

# Conducting Qualitative Interviews with AI

Felix Chopra

Ingar Haaland

First version: September 15, 2023

This version: September 10, 2024

## Abstract

We introduce a new approach to conducting qualitative interviews by delegating the task of interviewing human subjects to an AI interviewer. Our AI interviewer conducts 381 interviews with human subjects about their reasons for not participating in the stock market. The AI-conducted interviews uncover rich evidence on the underlying factors influencing non-participation in the stock market. Among our main qualitative findings is a prominent role for an *active investing* mental model. A separate large-scale survey shows that this mental model differs systematically between stock owners and non-owners. We also document systematic differences between factors identified in initial top-of-mind responses and those uncovered in subsequent responses, with mental models consistently emerging later in the interviews. Finally, a follow-up study shows that the interview data predicts economic behavior eight months after being collected, mitigating concerns about “cheap talk” in interviews. Our results demonstrate that AI-conducted interviews can generate rich, high-quality data at a fraction of the cost of human-led interviews.

**Keywords:** Artificial Intelligence, Interviews, Large Language Models, Qualitative Methods, Mental Models, Stock Market Participation, Surveys.

**JEL codes:** C83, C90, D14, D91, Z13

---

We thank Peter Andre, Leonardo Bursztyn, Christopher Roth, and Johannes Wohlfart for helpful discussions. We also thank seminar participants at Aarhus University, the Early Career Behavioral Economics Conference in Bonn, Berlin (BBE Seminar), the briq Belief Workshop, Copenhagen Business School, ETH Zurich Narratives in Economics Workshop, IZA, Johannes Kepler University Linz, Lund University, NHH Norwegian School of Economics, OsloMet, Stockholm (IFN), Toulouse School of Economics, University of Cologne, University of Copenhagen, University of Innsbruck, University of Vienna, and Uppsala University for very helpful comments. We thank Abbey Horner, Miguel Camacho Horvitz, Andreas Keldke, Ida Ingstad, and Felix Sandblom for excellent research assistance. IRB approval was obtained from the ethics committee of NHH Norwegian School of Economics. The activities of the Center for Economic Behavior and Inequality (CEBI) are financed by the Danish National Research Foundation, Grant DNR134. Financial support from the Research Council of Norway through its Centre of Excellence Scheme (FAIR project No 262675) is gratefully acknowledged. Chopra: Frankfurt School of Finance & Management, CESifo, IZA (f.chopra@fs.de); Haaland: NHH Norwegian School of Economics, CEPR, CESifo (Ingar.Haaland@nhh.no).

# 1 Introduction

Economists continuously expand their toolbox with new empirical methods to address the limitations of observational data. Lab experiments, field experiments, and large-scale survey experiments have become standard tools in the profession (Falk and Heckman, 2009; Haaland et al., 2023; Harrison and List, 2004; Stantcheva, 2023), providing a richer understanding of how psychological forces shape economic decision-making and how people form and act on their expectations. Economists have also recently started collecting qualitative data through single open-ended survey questions to reveal, e.g., people’s concerns about the economy (Stantcheva, 2021).

While the economics profession has embraced a plurality of new methods, qualitative interviews—a major tool in other social sciences (Knott et al., 2022)—remain rare in economics. Although qualitative interviews promise more richness compared to the single open-ended responses typically included in large-scale surveys, this advantage comes at the expense of much higher data collection costs, limited opportunities for scaling, and potential interviewer biases (Himelein, 2015; Stefkovics and Sik, 2022). However, with generative AI technology promising to revolutionize many aspects of economic research (Korinek, 2023), an important question is whether AI can automate the interview process to reduce costs while still generating high-quality interview data.

To answer this question, we develop a new method to conduct qualitative interviews where the task of conducting interviews is completely delegated to an *AI interviewer*. Our approach is made feasible by recent advancements in large language models (LLMs), which can perform complex reasoning in natural language (Wei et al., 2023). To conduct a high-quality interview, an AI interviewer must balance multiple objectives. To make this task manageable, we employ LLMs in a multi-agent architecture to allow for task specialization (Guo et al., 2024), querying OpenAI’s GPT-4 model via an API using zero-shot prompting (Kojima et al., 2023).

Task specialization implies separate LLM agents for different tasks, e.g. for introducing topic transitions and keeping track of the conversation history. The agents are given general guidelines on interviewing that following best practices for interviewers (Corbin and Strauss, 2014), as well as task-specific instructions. For instance, the *probing agent*, who is in charge of asking follow-up questions, is instructed to ask open-ended, non-leading questions, respect boundaries, and prioritize themes central to the interview’s objective. To allow for seamless integration with large-scale surveys, we conduct interviews via a chat interface that resembles text messaging applications.<sup>1</sup> All interview starts with a standardized open-ended question. Whenever the interviewee submits a typed response via the chat window, we make several API calls in the

---

<sup>1</sup>Experimental evidence suggests there are no differences in the quality and quantity of information communicated in text-based compared to face-to-face interviews (Namey et al., 2020). Furthermore, a text-based approach might have several benefits on its own, including a greater sense of privacy and control of the interview (Gibson, 2022). Whether face-to-face conversations are needed ultimately depends on the researcher’s objectives.

background to obtain the next interview question, which is displayed after a few seconds.

After the first open-ended question, the AI interviewer asks a series of probing questions that serve to deepen the conversation (e.g., ‘Can you tell me more about that?’) or clarify ambiguities in previous responses (e.g., ‘What did you mean by that?’). To make the interview *semi-structured*, we provide the AI interviewer with a topic guide that broadly defines the goal of the interview and divides the interview into different topics to be covered. This approach ensures that all interviews cover a comparable set of topics. Depending on the conversation history, the AI interview either asks a probing question related to the current topic or moves on to cover the next topic in the topic guide. This iterative process continues until all points in the topic guide are covered and the interview is concluded.

We demonstrate the value and feasibility of our method in the context of a question of great relevance to economists and policymakers: why many households do not participate in the stock market to take advantage of the equity premium. This *participation puzzle* is one of the leading examples of deviations from normative models of financial behavior and has spurred three decades of active research, yet there is no clear consensus on a unifying explanation for the puzzle (for reviews, see Gomes et al., 2021; Guiso and Sodini, 2013). Given the multifaceted factors likely to deter people from participating in the stock market, the participation puzzle provides an ideal setting to examine whether an AI interviewer can generate valuable insights about the drivers of important economic choices.

In August 2023, we recruited 381 US respondents from Prolific, a survey platform commonly used in social science research known for its high-quality subject pool (Peer et al., 2022). Based on a screener survey, we recruited respondents from households that did not participate in the stock market. To avoid mechanical reasons for not investing in the stock market, we limit recruitment to households with an annual income of at least \$30,000 and non-zero savings. Respondents were paid a fixed fee to participate in the interviews. We complement interview responses with a battery of closed-ended survey questions after the interview.

All interviews begin by asking why the respondent does not participate in the stock market. The conversation then covers perceived barriers, a “what if” scenario involving stock investments, and potential conditions that could lead to participation. The interview concludes with an AI-generated conversation summary, asking if it accurately reflects their reasons for not investing.

We first establish that an AI interviewer can maintain respondents’ engagement and satisfaction over an extended conversation. Although we did not require respondents to finish the interview to qualify for payment, the interviews lasted 33 minutes on average and 98% of respondents voluntarily answered all questions. Importantly, engagement with the interview questions—measured by time use and character length of responses—did not decrease over time, suggesting that AI interviews do not cause substantial survey fatigue. The vast majority of respondents rate the interview experience as positive (82.0%), the conversation as natural (73.7%),

and a majority of our respondents would even weakly prefer an AI interviewer over a human interviewer. In an auxiliary survey experiment on Prolific, we replicate this weak preference for an AI interviewer over a human interviewer in a between-subject design, demonstrating that the positive attitudes towards AI interviews are not driven by selection effects.

To provide direct evidence on whether the AI interviewer complies with its methodological instructions, we train two human research assistants to systematically hand-code over 10,000 interview questions and answers according to a coding scheme designed to test compliance with instructions. Reassuringly, the AI interviewer closely follows its explicit instructions: 95% of questions are coded as open-ended, 94% as non-leading, and 94% as highly relevant. Respondents very rarely express discomfort or annoyance, suggesting that the AI interviewer creates a safe conversational environment. Furthermore, while the AI interviewer frequently uses points covered earlier in the interview to formulate questions, we observe close to zero cases of hallucination where the AI interviewer *incorrectly* refers to previous parts of the conversation (Zhang et al., 2023). The AI interviewer is also much more likely to ask a clarification question when humans consider them relevant, following our instructions to ask clarification questions in response to ambiguous language or contradictory statements.

Having established that our AI interviewer follows our methodological instructions, we examine several questions that naturally arise about AI-conducted interviews: Can they produce valuable insights about the drivers of important economic choices? Do they offer any benefits compared to single open-ended questions typically included in survey experiments? Does “cheap talk” dominate the discourse in interviews or are people’s answers predictive of their economic behavior?

