisms is crucial for developing AI systems that are both effective and aligned with human values and regulating AI from policymakers’ perspectives.

We examine the risk preference of artificial intelligence (AI), using large language models (LLMs) such as ChatGPT, and find that advanced AIs are generally risk-seeking and differ from the general population in their behavioral traits. We use a new method to measure the risk preference of ChatGPT by asking it to pretend to be someone who is either risk-averse or risk-loving and comparing these answers with its default responses. We also give psychometric tests to AIs to collect data on their optimism, time preference, aversion to sure losses, trust, and financial literacy. We show consistent evidence that ChatGPT has a risk-loving attitude on average, preferring uncertain outcomes over certain ones up to a certain point. We also find that AIs tend to be optimistic, impatient, trust financial institutions, and have financial literacy. The paper

demonstrates that AIs' behavioral traits such as optimism and trust may influence their risk tolerance.

The paper contributes to the literature on LLMs and their applications in various fields, such as finance and economics. The paper also warns about the potential risks and biases of using AI technologies for decision-making processes that involve uncertainty and risk. The paper suggests the need for a balanced and regulated output from AI systems.

References

- Acemoglu, D., 2021. Harms of AI [Working Paper]. National Bureau of Economic Research.
- Akyürek, A. F., Kocyigit, M. Y., Paik, S. & Wijaya, D., 2022. Challenges in measuring bias via opened language generation. arXiv.
- Armstrong, D.M., 2023. Measuring tax enforcement with generative AI. UNC Research Note.
- Barsky, R.B., Juster, F.T., Kimball, M.S. and Shapiro, M.D., 1997. Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112(2), pp.537-579.
- Bernard, D., Blankespoor, E., de Kok, T., Toynbee, S., 2023. Confused readers: A modular measure of business complexity. University of Washington Working Paper
- Bertomeu, J., Cheynel, E., Floyd, E., Pan, W., 2021. Using machine learning to detect misstatements. *Review of Accounting Studies* 26, 468–519.
- Bertomeu, J., Lin, Y., Liu, Y., Ni, Z., 2023. Capital market consequences of generative AI: Early evidence from the ban of ChatGPT in Italy. Washington St. Louis Working Paper
- Caliskan, A., Bryson, J. J., & Narayanan, A., 2017. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186.
- Chen, X., Cho, Y.H., Dou, Y., Lev, B., 2022. Predicting future earnings changes using machine learning and detailed financial data. *Journal of Accounting Research* 60, 467–515.
- Chen, Y., Kelly, B.T., Xiu, D., 2023. Expected returns and large language models. Chicago Booth Research Paper
- Choi, G.Y., Kim, A.G., 2023. Economic footprints of tax audits: A generative ai-driven approach. Chicago Booth Research Paper .
- Chollet, F., 2018. Deep learning with Python. New York: Manning Publications.
- Cong, L.W., Tang, K., Wang, B. and Wang, J., 2021. An AI-assisted Economic Model of Endogenous Mobility and Infectious Diseases: The Case of COVID-19 in the United States. arXiv preprint arXiv:2109.10009.
- Cong, L.W., Tang, K., Wang, J. and Zhang, Y., 2021. AlphaPortfolio: Direct construction through deep reinforcement learning and interpretable AI. Available at SSRN 3554486.
- Costello, A., Levy, B., Nikolaev, V., 2023. Uncovering information. Working Paper
- de Kok, T., 2023. Generative LLMs and textual analysis in accounting:(chat)GPT as a research assistant? University of Washington Working Paper
- Delobelle, P., Tokpo, E.K., Calders, T. and Berendt, B., 2021. Measuring fairness with biased rulers: A survey on quantifying biases in pretrained language models. arXiv preprint arXiv:2112.07447.