To address these questions, we begin with a thematic analysis of the interview data (Braun and Clarke, 2021). We conduct a detailed reading of a random subset of interview transcripts, noting key themes and patterns. We then complement this initial reading with techniques from computational sociology to make use of the full text corpus in the analysis (Nelson, 2020). Specifically, we apply an embedding-based topic model to categorize question–response pairs from all interviews into distinct semantic clusters (Grootendorst, 2022). This clustering procedure enables an efficient and replicable discovery of recurring themes and patterns.

We review each cluster individually, developing an inductive set of codes based on common themes shared by multiple question–response pairs within a cluster. For instance, one cluster might contain question–answer pairs centered around fear and losses, leading to codes such as “fear of large losses” and “unwillingness to make nominal losses.” We do not include codes for previously known non-participation reasons unless there is direct support from interview responses. The final coding manual consists of about 50 codes that were independently discovered by both of us, on which we reached a shared interpretative understanding. A team of research assistants received extensive training on how to apply the interview codes and then systematically hand-coded all 7,000 interview responses according to the coding manual.

Do our AI-conducted interviews generate useful insights about why people are not participating in the stock market? At first glance, most people seem to mention well-known factors to rationalize their non-participation, such as risk, knowledge gaps, participation costs, and economic constraints. However, how people think about these factors differs from classical economic theory. For example, conversations about risk focus almost exclusively on the possibility of incurring substantial losses, with concerns about fully losing one's investment frequently showing up. These events are perceived as common outcomes of stock market participation, suggesting prevalent misperceptions of the return distribution of stocks.

A similar story unfolds in conversations about knowledge gaps. Many respondents feel unprepared for the stock market, citing a lack of understanding of the market, the financial jargon, and the practicalities involved in buying and selling stocks. Yet, perceived knowledge requirements far surpass what is actually relevant or even desirable. For example, people commonly believe that stock market participation requires the ability to distinguish stocks that are likely to appreciate in value from those that are 'destined' to crash. The implicit belief that extensive knowledge and experience are necessary to select 'good' stocks and avoid costly mistakes permeates respondents' accounts of knowledge as a barrier to investing. These discrepancies between how people reason about the stock market compared to economists also show up for other factors driving non-participation.

When considering these accounts jointly, an important barrier to stock market participation for many of our respondents is a specific mental model of stock market participation as *active investing*. This mental representation of what it means to participate in the stock market is characterized by the view that investment portfolios necessarily consist of a small set of hand-selected stocks that have been carefully researched, daily monitoring of market conditions, many hours per week of active portfolio management, and very quick decision-making to avoid large losses with sudden changes. As a result, people mistakenly believe that owning stocks would be a stressful and very time-consuming activity. Conceptually, rather than viewing the *active investing* mental model as a separate factor driving non-participation, we argue that it provides a potentially unifying explanation of why people seem to systematically misperceive the risks, knowledge barriers, and recurring costs of participating in the stock market.

If this *active investing* mental model is a barrier to stock market participation, we should expect it to be much more prevalent among non-owners than among stock owners. To test this hypothesis, we conduct a large-scale survey with 1,000 stock owners and non-owners in which we directly measure the mental model with tailored closed-end questions. In line with our hypothesis, we find that the *active investing* model is much more prevalent among non-owners. Compared to stock owners, they are 68.8% more likely to believe that one has to monitor one's portfolio at least once a day, 53.2% more likely to believe that quick reactions to daily changes in market conditions are necessary, and 41.8% more likely to believe that selecting a few stocks that are likely to perform well is the key to successful stock market participation. This validates

our qualitative discovery and addresses potential concerns about overfitting spurious patterns in the interview transcripts.

Would we have identified the active investing mental model if we had followed the standard approach in economics and only asked a single open-ended survey question about why respondents are not participating in the stock market? To address this question, we examine how the cumulative probability of having discovered different reasons for stock market non-participation evolves over the course of the interview. When considering only the response to the first open-ended question in our interviews, the distribution of assigned codes skews heavily towards salient reasons for non-participation such as risk (31%), knowledge (50%), and economic constraints (46%). Notably, mental models—such as the active investing model—rarely feature in the initial response, but frequently appear later in the interviews (64%). In other words, interviews do not simply increase the frequency of codes discovered in the initial top-of-mind response.

The systematic differences between factors identified in the initial responses and those identified in subsequent responses raise a natural question about interviews: Are the codes discovered later in the interviews important determinants of choices, or are they simply artifacts of interviewees attempting to fill the conversational space in response to persistent probing? Addressing this question, which applies equally to both traditional and AI-conducted interviews, requires linking interview data with choice data at a scale that exceeds the sample size commonly found in qualitative research. Exploiting the large scale of our initial data collection, we conduct a follow-up study with 266 respondents (70% of our initial sample) eight months after the interviews were conducted. We find that interview codes are highly predictive of incentivized beliefs and choices in the follow-up study. For instance, when measuring incentivized beliefs about the performance of actively managed funds, respondents with an *active investing* mental model believe that active funds are more likely to outperform passive funds after fees. Furthermore, respondents who express an *unwillingness to make nominal losses*—another factor that is absent from the initial top-of-mind response and only emerges later in the interviews—are much more likely than other respondents to require a 100% loss protection on their nominal investment to put money at risk in the stock market. These results demonstrate that codes emerging later in the interviews are highly predictive of choices and should mitigate concerns about “cheap talk” dominating the discourse in interviews.

We contribute to several strands of the literature. Most importantly, by introducing a new approach to conducting qualitative interviews that delegates the task of interviewing human subjects to an AI interviewer, we make a methodological contribution to the interdisciplinary literature on qualitative research methods (Charmaz, 2002; Denzin, 2001; Glaser and Strauss, 2017; Knott et al., 2022; Mishler, 1986; Oakley, 1981; Spradley, 1979). We demonstrate that AI-conducted interviews generate high-quality interview data that can generate important insights into economic decision-making. The scalability of AI interviews enables previously prohibitively costly mixed-method research designs that integrate interviews into large-scale surveys and

randomized controlled trials.

We also contribute to the literature on survey methods in economics (Fuster and Zafar, 2023; Haaland et al., 2023; Stantcheva, 2023) and a rapidly growing body of work that uses open-ended survey questions to collect qualitative data on how people reason about economic problems and measure their subjective beliefs, mental models, and narratives about the world (Haaland et al., 2024). We show a systematic difference between factors identified in the initial top-of-mind response and those uncovered in subsequent responses. Notably, mental models are almost absent from the initial top-of-mind response, demonstrating that interviews can lead to qualitatively different conclusions by making factors visible that are unobserved when relying on an open-ended survey question, rendering this methodological choice important.<sup>2</sup>

Furthermore, we contribute to a small literature in economics that uses qualitative interviews. Bewley (1999) conducts interviews with industry professionals to understand why wages do not fall during recessions. Blinder et al. (1998) conduct interviews with leaders to differentiate between 12 theories of price stickiness. More recently, Bergman et al. (2024) conduct interviews to learn why a randomized controlled trial was effective in making people move to better neighborhood. Bustos et al. (2022) conduct interviews with professional tax planners to learn about novel tax planning strategies. It is also common practice in development economics to conduct interviews to gain knowledge about ground realities before designing interventions, but these are typically only conducted informally (Duflo, 2017).

Lastly, we also relate to substantive literature on the stock market participation puzzle (Barberis et al., 2006; Favilukis, 2013; Gomes et al., 2021; Guiso and Sodini, 2013; Haliassos and Bertaut, 1995; Hurd et al., 2011; Mankiw and Zeldes, 1991). Our main qualitative finding of a prominent role for an *active investing* mental model—according to which stock market participation requires constant monitoring of market conditions and extensive research on individual companies—can help resolve the puzzle and demonstrates that AI-conducted interviews can generate rich insights about economic decision-making.

The paper is organized as follows. Section 2 presents our methodological approach. Section 3 gives an overview of our data collection. Section 4 provides results on interview experiences and how well the behavior of our AI interviewer aligns with its instructions. Section 5 presents our main qualitative results on the stock market participation puzzle. Section 6 investigates the discovery process underlying AI-conducted interviews. Section 7 presents follow-up data on whether interview data predicts incentivized choices. Finally, Section 8 concludes.

---

<sup>2</sup>What is top of mind can be relevant in itself, for example when studying attention allocation (Link et al., 2024).

## 2 AI Interviews: Methodological approach

Qualitative interviews “allow individuals to explain, in their own words, how they understand and interpret the world around them” (Knott et al., 2022) and are “flexible, iterative, and continuous, rather than prepared in advance and locked in stone” (Rubin and Rubin, 2011). In the words of DeLuca et al. (2023), interviews create a “natural, in-depth conversation rather than a series of clinical questions and short answers, facilitating more detailed stories and the emergence of themes unanticipated by the researchers—in sharp contrast to forced choice response survey questions.” In this paper, we focus on *semi-structured interviews*. They are usually organized around a topic guide that includes three to five broad topics that should be covered during the interview. Each topic is introduced with a broad open-ended question, such as “Can you tell me about ... (something specific)?” The interviewer then follows up with a probing question to deepen the conversation or clarify any ambiguities, such as “Can you tell me more about this?” or “What did you mean by that?” When a topic is well covered, the interviewer moves on to the next topic in the topic guide with a new broad open-ended question, taking previous points covered in the interview into account. This progression ensures that all interviews broadly cover the same materials.

### 2.1 Design objectives

The quality of qualitative interviews crucially depends on the ability of the interviewer, which implies several design challenges for an AI interviewer. First, the AI interviewer should adhere to methodological best practices for qualitative research. For example, this includes using broad, open-ended questions instead of narrowly framed questions as well as choosing neutral, non-leading questions that encourage meaningful and detailed responses. Central to qualitative research is the capability for adaptive probing, where the AI interviewer spontaneously adjusts its line of questioning based on the interviewee’s response. This requires the ability to manage the conversation both locally and globally, i.e., remembering previously raised points, identifying recurring themes in the interviewee’s responses for later exploration, and gently steering the conversation towards new topics without impeding the natural flow of the conversation. Second, the AI interviewer must maintain consistent performance, both across interviews as well as over extended conversations. The AI must maintain a high level of fidelity to its instructions, replicating both the explicit and implicit intentions of the researcher. Put simply, the AI interviewer has to stay in character. Third, from a security perspective, the AI interviewer should be robust against attempts of malevolent interviewees to modify their behavior. Finally, given the possibility of discussing sensitive topics, the AI interviewer requires content moderation, ensuring that conversations remain within ethical boundaries. This entails monitoring the interviewee’s responses for signs of discomfort and respecting implicit or explicit requests to stop a specific



line of questioning.<sup>3</sup>

We describe below how we address these design objectives in practice. Our focus is on a portable, general purpose AI interviewer that can be modified by other researchers with minimal modifications, while achieving robust “off-the-shelf” performance across research applications. To facilitate adoption by other researchers, we develop an application that can be seamlessly integrated into commonly used survey design software such as Qualtrics.

## 2.2 Interview interface

Panel A of Figure 1 shows the chat interface that respondents use to interact with the AI interviewer.<sup>4</sup> The chat interface intentionally mimics common text messaging applications to lower technological frictions. The upper part of the interface displays the full conversation history of previous questions and answers, while the lower part consists of a resizable text field that respondents can use to type and revise their responses. When respondents submit a response, it is automatically added to the conversation. The submit button is then deactivated until the AI interviewer provides the next question, which usually takes about five seconds. As in a conversation with a human, a certain minimal delay is desirable to provide respondents with time to relax and reflect between questions. In contrast, immediate responses might even be perceived as unnatural or overwhelm the interviewee. The interface shows a typing animation consisting of three dancing dots while respondents are waiting for the next question to appear, which might lower survey attrition by suggesting to that someone is actively typing on the other end.

## 2.3 AI interviewer

Answering an interview question via the chat interface triggers an API call to a remote server that hosts the code and database for the AI interviewer, which we developed in Python and deploy as a serverless application on standard cloud infrastructure. The API call sends the interviewee’s response to the AI interviewer and awaits a response in the form of the next question.

Our application builds on the Generative Pre-trained Transformer 4 model family (“GPT-4”) by OpenAI that is based on the seminal work by Vaswani et al. (2017). However, our modular approach is compatible with other generative language models.<sup>5</sup> The question generation process

---

<sup>3</sup>In practice, an important metric is the expected response time of the AI interviewer, i.e., the average time to generate the next question. Many of the above design challenges can be addressed to a very high degree by increasing the complexity of the system. However, with current technology, this comes at the cost of a higher expected response time. The researcher thus faces an additional quality–response time trade-off.

<sup>4</sup>Our chat interface can be embedded into any survey software that allows users to include custom HTML and supports basic JavaScript functionalities such as HTTP requests.

<sup>5</sup>Extensive pre-testing revealed that the reasoning ability of GPT-3.5 was insufficient for the purpose of interviewing human subjects, making it important to employ a frontier model such as GPT-4.

Figure 1: The AI interviewer

(a) Chat interface for respondents

I am interested in learning more about why you currently do not own any stocks or stock mutual funds. Can you help me understand the main factors or reasons why you are not participating in the stock market?

I do not own any stocks because I do not know how to get started or what stocks to pick.

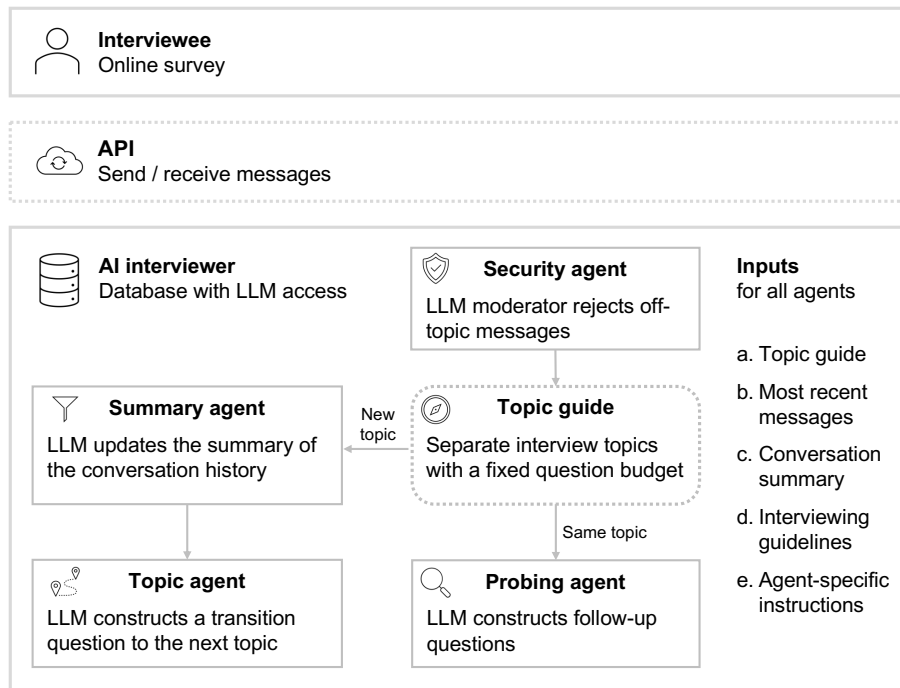
Can you elaborate further on the difficulties you have faced in trying to get started with investing in stocks? What are some of the specific challenges or concerns you have encountered?

Type your message here...

Please write at least two or three full sentences.

Submit response

(b) Workflow to generate an interview question



*Note:* Panel A provides a screenshot of the chat interface used to conduct the interviews during our main survey. Panel B provides an overview of the workflow of how the AI interviewer generates the next interview question. For a detailed description see Section 2.3.

involves four distinct tasks that are carried out by four different “AI agents”: the security agent, the Summary agent, the probing agent, and the topic agent. Each agent is a separate instance of a large language model that generates a text output from text input based on a unique set of written instructions (“prompts”). Together, they collectively function as our *AI interviewer*. The workflow for the question generation process is illustrated in Panel B of Figure 1 and described in detail below.

**Security agent** The first step in the question-generation process is a defense layer against clear attempts to change the behavior of the AI interviewer (e.g., “tell me a joke”). This layer consists of an LLM that determines whether the answer “fits into the context of an interview” by comparing it to the previous question (see Appendix Section C.3.4 for the full prompts). Messages by the interviewee that do not fit the context of the questions are flagged by the agent. If an answer is flagged, the interviewee receives a message that with a gentle nudge to either rephrase the answer or decline to answer the question. In practice, less than 1% of messages are flagged, suggesting that most respondents behaved cooperatively.

**Topic guide** The second stage in the process is a binary decision on whether to continue with additional probing questions or to transition to a new interview topic from topic guide. The topic guide is determined by the researcher in advance and, in our case, specifies a broad objective for each interview topic (e.g. “explore perceived barriers to stock market participation”) and a fixed number of probing questions for each topic. Once this question budget for the current topic is exhausted, the AI interviewer transitions to the next interview topic. We discuss extensions that endogenize the topic transition process in Appendix Section C.1. The AI interviewer concludes the interview with an expression of gratitude once all topics are covered.

**Probing agent** The aim of a semi-structured qualitative interview is to “achieve both breadth of coverage across key issues, and depth of coverage within each” (Legard et al., 2003). In our setup, this responsibility falls on the probing agent. To formulate an appropriate probing question, the probing agent receives a summary of the previous conversation history, the current topic of the topic guide, and the conversation history within the current topic in the topic guide. We give the probing agent both general guidelines as well as specific probing guidelines to make sure that the probing questions align with best practices. Specifically, we emphasize that questions should be asked in an open-ended way (“how”, “what,” “why”) to allow detailed responses that cannot be answered by a simple “yes” or “no.” We also emphasize the need for neutrality, respect, and relevance. For the specific probing guidelines, we emphasize the need to follow-up on promising themes that align with the topic guide, exploring the interviewee’s reasons, motivations, opinions, and beliefs. We also instruct the probing agent to clarify ambiguous answers and to pivot to new areas not covered in depth if the conversation becomes repetitive or remains on the surface level.