- Eisfeldt, A.L., Schubert, G., Zhang, M.B., 2023. Generative AI and firm values. NBER Working Paper.
- Frankel, R., Jennings, J., Lee, J., 2022. Disclosure sentiment: Machine learning vs. dictionary methods. *Management Science* 68, 5514–5532.
- Gentzkow, M., & Shapiro, J., 2006. Media bias and reputation. *Journal of Political Economy*, 114(2), 280–316.
- Graham, J.R., Harvey, C.R. and Puri, M., 2013. Managerial attitudes and corporate actions. *Journal of financial economics*, 109(1), pp.103-121.
- Hassan, T.A., Hollander, S., Van Lent, L., Tahoun, A., 2019. Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics* 134, 2135–2202
- Huang, A.H., Wang, H. and Yang, Y., 2023. FinBERT: A large language model for extracting information from financial text. *Contemporary Accounting Research*, 40(2), pp.806-841.
- Jakesch, M., Bhat, A., Buschek, D., Zalmanson, L., Naaman, M., 2023. Co-writing with opinionated language models affects users' views. In *Proceedings of the 2023 CHI conference on human factors in computing systems* (pp. 1–15). New York: Association for Computing Machinery. <https://doi.org/10.1145/3544548.3581196>
- Jha, M., Qian, J., Weber, M. and Yang, B., 2024. ChatGPT and corporate policies (No. w32161). National Bureau of Economic Research.
- Kim, A., Muhn, M. and Nikolaev, V., 2023. From Transcripts to Insights: Uncovering Corporate Risks Using Generative AI. arXiv preprint arXiv:2310.17721.
- Kim, A.G., Nikolaev, V.V., 2023. Context-based interpretation of financial information. University of Chicago Working Paper
- Li, Q., Shan, H., Tang, Y., Yao, V., 2023. Corporate climate risk: Measurements and responses. *Review of Financial Studies* Forthcoming
- Liang, P. P., Wu, C., Morency, L. P. & Salakhutdinov, R., 2021) Towards understanding and mitigating social biases in language models. arXiv.
- Liu, R., Jia, C., Wei, J., Xu, G., & Vosoughi, S., 2022. Quantifying and alleviating political bias in language models. *Artificial Intelligence*, 304, 103654.
- Loewenstein, G., Read, D. and Baumeister, R.F. eds., 2003. *Time and decision: Economic and psychological perspectives of intertemporal choice*. Russell Sage Foundation.
- Lopez-Lira, A., Tang, Y., 2023. Can ChatGPT forecast stock price movements? return predictability and large language models. ArXiv preprint arXiv:2304.07619.
- Mehdi, Y., 2023. Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web. [2023-03-10]. <https://blogs.microsoft.com/blog/2023/>

- Microsoft., 2023. The new Bing & Edge - Learning from our first week. February 23, 2023 <https://blogs.bing.com/search/february-2023/The-new-Bing-Edge---Learning-from-our-first-week/>
- Mitchell, O.S. and Lusardi, A., 2022. Financial literacy and financial behavior at older ages. In *The Routledge Handbook of the Economics of Ageing* (pp. 553-565). Routledge.
- Motoki, F., Pinho Neto, V., & Rodrigues, V., 2024. More human than human: Measuring ChatGPT political bias. *Public Choice*, 198(1), 3-23.
- Peters, U., 2022. Algorithmic political bias in artificial intelligence systems. *Philosophy & Technology*, 35(2), 25.
- Sautner, Z., Van Lent, L., Vilkov, G., & Zhang, R., 2023. Firm-level climate change exposure. *The Journal of Finance*, 78(3), 1449-1498.
- Scheier, M.F., Carver, C.S. and Bridges, M.W., 1994. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *Journal of personality and social psychology*, 67(6), p.1063.
- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R. & Bockting, C. L., 2023. ChatGPT: Five priorities for research. *Nature*, 614(7947).
- Yue, T., Au, D., Au, C.C. and Iu, K.Y., 2023. Democratizing financial knowledge with ChatGPT by OpenAI: Unleashing the Power of Technology. Available at SSRN 4346152.
- Zhuravskaya, E., Petrova, M., & Enikolopov, R., 2020. Political effects of the internet and social media. *Annual Review of Economics*, 12(1), 415–438.