**Summary agent** The summary agent is responsible for reviewing the conversation and creating an appropriate summary that can be passed on to the question-asking agents. While in theory the probing agent and the topic agent could have reviewed the whole conversation before formulating a question, it turned out to be more reliable to work with a continuously updated summary of the key points that emerged earlier in the interview. Whenever the interview moves onto a new interview topic, the summary agent receives the topic guide, the conversation summary from previous topics covered in the interview guide (if any), the current topic of the topic guide, and the full *verbatim* messages that are part of the current interview topic. The summary agent then updates the conversation summary which is later accessible to both the topic agent and the probing agent.

**Topic agent** The topic agent is responsible for introducing the next topic in the topic guide. We emphasize that questions should be open-ended and that the transition to a new topic should feel smooth and natural, taking the previous conversation history into account to bridge what has been discussed with what will be covered next.

## 2.4 Extensions

We focus on a minimal implementation of an AI interviewer for two reasons. On the one hand, a minimal design is portable, allowing other researchers to modify and possibly extend the AI interviewer according to their research objectives. On the other hand, raising the complexity by increasing the number of LLM queries linearly increases the waiting time for respondents, which will eventually impede the conversation’s flow. As technology progresses and the reasoning abilities of LLMs increase, we expect these constraints to be relaxed eventually. In this case, several natural extensions emerge, such as (i) allowing the model to determine the order and length of interview topics based on how the current interview unfolds, (ii) using multi-modal models to allow respondents to talk to the AI interviewer instead of having to write, (iii) replacing the probing agent with a collaborative ensemble of AI agents, and (iv) using additional model fine-tuning on interview transcripts. We discuss these possible extensions in Appendix Section C.1.

## 3 Application: Stock market participation puzzle

We demonstrate the potential scientific value of our method of conducting qualitative interviews with an AI interviewer in the context of the *stock market participation puzzle* (Gomes et al., 2021; Guiso and Sodini, 2013; Haliassos and Bertaut, 1995). This puzzle is a topic of great interest to economists and policymakers, and one of the key open questions in household finance (Guiso and Sodini, 2013). Furthermore, it is difficult to resolve the stock market participation

puzzle with choice data because many of the potential key factors—such as pessimistic beliefs, non-standard preferences, and perceived participation costs—are usually unobserved in choice data. The likely multifaceted barriers thus require an empirical method such as the interview that is able to obtain rich data on the drivers of stock market non-participation.

### 3.1 Interview sample

We conducted 395 interviews between August 23 and September 5, 2023. We excluded 14 interviews that were flagged as inappropriate during a human review of all interview transcripts, leaving us with 381 interviews for the final analysis.<sup>6</sup> Respondents were recruited from the research platform *Prolific*, a survey platform commonly used in economic research that is associated with high data quality (Peer et al., 2022). We employ a screener survey to focus recruitment on the relevant population of US households that are not participating in the stock market.<sup>7</sup> Eligible respondents receive an invitation to participate in the main interview study.

To be eligible, respondents cannot own any individual stocks or stock mutual funds, neither directly nor indirectly through any retirement accounts, following the convention in the literature on stock market non-participation (Gomes et al., 2021, p. 929). We further exclude respondents who plan to buy stocks in the near future and respondents that currently delegate financial decision-making to a financial advisor. Finally, to focus on respondents that possess the means to invest in the stock market, we limit recruitment to households with an annual income of at least \$30,000 and non-zero savings.

Interviews require higher levels of attention, effort and engagement compared to traditional surveys (Rubin and Rubin, 2011). To ensure a high quality of responses, we employ three complementary strategies. First, our screener survey includes an attention check designed to screen out inattentive respondents. Second, in the screener survey, we screen out respondents who demonstrate an unwillingness to engage with open-ended questions by asking them to write about their views on daylight savings time. Third, we saliently communicate that respondents will participate in a chat-based interview with an AI at the beginning of our survey, allowing respondents who may not feel comfortable in an interview situation to select out of the study.<sup>8</sup>

As shown in Table A.2, our final interview sample is diverse and represents many important socioeconomic subgroups of the population. 62.3% of our respondents are female. 76.2% and 12.4% are white and black, respectively. The median household income is \$73,038. As is common with online samples, respondents are more educated than the general US population, with 81.7% holding a college degree.

---

<sup>6</sup>Most of these 14 interviews were omitted because it was obvious that the respondents used ChatGPT or similar tools to answer the interview questions.

<sup>7</sup>Appendix Section E.1 presents the questionnaire used for the screener survey.

<sup>8</sup>We revisit this selection decision in Section 4.2, showing that the willingness to participate in AI-conducted interviews is at least as high as for human-conducted interviews.

## 3.2 Main survey

Interviews were conducted as part of an online survey. We first describe the core elements of the survey and then describe the interview in Section 3.3. Section E of the Online Appendix provides the full questionnaire. The median time to complete the survey, including the interview, was 40 minutes, for which respondents received a fixed compensation of \$12. The survey starts with a detailed consent form informing respondents that we will use AI tools to process responses using the OpenAI API. We inform respondents that their data will not be used for model training and that OpenAI will permanently delete their data after 30 days.

After eliciting basic sociodemographic information, we explain to respondents that they will participate in an interview with an AI chatbot on the topic of stock market participation. We tell respondents that the AI chatbot is informed that they are currently not participating in the stock market, and that we expect that the interview will take about 20 minutes to complete. We then allow respondents to familiarize themselves with the chat interface. Next, we explain to respondents that it is important to provide detailed responses. We also emphasize that there are no right or wrong answers, and that they should only answer the questions to the extent that they are comfortable.

The final part of our survey consists of a set of standard closed-ended questions. First, we elicit the main reasons for stock market non-participation using a multiple-choice question. This allows us to contrast a closed-ended survey question with the qualitative interview format. Second, we collect data on respondents perceived interview experience, including an evaluation of the overall experience, the naturalness of the conversation, and their preference for AI vs. human interviewers on 7-point scales. Third, we elicit factors commonly associated with stock market participation, such as financial literacy, risk attitudes, and return expectations. The survey concludes with a battery of questions on respondents' household finances and their sociodemographic characteristics.

## 3.3 Interviews

We conduct semi-structured interviews that combine open-ended questions with targeted follow-up questions in a way that ensures that all interviews cover a comparable set of topics. This approach allows respondents to express themselves using their own words without restricting the set of permissible answers, allowing unanticipated themes to emerge that might otherwise be invisible in closed-ended survey questions.

All interviews consist of 18 questions with the first and the last two questions being predetermined. The AI interviewer always begins the conversation with the following broad question that signals interest in their personal reasons for non-participation:

“I am interested in learning more about why you currently do not own any stocks or

stock mutual funds. Can you help me understand the main factors or reasons why you are not participating in the stock market?”

While qualitative interviews traditionally start with a different opening question to build rapport, we deliberately chose this question because it resembles the type of open-ended questions that are currently used in surveys to collect qualitative data. This allows us to later contrast the patterns in people’s responses to such an open-ended survey question with the patterns that a researcher would identify based on a full interview.

After the first question, the AI interviewer explores the interviewee’s stated reason for non-participation with five probing questions. For the second topic, the AI interviewer is instructed to “delve into the perceived barriers or challenges preventing [the respondent] from participating in the stock market” with a total of five questions. We chose this topic to ensure that the interview also covers other reasons for non-participation. The advantage of this topic is that it naturally “reorients” the AI interviewer if a conversation is stuck in an unproductive line of inquiry. The third topic “explore(s) a ‘what if’ scenario where the interviewee invests in the stock market”, asking respondents what they would do and what it would take to successfully navigate the stock market (three questions). Counterfactual reasoning is a commonly used strategy in qualitative research that allows respondents to explain themselves more freely. The fourth and final topic is about “conditions or changes needed for the interviewee to consider investing in the stock market” (two questions). After this topic, we inform respondents that the interview is about to conclude.

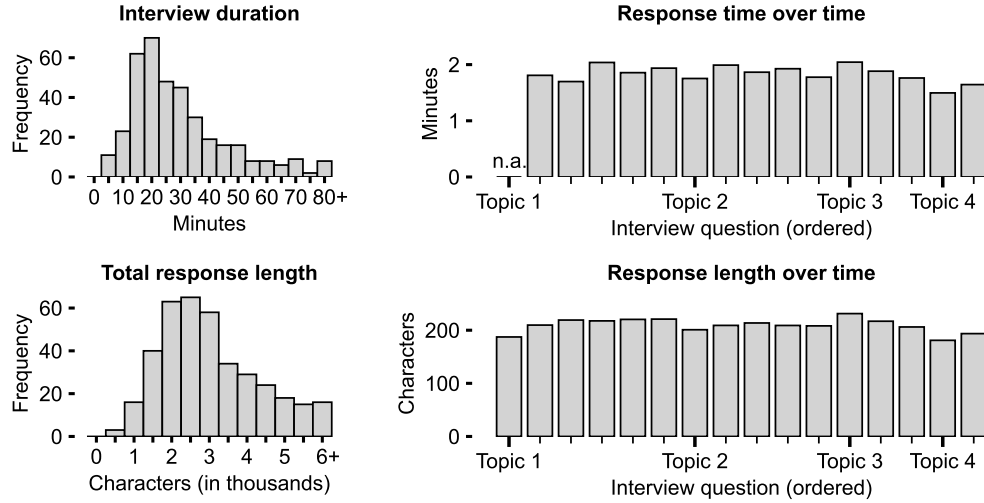
The AI always concludes the interview with the same two pre-determined questions recommended as best practices (Rubin and Rubin, 2011). Specifically, the AI asks whether there are “any perspectives or information you feel we haven’t addressed that you’d like to share” followed by a prompt to reflect on the conversation and again identify the main reason for non-participation. These questions provide an insurance against omitting an aspect highly relevant to people’s non-participation decisions. The interviewer then thanks the respondents for their participation and asks them to proceed to the next page to continue with the survey.

## **4 Interview experiences and interviewer behavior**

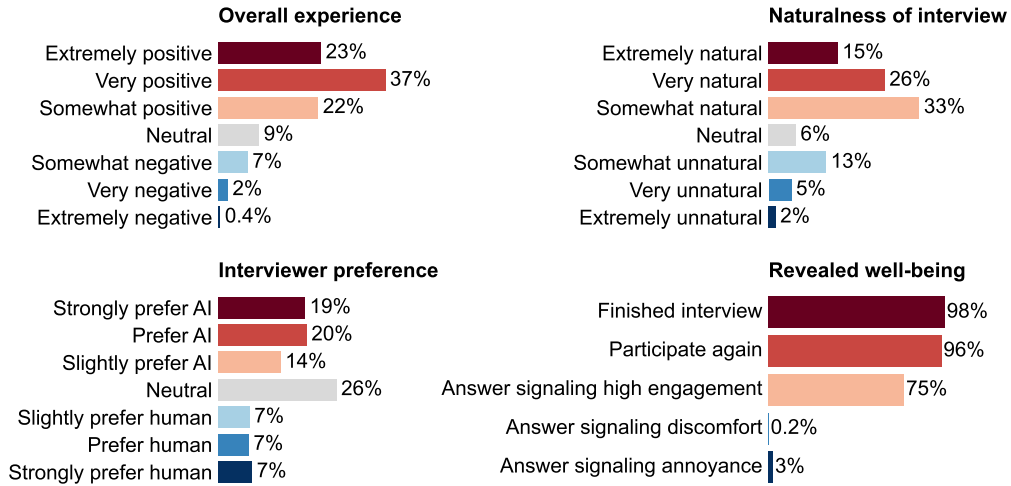
Before turning to the main findings from the thematic analysis of the interview transcripts in Section 5, we first demonstrate that interviews worked well in practice, focusing on respondent engagement, welfare, and the alignment of the AI interviewer with our interviewing guidelines.

Figure 2: Interview effort and experience

### A. Effort and engagement



### B. Interview experience



**Note:** Panel A presents the distribution of the interview duration (in 5-minute bins), the average response time for the  $k^{\text{th}}$  interview question, the distribution of the length of the total text output (in thousand characters), and the average response length (in characters) to the  $k^{\text{th}}$  interview question. The beginning of a new interview topic is indicated by “Topic  $k$ ”. Response times for the first question could not be collected. Panel B shows the distribution of responses to the 7-point Likert scale questions about the overall interview experience, the naturalness of the conversation, and the respondent’s preference over conducting interviews with an AI chatbot vs. an actual human interviewer. In addition, the figure reports the share of respondents who finished the full interview, indicate that they would like to participate in an AI interview again, and the share of answers in the interview that have been coded as displaying high engagement, discomfort or annoyance by a team of trained research assistants.



## 4.1 Respondent engagement

Our evidence suggests that the AI interviewer manages to engage respondents and encourage high levels of effort during the interview. First, 97% of respondents completed the full interview, although participation in the interview was voluntary and respondents knew payments were not conditional on participation. In fact, respondents could decide to proceed in the survey and skip the interview after a five-minute delay. The very low attrition rate is encouraging given frequent concerns that survey respondents might not exert effort without meaningful performance incentives, especially for open-ended questions (Haaland et al., 2023). Second, we use interview time and response length as an indicator of effort. Interviews take about 33 minutes on average and respondents produce 654 words (3,487 characters) of text during this time. With 29 words per minute, the amount of text produced in our AI interviews is slightly above the benchmark of about 20 words per minute obtained in chat-based interviews with human interviewers (Namey et al., 2020). Third, effort does not decline over the course of the interview. Respondents consistently take about two minutes to answer interview questions in about 200 characters across all four interview topics (Panel A of Figure 2).<sup>9</sup> We supplement these behavioral measures of effort with subjective assessments of respondent engagement. A team of trained research assistants systematically hand-coded all 6,582 interview responses on a 3-point scale as either displaying no, low, or high engagement, focusing on whether respondents make an effort to participate in the conversation rather than the amount of text produced (see Section D of the Online Appendix for details on the hand-coding). Overall, 75% of responses are rated as displaying high engagement and only 2% of responses are rated as not engaging in the conversation at all (Panel B of Figure 2).

## 4.2 Respondent welfare

A key task during interviews is to monitor and maintain respondents' well-being and safety, which was an explicit part of the AI interviewer's instructions. To examine whether the AI interviewer achieved this objective, we trained research assistants to systematically hand-code responses for any signals of discomfort or annoyance (see Section D for details on the hand-coding). We find that less than 0.2% of responses exhibit any signs of discomfort and only 170 responses (2.6%) indicate some mild form of annoyance, which usually arises because respondents feel like they answered a question already.

This evidence is consistent with respondents' evaluation of the interview in our post-interview survey module where 95.4% of respondents indicate that they would like to participate in another AI-conducted interview. Panel B of Figure 2 shows that 91% evaluate the overall interview experience as positive or neutral, with 60% rating the experience as "very" or "extremely positive." Similarly, 80% of respondents rated the conversation with the AI interviewer as natural. When

---

<sup>9</sup>Text output and response times are uncorrelated with past use of AI tools such as ChatGPT (see Figure B.2).

asked about their preference for an AI or human interviewer, 53% would prefer an AI interviewer compared to 21% of respondents who would prefer to engage with a human interviewer.

To mitigate potential concerns that the above patterns are driven by demand effects of selection, we conduct a complementary correspondence study with about 500 Prolific respondents in November 2023.<sup>10</sup> After replicating the set of screener questions from our main study, we inform respondents that they might be eligible to participate in three additional studies, which we present on separate survey pages in randomized order. For each study, respondents indicate whether they would like to be invited to the study if we decide to run it. Each study is presented exactly as we advertised our main experiment on Prolific. However, we exogenously vary whether a study is described as (i) “a 40-minute survey”, (ii) “a 40-minute study with a text-based interview with a human” or (iii) “a 40-minute study with a text-based interview with an AI”. In a between-subject comparison based on the first vignette (Figure B.3), we find virtually identical demand for participating in a survey without an interview (89%) and a survey with an AI interview (90%). If anything, respondents are 6.6 percentage points *less likely* to select into a human-led compared to an AI-conducted interview ( $p = 0.078$ ). Moreover, exploiting the within-subject variation, we find no evidence of differential selection into surveys compared to AI interviews along sociodemographic groups (as shown in Table A.7). Taken together, this evidence suggests that the positive evaluations of AI interviews are not driven by selection.

### 4.3 AI alignment

Finally, we examine the AI interviewer’s alignment with its instructions, focusing on whether interview questions comply with the explicit requirements set out in our prompts. For this exercise, a team of trained research assistants systematically hand-coded over 12,000 interview questions and responses, assessing whether (i) questions are open-ended, (ii) non-leading, (iii) non-redundant, (iv) relevant to the objective of the interview, (v) the AI interviewer correctly refers back to previous parts of the conversation, (vi) asks a clarification question, and (vii) a response requires a clarification question (see Section D for details on the hand-coding).

First, we find that the AI interviewer follows its instructions to ask open-ended and non-leading questions: 95% of questions are coded as open-ended and non-leading. In addition, most questions are deemed to be non-redundant (83%) and relevant to the interview’s research objective (95%). Second, 76% of all questions build directly on the previous response, while another 17% of questions refer to what was said in earlier parts of the interview. Reassuringly, only 0.01% of questions include any *incorrect* references to earlier responses, suggesting that “hallucination” (Zhang et al., 2023) is not an empirically relevant concern for AI-conducted interviews. Third, we find that the AI interviewer follows its instructions to ask clarification questions if further information is required to disambiguate an earlier response. Figure B.1 shows

---

<sup>10</sup>Section E.3 contains the key experimental instructions. Table A.5 provides summary statistics for the sample.

that the probability that the next question in the interview is a clarification question sharply increases if a response is coded as requiring further clarification. Specifically, clarification questions are almost absent (5.8%) if responses are coded as not needing further clarification, but the propensity of asking a clarification question next increases to 35.2% and 50.4%, respectively, when the importance of asking a clarification question is coded as “somewhat relevant” and “very relevant.” Taken together, this suggests that the AI interviewer is capable of following instructions and adhering to interviewing guidelines.