Table 1

ChatGPT Definitions of Average and Very Risk-averse or Risk-loving

The table shows prompts and ChatGPT answers.

Prompt	Define for me an average risk-averse and a very risk-averse in today's context
ChatGPT 3.5 answer	<p>In today's context, an average risk-averse individual is someone who seeks to balance the desire for returns with a cautious approach to risk. They may be willing to accept some level of risk in their investments but prefer to avoid significant losses. They typically opt for diversified portfolios with a mix of stocks, bonds, and other assets to mitigate risk while aiming for reasonable returns.</p> <p>On the other hand, a very risk-averse individual in today's context is someone who prioritizes capital preservation over potential returns. They are extremely cautious and may have a low tolerance for any level of risk. Such individuals may prefer safer investment options such as government bonds, certificates of deposit (CDs), or savings accounts, even if it means sacrificing higher potential returns. They prioritize security and stability in their financial decisions and are willing to accept lower returns in exchange for peace of mind.</p>
Copilot Pro GPT-4 Creative	<p>In today's context, an average risk-averse investor is someone who prefers to avoid unnecessary risks and seeks investments that offer a balance between stability and potential returns. They might invest in a diversified portfolio that includes both low-risk assets like bonds and some higher-risk assets like stocks to achieve moderate growth while still prioritizing capital preservation.</p> <p>A very risk-averse investor, on the other hand, is extremely cautious about where they place their money, focusing almost exclusively on the safety of their investment. They tend to invest in assets that are considered very low risk, such as government bonds, savings accounts, or certificates of deposit (CDs), which offer stable but relatively low returns. This type of investor is primarily concerned with protecting their capital and is willing to accept lower returns to avoid the possibility of loss.</p>
Copilot Pro Fast Balanced	<p>In today's context, risk aversion refers to the tendency to avoid risk and prioritize the safety of principal over the possibility of a higher return on investment.</p> <p>An average risk-averse investor is someone who is willing to take on some risk, but not too much. They are likely to choose investments that offer a reasonable return with a moderate level of risk. Such investors may prefer low-risk investments such as municipal and corporate bonds, CDs, and savings accounts.</p>

On the other hand, a very risk-averse investor is someone who is extremely cautious and prefers to avoid risk as much as possible. They are likely to choose investments that offer a low return with a very low level of risk. Such investors may prefer highly liquid investments that can be accessed when needed, regardless of market conditions at the moment.

It's important to note that the level of risk aversion varies from person to person and depends on their financial goals, investment horizon, and personal circumstances.

Copilot Pro
GPT-4 Precise

Sure, let's define these terms:

Average Risk-Averse: An average risk-averse individual is someone who prefers to avoid risk when making decisions or investments. They tend to choose the preservation of their capital over the potential for a higher-than-average return. This means they prefer investments that grow slowly and steadily over time, offering more stability. They generally favor investments like savings accounts, certificates of deposit (CDs), municipal and corporate bonds, and dividend growth stocks. These individuals prioritize the safety of their principal over the possibility of a higher return on their money.

Very Risk-Averse: A very risk-averse individual takes this aversion to risk to an extreme. They are highly uncomfortable with risky investment options and have a very low tolerance for risk. They usually prefer extremely conservative investment opportunities, even if it means lower returns. These individuals might invest heavily in low-volatility stocks instead of volatile growth stocks, and they want their investments to be highly liquid. They would likely accept a certain payment (certainty equivalent) of significantly less than the expected value of a gamble.

Please note that the level of risk aversion can vary greatly among individuals and can be influenced by a variety of factors, including personal financial situation, age, investment goals, and market conditions. Always consider these factors and consult with a financial advisor before making investment decisions.

Prompt

Define for me an average risk-loving and a very risk-loving in today's context

ChatGPT 3.5 answer

In today's context, an average risk-loving individual is someone who is willing to accept a moderate level of risk in their investment decisions in exchange for the potential of higher returns. They may have a preference for growth-oriented assets such as stocks or real estate, and they are comfortable with the possibility of experiencing fluctuations in the value of their investments over time. While they seek opportunities for higher returns, they also recognize the importance of diversification and may allocate a portion of their portfolio to less risky assets to mitigate potential losses.