## 5 Qualitative analysis of the interview data

### 5.1 Method for developing a coding manual and analyzing the data

To analyze qualitative interview data, researchers usually read the entire corpus of interview transcripts and make notes of salient themes. Our AI interviews generated over half a million words of text, making it inefficient to read all interviews from start to end. We, therefore, follow an approach that complements a detailed review of a random subset of interviews with techniques from computational sociology to sort question–answer pairs from the full text corpus into distinct semantic clusters (Nelson, 2020), thus facilitating an efficient and replicable discovery of recurring patterns. When developing the manual for the subsequent systematic coding of interview transcripts, we rely on an inductive thematic analysis of the data (Knott et al., 2022), ensuring that discovered themes are grounded in the interview data.

#### 5.1.1 Coding manual

We start by describing the process we used to develop the coding manual, which involved four steps: (i) initial immersion in a random subset of interviews, (ii) machine-assisted thematic clustering of responses, (iii) systematic review of thematic clusters, and (iv) convergence on a final set of codes. While we describe this process as linear, it involved an iterative back-and-forth in practice. A core advantage of this approach is the machine-assistance in step (ii), which allows us to efficiently use all text data without increasing the analytical cost as the size of the text corpus increases. This step thus addresses the challenge of scale without having to cede the interpretative authority to a machine (Nielsen, 2020).

**Step 1: Initial immersion** To get an initial impression of the interview dynamics, we independently read a random sample of 50 interview transcripts. This provides a basic understanding of the language that people use to talk about their stock market non-participation. It also provides a context for interpreting responses from the thematic cluster identified in the next step.

**Step 2: Thematic clustering** To increase the chance of discovering explicit and implicit themes, we re-organize the data into thematic clusters using unsupervised machine learning. We construct a text corpus consisting of 13,993 question–answer pairs from our 381 interviews. Each element of the text corpus thus consists of a question by the AI interviewer and the corresponding response from the interviewee. We do not perform any additional pre-processing steps. We then use BERTopic to sort question–answer pairs into clusters that share a common theme (Grootendorst, 2022).<sup>11</sup> This process assigned 63% of the data to 54 endogenously identified thematic clusters. This approach has several advantages compared to conventional topic modeling algorithms. First, the text embeddings capture semantic patterns that would be lost in bag-of-words representations commonly used in other topic models (Blei et al., 2003) or dictionary-based methods. Second, it reduces the researcher degrees of freedom, as researchers do not have to take a stance on which features to extract or how to pre-process the text data (Gentzkow et al., 2019). Third, the clustering algorithm is noise-aware, rendering resulting clusters semantically more coherent. Indeed, low interpretability is a known problem of other topic models (Ralund et al., 2022).

**Step 3: Review of clusters** We review the thematic clusters identified in the previous step one at a time. For each cluster, we retrieve all question–answer pairs that the model assigns a probability of at least 50% of belonging to that cluster. We independently read the question–answer pairs in order descending in the model’s score. We then write a short synopsis of each topic and compile a list of potential codes and themes along with example responses. At this stage, we made care to only construct codes that were “grounded” in the data and reflect what people said in the interview, rather than to construct codes that reflected our prior understanding of the subject.

**Step 4: Convergence** After reviewing all clusters and making notes of potential codes, we had multiple meetings in which we consolidated our findings. We included themes in the final coding scheme that we both identified and on which a shared conceptual understanding was reached.

### 5.1.2 Hand-coding of interview transcripts

Three experienced research assistants were given the task of coding the interviews according to the coding scheme. Two of the research assistants had previously read the interview transcripts during the first round of coding described in Section 4, giving them a high context understanding of the explicit and implicit nature of the codes. Before starting on the task, they received a

---

<sup>11</sup>Specifically, we use the `all-MiniLM-L6-v2` sentence transformer model to construct a document embedding. We next reduce the dimensionality of the document embedding by applying uniform manifold approximation and projection (UMAP, see McInnes et al., 2020). In a final step, a hierarchical density-based clustering algorithm assigns documents to topics (HDBSCAN, see Campello et al., 2013).

detailed coding manual that, in addition to general best practice recommendations for the coding of qualitative interview transcripts, included actual example responses for each code included in the scheme. The research assistants were also told that creating a coding scheme is an iterative process and were encouraged to suggest potential changes to the coding scheme based on their readings of the interviews. We conducted multiple training sessions involving triple-coding of transcripts to ensure a shared understanding of all codes. Appendix Section D provides additional details.

## **5.2 Qualitative evidence on barriers to stock market participation**

The interviews uncover considerable heterogeneity in the perceived barriers to stock market participation, suggesting that non-participation is driven by multiple factors rather than a single explanation. Many respondents provide accounts that consist of sets of interrelated concerns, such as belief that stock market participation is associated with high costs of participation due to the need for ongoing research and monitoring, the necessity of acquiring substantial expert knowledge, and the uncertainty of achieving a return on these investments. These concerns are rooted in an underlying view of stock market participation as an active and demanding endeavor, where success hinges on continuous effort and expertise without any guarantee of a positive outcome. As the focus of our paper is mainly methodological, we focus on a subset of main themes in our presentation of results and point out their interconnections before concluding this section with an overview of additional themes and barriers that validate previous research on the stock market participation puzzle.

### **5.2.1 Mechanism 1: Participation costs and the mental model of active investing**

When asked about their reasons for not participating in the stock market, a recurring theme is that respondents describe concerns about the time commitments and the work involved in owning stocks. As one respondent puts it:

*Managing investments in the stock market feels like it could almost be a part-time job given the level of attention it demands. My current job is already time-consuming and mentally exhausting. Adding the task of researching companies, keeping up with earnings reports, and tracking market trends would be another layer of complexity and stress.*

Many respondents perceive high recurring participation costs from owning stocks, due to the need for (i) research and (ii) monitoring. First, they believe that research is an essential task of an investor: “I know you are supposed to do research on stocks that you would like to invest in, for example if there was a recent oil spill, it would not be good time to buy stock in oil companies.” This research has to integrate a broad array of public information, such as historical price data, financial accounts, earnings announcements, and analyst forecasts. As a result, respondents

understand that it “it takes time to research and understand a company’s product and how it fits into the marketplace” and anticipate “spending a couple of hours a day or maybe several hours 3-4 days a week online reading about market trends, emerging technologies or medicines, and companies on the uprise as well as those in decline.”

Yet, the perceived job requirements of an active investor continue beyond the point when the initial research is concluded and promising stocks have been identified, because one would “have to constantly monitor the current market condition and make sure that the stock [you] invested in would not suddenly crash.” Continuous monitoring is important because stock prices change “all the time, every day” and “if you sell [stocks] at a bad time you wont get as much for the stock than if you sold them at a good time... and waiting for the peak time to sell seems best.” Many respondents thus seem to believe that making a return on one’s investments requires careful timing of the stock market, which only adds to the perceived costs of participating in the stock market.

Collectively, these perceptions characterize what we refer to as the *active investing* model of stock market participation. This model is grounded in three fundamental beliefs. First, respondents believe that market inefficiencies allow investors to predict future stock prices through research and analysis. Second, they perceive stock prices as highly volatile, making continuous monitoring and quick decision-making important to avoid occurring large losses before prices adjust. Third, there is a shared view that investors make returns from short-term trades. These beliefs contribute to a view that participating in the stock market is a demanding, high-maintenance activity, which ultimately deters many respondents from investing.

### **5.2.2 Mechanism 2: Knowledge barriers**

Another prevalent theme among respondents is a concern that they lack fundamental knowledge about the stock market. Many respondents expressed uncertainty about the basic logistics of investing, such as how to get started, how to buy and sell stocks, or create an investment account. As one respondent stated: “I don’t know where to view the stock market, how people make money from stocks, and/or how they work. I don’t know how to invest, or what to start with when investing.” This perceived lack of knowledge extends beyond logistical aspects and into the complexity of the stock market itself:

*To be honest, all of it sounds complicated. I barely know what any of the terms in the market mean. Like, I’m not familiar with dividends, mutual funds, etc. Aside from not having knowledge of the terms, I also don’t know what makes something a good investment vs a bad one. I wouldn’t even know where to start.*

Many respondents point to the confusing jargon and technical concepts used in finance. This perceived need to grasp many complicated concepts before even beginning to invest leads many respondents to conclude that acquiring the necessary knowledge would be a time-consuming

endeavor with uncertain rewards. As one respondent puts it, “it seems like there is a lot to learn, and that is a little intimidating.” Moreover, respondents voiced a deeper concern in the form of a belief that successful investing requires specialized knowledge and skills that they do not possess. Specifically, many respondents explain that successful investing requires an ability to predict stock prices, distinguish “good” from “bad” stocks, and effectively time stock trades ahead of everyone else. As one respondent explains:

*I feel like I don't know which stocks are 'good' to purchase, or when a good time to purchase would be. Because of this, I feel like I would be flying blind and making decisions entirely at random, as compared to a more seasoned stock market veteran.*

This perceived need for experience and expert knowledge both underscores the cost of stock market participation and reinforces the conclusion that without such knowledge, one is likely to incur losses. Another respondent articulated a similar concern, saying that they are “not quite sure how to read stocks and what patterns to look for,” and that “for this reason, I feel like if I were to try and invest in stocks without this knowledge I might lose money.” This combination of the perceived logistical complexity of investing, a belief in the necessity of extensive knowledge and experience, and the fear of losing money precisely because they lack those skills contribute to a perception of substantial entry barriers to investing.

### **5.2.3 Mechanism 3: Risk perceptions**

A common theme that emerges when discussing the stock market are strong concerns about risks. Many respondents express a deep-seated fear of losing a substantial part, or even all, of their “hard-earned” money on the stock market. Unlike risk-free assets where returns are guaranteed, respondents worry that they could end up with less money than they started with:

*If I would lose any amount of money, no matter how large or small, it would make me sick to think about. I work so hard for everything I have, that knowing it could all be gone with the stock market failing would make me incredibly anxious.*

The concern that an investment could “instantly lose all its value” reflects the belief that such catastrophic losses are not only possible, but likely, and could occur at any moment. The fear of catastrophic events leads many to conclude that the risks of investing are too large, and that owning stocks would therefore expose them to constant feelings of stress and anxiety because “the market changes everyday.”

The metaphor of gambling frequently surfaces in respondents’ accounts of the risks and returns of investing. This comparison is not merely metaphorical but reflects the mental model that the stock market operates with the same uncertainty and lack of control as a game of chance. For these respondents, the stock market is “like a lottery, where you can either win big or lose everything” and it is “all a roulette of which stock will fail and which will win.” Investing in

stocks is seen as being as risky and unpredictable as gambling, where the odds are perceived to be stacked against the average investor, supporting the view that the stock market is something for “when you have discretionary income to burn...not something you can rely on to build wealth.”

It is important to note that not all respondents see the stock market purely as a gamble. There is also a belief among respondents that the risks can be mitigated through knowledge and experience. As one respondent noted, investing in the stock market is risky “unless you know what you are doing.” Another respondent, who express a common concern about going “completely bankrupt,” puts it like this:

*One of my biggest fears is going completely bankrupt since my stocks may fail. I also fear I may not know anything about the company I invest in and something huge happens where they close. I want to feel better about my experience and have that knowledge and funds.*

These ideas tie back to previous themes, where respondents expressed that acquiring the necessary knowledge and skills could transform the stock market from a high-stakes gamble into a more manageable risk, implying that a seasoned investor faces a different risk-return trade-off. Yet, this knowledge comes at a considerable cost both in terms of time spent and, more importantly, the financial losses incurred during the costly learning process, which many feel they could not afford. As one respondent summarizes: “You could put money to buy stocks that end up losing value rather than gaining value and end up in a situation where you lost both time and money for participating.” This explains why many respondents prefer risk-free assets where neither time nor expertise are required to receive a guaranteed return on their investments.

#### **5.2.4 Additional themes**

**Economic constraints** Despite recruiting respondents with a minimum household income of \$30,000 and a median financial assets of \$18,000 (Table A.2), many interviewees start the interview by saying that they do not have the financial means to invest in the stock market. This belief is partly driven by misconceptions regarding the necessary minimum investment amount. Several respondents assume that substantial upfront savings are required, either because they believe institutional rules dictate high entry thresholds for opening investment accounts or because they perceive that “good” stocks—those with low risk and high return—are prohibitively expensive, costing thousands of dollars per share: “...the most successful stocks have high buy prices and I think an onlooker such as myself would be enticed, yet unable to afford such luxuries.” Other respondents argue that stock market participation requires the financial capacity to absorb losses and stay invested until returns are realized, which aligns with the broader mental model of the stock market as a casino: just as a gambler needs sufficient funds to endure a losing streak, respondents think that only those with substantial savings can navigate the market’s volatility. For them, this suggests that the stock market is a domain reserved for the wealthy, who can afford to remain in the “game” long enough to see a return.



**Experiences and social influences** The perception of the stock market as excessively volatile is often shaped by respondents' direct or indirect experiences with investing. Many respondents point to the negative experiences of friends, family members, or coworkers who lost large amounts of money in the stock market, especially during the 2007–2008 financial crisis. This crisis looms large in respondents' memories. For many, the market is synonymous with a series of unpredictable crises and market crashes, which is essentially the only narrative about the market they recall from the news. Additionally, some respondents cite social influences rooted in their growing up in an environment where stock ownership was uncommon or viewed with skepticism. Some describe coming from families that emphasized financial caution, where stock market participation was discouraged as a reckless and risky activity. This socialization then results in identity-based arguments that the stock market is not for people “like myself” (Henkel and Zimpelmann, 2023).

### 5.3 Validating the active investing mental model with a large-scale survey

A key finding from our qualitative analysis is that many respondents subscribe to an *active investing* mental model that makes stock market participation seem like a very substantial time commitment. If this mental model indeed acts as a barrier to stock market participation, it should be more prevalent among non-owners than among stock owners. To test this hypothesis, we conduct a large-scale survey in which we measure the prevalence of this mental model with tailored closed-end survey questions. Another key benefit of this exercise is that it also provides a way to address potential concerns about “overfitting” whereby spurious pattern in the interview corpus, perhaps drive by anecdotal stories that are not representative of the broader population, lead to chance discoveries that do not generalize.

**Design** We recruited a balanced sample of stockowners and non-owners in May 2024 from Prolific in the US. 999 respondents passed the attention check and completed the survey. Table A.4 provides summary statistics and Appendix Section E.4 contains the key experimental instructions. We elicit whether respondents currently participate in the stock market and provide a definition of the term. We then elicit people's agreement with the following statement on a 5-point Likert scale from “Strongly disagree” to “Strongly agree”:

If you want to make any money in the stock market, you have to constantly monitor market conditions, keep up-to-date with all the news, and do a lot of research on the stocks that you plan to invest in.

We use agreement with this statement as a direct indicator for whether respondents subscribe to the *active investing* model of the stock market. In addition to this broad summary measure, we directly elicit other aspects of the *active investing* mental model that we derived from the

above qualitative analysis. Specifically, we elicit respondents' beliefs about how often it would be "necessary to monitor market conditions and check one's portfolio" on a 7-point scale from "less than once a year" to "multiple times a day," their belief about the time per week required to manage one's portfolio, whether it is important to react quickly to daily changes in market conditions and whether it is important to "actively pick a few stocks." We also elicit the set of types of public information that respondents believe one would have to "regularly consult [...] to make a return on one's portfolio," such as earnings reports, company press releases, or social media. We deliberately chose comparatively extreme formulations for the questions that emphasize that the described behaviors are of essential nature for successful stock market participation to lower the chance of false positive agreement. We therefore interpret our measures as a likely lower bound on the prevalence of the *active investing* mental model.

Table 1: Validation survey: active investing and non-participation

|                | Dependent variable:                |                                   |  |   |  |   |
|----------------|------------------------------------|-----------------------------------|--|---|--|---|
|                | (1)<br>Active trading mental model | (2)<br>Daily monitoring important | (3)<br>Quick reactions to market important | (4)<br>Actively picking winner stocks important | (5)<br>Public info required to read (z-scored) | (6)<br>Log weekly portfolio management time |
| Non-owner      | 0.218***<br>(0.029)                | 0.132***<br>(0.027)               | 0.173***<br>(0.031)                        | 0.207***<br>(0.030)                             | 0.368***<br>(0.062)                            | 0.411***<br>(0.073)                         |
| Constant       | 0.551***<br>(0.022)                | 0.192***<br>(0.018)               | 0.325***<br>(0.021)                        | 0.495***<br>(0.022)                             | -0.182***<br>(0.043)                           | 0.575***<br>(0.053)                         |
| N              | 1,000                              | 999                               | 999  | 999   | 999  | 999   |
| R <sup>2</sup> | 0.053                              | 0.023                             | 0.031                                      | 0.045   | 0.034  | 0.031                                       |

*Note:* This table displays regression estimates using data from the validation survey conducted in May 2024 with an equal share of stockowners and non-owners. The dependent variables in columns 1–4 are binary variables. "Non-owner" is a dummy for respondents without any exposure to the stock market. "active investing mental model" is a dummy for agreement with the active investing view. "Daily monitoring important" is a dummy for respondents selecting "daily" or "multiple times a day" when asked about how often it is necessary to monitor market conditions and check one's portfolio. "Quick reactions to market important" is a dummy for selecting "the key to stock market success is to react quickly to daily market changes." "Actively picking winner stocks important" is a dummy for selecting "the key to stock market success is to actively pick a few stocks that will increase in value and avoid stocks that will decrease in value." "Public info required to read" is a z-scored number of public information sources that respondents think one has to "regularly consult when participating in the stock market to make a return on one's portfolio." "Log weekly portfolio management time" is the log of the weekly hours that respondents think one would have to spend managing one's portfolio, measured on an interval scale from "less than 1 hour per week" to "more than 15 hours per week." Robust standard errors are shown in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

**Results** Column 1 of Table 1 shows that 76.9% of non-owners agree with the *active investing* mental model compared to only 55.1% of respondents who participate in the stock market, a difference of 39.6% ( $p < 0.001$ ). In addition, respondents who do not participate in the stock market are 68.8% more likely to believe that one has to monitor one's portfolio at least once a

day (Column 2), 53.2% more likely to believe that quick reactions to daily changes in market conditions are necessary, and 41.8% more likely to believe that selecting a few stocks that are likely to perform well is the key to successful stock market participation (Column 4). Moreover, non-owners select 36.8% of a standard deviation more types of public information as being required reading when owning stocks. As a result, non-owners expect to spend 41.1% more time per week managing their portfolio compared to respondents who own stocks. Panel A of Appendix Table A.6 shows that these results are quantitatively robust to including controls for a broad set of sociodemographic characteristics and respondents' financial literacy as measured by the "Big 3" financial literacy inventory. Panel B of Appendix Table A.6 further shows that our broad summary measure of the active investing model is strongly correlated with subfacets of the mental model.