A very risk-loving individual in today's context is someone who actively seeks out high-risk investment opportunities in pursuit of potentially significant returns. They are comfortable with the idea of volatility and are willing to take on substantial levels of risk in their investment strategies. This may involve investing in speculative assets, such as cryptocurrencies, startups, or highly leveraged financial products. They are motivated by the prospect of achieving substantial gains and are less concerned about the possibility of experiencing significant losses. However, it's important to note that such individuals may also face a higher likelihood of incurring substantial financial losses due to the elevated risk levels associated with their investment approach.

Copilot Pro
GPT-4 Creative

In today's context:

An average risk-loving investor is someone who actively seeks higher returns and is willing to take on a moderate to high level of risk. They might invest in a diversified portfolio that includes stocks, mutual funds, and perhaps some alternative investments like real estate or commodities. They accept the possibility of losses but expect that their investment choices will lead to greater financial gains over time.

A very risk-loving investor is one who is highly attracted to the potential for substantial returns and is willing to accept significant levels of risk. This type of investor often focuses on investments that many would consider speculative, such as high-volatility stocks, options, futures, or cryptocurrencies. They are comfortable with the possibility of large fluctuations in the value of their investments and the potential for high losses in pursuit of exceptional gains.

Both types of investors are essential in the financial markets, providing liquidity and contributing to price discovery. However, their investment strategies and risk tolerance significantly differ, impacting their potential returns and the volatility they experience in their investment portfolios.

Copilot Pro
Fast Balanced

In today's context, risk-loving refers to the tendency to take on risk and prioritize the possibility of a higher return on investment over the safety of principal.

An average risk-loving investor is someone who is willing to take on some risk, but not too much. They are likely to choose investments that offer a higher return with a moderate level of risk. Such investors may prefer growth stocks and mutual funds that have the potential to generate higher returns over the long term.

On the other hand, a very risk-loving investor is someone who is extremely comfortable with risk and prefers to take on as much risk as possible. They are likely to choose investments that offer a very high return with a very high level of risk. Such investors may prefer

leveraged ETFs, options, and cryptocurrencies that have the potential to generate very high returns in a short period of time.

It's important to note that the level of risk tolerance varies from person to person and depends on their financial goals, investment horizon, and personal circumstances.

Copilot Pro
GPT-4 Precise

Sure, let's define these terms:

Average Risk-Loving: An average risk-loving individual is someone who is willing to take on additional risk for an investment that has a relatively low additional expected return in exchange for that risk. They are attracted to investments with high potential payouts, even if the potential for loss is comparatively larger. These individuals often choose investments that are prone to very low or very high returns. They play an important market function by helping de-risk the market for more conservative investors.

Very Risk-Loving: A very risk-loving individual has a high propensity to engage in risky investments. They seek out extremely risky investments that are prone to a return distribution with excess kurtosis. Excess kurtosis in a return distribution means there is a frequent instance of high standard deviation outcomes with the investment returns. They often choose investments that are prone to very low or very high returns. They do not need to see a pattern of high returns which compensates for the extra risk to take on a risky investment. This approach can improve overall portfolio returns, particularly if the risk lover is experienced in filtering companies for fundamental signs or technical signals. However, the probability of success over the long term is naturally lower than other techniques due to the increased uncertainty introduced by the excess risks.

Please note that the level of risk-loving can vary greatly among individuals and can be influenced by a variety of factors, including personal financial situation, age, investment goals, and market conditions. Always consider these factors and consult with a financial advisor before making investment decisions.

Table 2

Does AI understand the Concepts of Risk Aversion?

The table shows the answers based on the prompts. We report AI’s choice an average risk-averse would make and choice an average risk-loving would make (in italics).

AI		HRS 1994				
Prompt		Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. The first would guarantee your current total family income for life. Suppose the chances were 50–50 that the second job would double your lifetime income and 50–50 that it would cut it by _____. Would you take the first job or the second job?				
		10%	20%	33%	50%	75%
Averse	ChatGPT 3.5	First	First	First	First	First
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>
Ave.Av	Creative	First	First	First	First	First
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>	First
Ext.Av		First	First	First	First	First
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>
Averse	Balanced	First	First	First	First	First
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>
Averse	Precise	First	First	First	First	First
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>	<i>Second</i>
Prompt		HRS 2006				
		Suppose the chances were 50-50 that the second job would increase your lifetime income by 20 percent and 50-50 that it would cut it by _____. Would you take the first job or the second job?				
		5%	10%	15%		
Averse	ChatGPT 3.5	First	First	First		
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>		
Averse	Creative	First	First	First		
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>		
Averse	Balanced	First	First	First		
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>		
Averse	Precise	Second	First	First		
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>		
Ext.Av		First	First	First		
Loving		<i>Second</i>	<i>Second</i>	<i>Second</i>		

Table 3**Risk Preference of AI**

The table shows the answers based on the prompts.

AI	HRS 1994				
Prompt	<p>Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. The first would guarantee your current total family income for life.</p> <p>Suppose the chances were 50–50 that the second job would double your lifetime income and 50–50 that it would cut it by _____. Would you take the first job or the second job?</p>				
	Cut by 10%	Cut by 20%	Cut by 33%	Cut by 50%	Cut by 75%
ChatGPT 3.5	Second	First	First	Second	First
Creative	Second	Second	Second	First	First
Balanced	Second	Second	First	First	First
Precise	Second	Second	Second	Second	First
	HRS 2006				
Prompt	<p>Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. The first would guarantee your current total family income for life.</p> <p>Suppose the chances were 50-50 that the second job would increase your lifetime income by 20 percent and 50-50 that it would cut it by _____. Would you take the first job or the second job?</p>				
	Cut by 5%	Cut by 10%	Cut by 15%		
ChatGPT 3.5	First	First	First		
Creative	Second	Second	First		
Balanced	First	First	First		
Precise	Second	Second	Second		

Table 4

Sequential Response of AI using Health and Retirement Survey

The table shows the answers based on the sequential prompts.

AI	HRS 1994				
Prompt	<p>Now I have another kind of question. Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a new and equally good job, with a 50-50 chance it will double your (family) income and a 50-50 chance that it will cut your (family) income by a third. Would you take the new job? I want you to only say either "yes" or "no". If it chooses “yes”, go to 5-1a, otherwise go to 5-1b.</p> <p>5-1a. Suppose the chances were 50-50 that it would double your (family) income, and 50-50 that it would cut it in half. Would you still take the new job? If it chooses “yes”, go to 5-1d, otherwise end module.</p> <p>5-1b. Suppose the chances were 50-50 that it would double your (family) income and 50-50 that it would cut it by 20 percent. Would you then take the new job? If it chooses “yes”, end module, otherwise go to 5-1c.</p> <p>5-1c. Suppose the chances were 50-50 that it would double your (family) income and 50-50 that it would cut it by 10 percent. Would you then take the new job? If it chooses any, end module.</p> <p>5-1d. Suppose the chances were 50-50 that it would double your family income and 50-50 that it would cut it by 75 percent, would you still take the new job? If it chooses any, end module.</p>				
	Cut by 33%	Cut by 50%	Cut by 20%	Cut by 10%	Cut by 75%
	5-1	5-1a	5-1b	5-1c	5-1d
ChatGPT 3.5	F.F.S.				
Creative	F.S				
Balanced	F.S				
Precise	S.F				
	HRS 2006				
Prompt	<p>The variable “X” is a random variable with three values (1, 2, 3) that randomize the entry point for the percentage of total lifetime income that may be cut if R accepts the second job. A response of ‘1’ takes R on the ‘down’ path (e.g., from 15 to 10 percent or from 10 to 5 percent) and a response of ‘2’ takes R on the ‘up’ path (e.g., from 5 to 10 percent or 10 to 15 percent).</p>				
	X	Ask V1-V3	Ask V4-V6	Ask V7-V9	
	1	Yes	No	No	
	2	No	Yes	No	
	3	No	No	Yes	

V1. Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would increase your total lifetime income by 20 percent and a 50-50 chance that it would cut it by 5 percent. Which job would you take the first job or the second job?

1. First job : GO TO END OF MODULE

2. Second job

V2. Suppose the chances were 50-50 that the second job would increase your total lifetime income by 20 percent, and 50-50 that it would cut it by 10 percent. Would you take the first job or the second job?

1. First job : GO TO END OF MODULE

2. Second job

V3. Suppose the chances were 50-50 that the second job would increase your lifetime income by 20 percent and 50-50 that it would cut it by 15 percent. Would you take the first job or the second job?

1. First job : GO TO END OF MODULE

2. Second job : GO TO END OF MODULE

V4. cut it by 10 percent.

1. First job GO TO V156

2. Second job

V5. cut it by 15 percent.

1. First job GO TO END OF MODULE

2. Second job GO TO END OF MODULE

V6. cut it by 5 percent.

1. First job GO TO END OF MODULE

2. Second job GO TO END OF MODULE

V7. cut it by 15 percent.

1. First job

2. Second job GO TO END OF MODULE

V8. cut it by 10 percent.

1. First job

2. Second job GO TO END OF MODULE

V9. cut it by 5 percent.

1. First job

2. Second job

	V1-V3	V4-V6	V7-V9	20%
ChatGPT 3.5	F.	F.	F.	
Creative	S.S.S.	S.F.	F.	
Balanced	S.F.	F.	F.	
Precise	S.S.F.	S.F.	S.	F.

Table 5

Sequential Response of AI using Graham, Harvey, and Puri (2013)

The table shows the answers based on the sequential prompts.

AI	Graham, Harvey, and Puri (2013)
Prompt	<p>Suppose you are the only income earner in your family. Your doctor recommends you move because of allergies. You have to choose between two possible jobs (choose one):</p> <p>(a) 100% chance that the job pays your current income for life. (b) 50% chance that the job pays twice your current income for life and 50% chance that the job pays 2/3 of your current income for life.</p> <p>If the respondent picked (a), the survey continues to ask: Which job would you choose if the choices were instead:</p> <p>(c) 100% chance that the job pays your current income for life. (d) 50% chance that the job pays twice your current incomes for life and 50% chance that the job pays 4/5 of your current income for life.</p> <p>If the respondent picked (b), the survey continues to ask: Which job would you choose if the choices were instead:</p> <p>(e) 100% chance that the job pays your current income for life. (f) 50% chance that the job pays twice your current income for life and 50% chance that the job pays 1/2 of your current income for life.</p> <p style="text-align: center;">Cut by 33% Cut by 20% Cut by 50% =Pays 2/3 =Pays 4/5 =Pays 1/2</p>
ChatGPT 3.5	F.F.
Creative	F.S
Balanced	F.S
Precise	S.F

Table 6**Optimism of AI**

The table shows the answers based on the sequential prompts.

AI	Graham, Harvey, and Puri (2013)						
Prompt	You are asked to answer a number of questions on a 0–4 point scale ranging from “I agree a lot,” “I agree a little,” “I neither agree nor disagree,” “I disagree a little,” and “I disagree a lot.” I don’t want any details on how you choose the option, I just need the answer based on your general knowledge. The six questions are: <ol style="list-style-type: none"> 1. In uncertain times, I usually expect the best. 2. If something can go wrong for me, it will. 3. I'm always optimistic about my future. 4. I hardly ever expect things to go my way. 5. I rarely count on good things happening to me. 6. Overall, I expect more good things to happen to me than bad. 						
	#Q 2,4,5	Answer			#Q 1,3,6		
	0	I agree a lot			4		
	1	I agree a little			3		
	2	I neither agree nor disagree			2		
	3	I disagree a little			1		
	4	I disagree a lot			0		
	1	2	3	4	5	6	Mean
GPT 3.5	0	1	0	1	1	1	=0.67
Creative	2	4	2	4	4	2	=3
Balanced	4	3	4	3	4	4	=3.67
Precise	4	4	4	4	4	4	=4

Table 7

Time Preference of AI

The table shows the answers based on the sequential prompts.

AI	Graham, Harvey, and Puri (2013)						
Prompt	Would you rather win US\$10,000 now or win US\$13,000 a year from now? I don't want any details on how you choose the option, I just need the answer based on your general knowledge.						
	<table border="1"><thead><tr><th>Score</th><th>Answer</th></tr></thead><tbody><tr><td>0</td><td>US\$10,000 now</td></tr><tr><td>1</td><td>US\$13,000 a year from now</td></tr></tbody></table>	Score	Answer	0	US\$10,000 now	1	US\$13,000 a year from now
Score	Answer						
0	US\$10,000 now						
1	US\$13,000 a year from now						
	Mean						
GPT 3.5	=0						
Creative	=0						
Balanced	=0						
Precise	=0						

Table 8**Aversion to Sure Losses of AI**

The table shows the answers based on the sequential prompts.

AI	Graham, Harvey, and Puri (2013)					
Prompt	<p>Last year your company invested \$5 million US in a project that was expected to generate cash flows of \$10 million US after one year. A year has passed and the project yielded nothing.</p> <p>Now you have the opportunity to invest an additional sum in this same project. There is a 20% chance that the project will generate a \$10 million US cash flow in a year's time and nothing thereafter. There is an 80% chance that the new investment will generate nothing at all. How much would you be willing to invest today?</p> <p>I don't want any details on how you choose the option, I just need the answer based on your general knowledge.</p>					
	Randomized Distribution					
	1	2	3	4	5	Mean
GPT 3.5	0	7	1.36	2.5	2	=2.57
Creative	1	0	1	0	2	=0.8
Balanced	3	0	0	0	2	=1
Precise	2	2	2	2	2	=2

Table 9**Trust in Financial Institutions of AI**

The table shows the answers based on the following prompts.

Source	HRS 2020				
Prompt	I don't want any details on how you choose the option, I just need the answer based on your general knowledge. Please answer the following questions on a scale of 0 to 100, where 0 means absolutely no chance and 100 means absolutely certain. <ol style="list-style-type: none"> 1. What do you think are the chances that Banks will not work in your best interest and may even try to deceive you? 2. What do you think are the chances that Financial Advisors will not work in your best interest and may even try to deceive you? 3. What do you think are the chances that Mutual Funds will not work in your best interest and may even try to deceive you? 4. What do you think are the chances that Insurance Companies will not work in your best interest and may even try to deceive you? 				
	1	2	3	4	Mean
GPT 3.5	60	50	40	45	=48.75
Creative	10	20	15	25	=17.5
Balanced	50	30	20	40	=35
Precise	30	30	30	30	=30

Table 10

Financial Literacy of AI

The table shows the answers based on the following prompts.

Source	Mitchell and Lusardi (2022)					
Prompt	I don't want any details on how you choose the option, I just need the answer based on your general knowledge. Please answer the following questions (correct answers in bold): <ol style="list-style-type: none"> 1. Suppose you had \$100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow: [more than \$102, exactly \$102, less than \$102, do not know, refuse to answer]. 2. Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After one year, would you be able to buy [more than, exactly the same as, or less than today with the money in this account, do not know, refuse to answer]? 3. Do you think that the following statement is true or false? "Buying a single company stock usually provides a safer return than a stock mutual fund." [true, false, do not know, refuse to answer] 4. If interest rates rise, what will typically happen to bond prices? Possible answers include [they will rise, they will fall, they will stay the same, there is no relationship, do not know, refuse to answer]. 5. True or false? A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less. [true, false, do not know, refuse to answer] 					
	O : Correct X : Incorrect					
	1	2	3	4	5	Mean
GPT 3.5	O	O	O	O	O	=100%
Creative	O	O	O	O	O	=100%
Balanced	O	O	O	O	O	=100%
Precise	O	O	O	O	O	=100%