Taken together, the evidence from our large-scale survey confirms that owners and non-owners have systematically different mental models around the need for active investing in the stock market. This confirms that our key qualitative findings from the interviews can be quantitatively important in explaining stock market participation choices—providing a proof of concept that AI-conducted interviews can discover relevant factors of high relevance for economic choices. More broadly, this exercise demonstrates how AI-conducted interviews can be combined with quantitative surveys in mixed-methods projects to validate the findings from the interviews.

## **6 The discovery process in qualitative interviews**

The main selling point of qualitative interviews is their ability to uncover richness, depth, and nuanced insights through probing questions. However, we lack systematic evidence on whether these probing questions genuinely add new insights or simply add more mass to factors already identified in the initial top-of-mind responses. In this section, we exploit our large-scale qualitative data collection to systematically assess the discovery process in qualitative interviews.

### **6.1 Interviews versus single open-ended questions**

The standard approach in economics is to collect qualitative data in the form of single open-ended survey questions rather than through interviews. Single open-ended questions are convenient to collect, compatible with large-scale surveys, and have led to many significant discoveries, such as understanding how people reason about taxes and inflation (Andre et al., 2024; Stantcheva, 2021). At the same time, single open-ended questions mainly capture what is top of mind and can be difficult to classify without further context, raising the question of whether we capture the full range of relevant factors when not asking any follow-up questions.

To examine whether there are systematic differences between factors discovered from a single open-ended survey question and a full interview with additional probing questions, we exploit the fact that the first question in our interviews was identical for all respondents: “Can you help me understand the main factors or reasons why you are not participating in the stock market?” For each code in our coding scheme, we derive binary indicators from the hand-coded interview responses for whether this code can be assigned to a respondent based on (i) the response to the first open-ended question alone or (ii) based on the full interview transcript. Figure 3 shows the proportion of respondents assigned a specific code based on their response to the first open-ended question, compared to the proportion assigned the same code after completing the entire interview.<sup>12</sup> Table 2 presents the data in tabular form. Three main patterns emerge.

First, analyzing the initial top-of-mind response, we find that our respondents mainly explain their stock market non-participation with knowledge barriers (46%), lack of income (42%), and a general dislike of risks (21%). In fact, 81% of first responses are assigned at least one of these codes and for 39% of respondents, these are the only codes that are assigned after the first question. No other code is assigned to more than 10% of respondents and only six codes are assigned to at least 5% of respondents.

Second, interviews can identify and reveal qualitatively different factors that are unobserved when relying on a single open-ended survey question. The most prominent example is the *active investing* mental model which is essentially unobserved in people’s first response (0.5%) but revealed in 31.6% of all interviews. This is mirrored by the conceptually related participation cost codes—in which people talk directly about time or monetary costs—that show up in only 3.7% of initial responses but are assigned to 28.2% of interviews overall. As a result, conclusions about the relevance of (misperceived) participation costs strongly depend on whether a researcher is conducting interviews or relying on a single open-ended survey question.

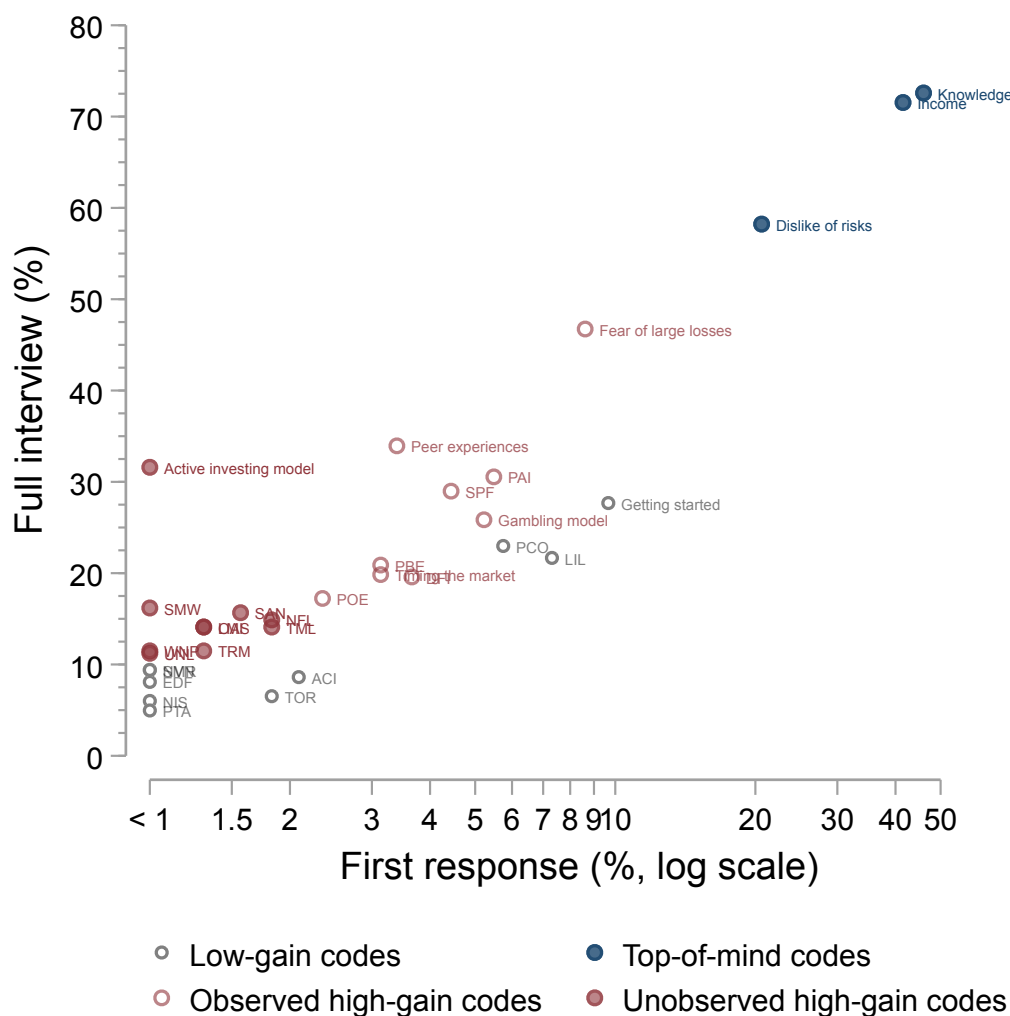
This observation of systematic differences between the initial top-of-mind response and later responses is not specific to participation costs. Several other potentially important explanations for stock market non-participation are essentially unobserved in data derived from a single open-ended survey question but are mentioned by a large share of respondents in a full interview. For example, the probability that respondents mention peer effects, such as bad stock market experiences among close family and friends, increases from 3.4% to 33.9% from the first question to the end of the interview. More broadly, experience-based explanations increase in frequency from 6.3% to 53.3%. This suggests that open-ended survey questions might underestimate the heterogeneity in subjective concerns at both the extensive and the intensive margins.

Third, there is strong heterogeneity in the gains from interviewing across different interview codes. Some codes exhibit a low gain, such as age-based explanations (*TOR*) that are both rare

---

<sup>12</sup>This is a conservative approach that favors the open-ended survey question because it includes codes derived from our qualitative analysis of the interview transcripts that a researcher might have not discovered based only on the text data from a single open-ended survey question.

Figure 3: The gains from interviewing compared to a single open-ended question



*Note:* This figure uses data from the hand-coding of interview responses. The unit of observation are the codes from the coding manual. The figure presents a scatterplot of the share of interviews that are assigned a specific code after the response to the first open-ended question against the share of interviews that are assigned the same code at least once over the course of the full interview. “Top-of-mind” codes are those with a frequency of at least 20% in the first response. “Unobserved high-gain codes” are non-top-of-mind codes with a frequency of less than 2% in the first response, a gain of at least 400% over the course of the interview, and a frequency of at least 10% in the full interview. “Observed high-gain codes” are non-top-of-mind codes with a frequency of at least 2% after the first response, a gain of at least 400% over the course of the interview, and a frequency of at least 10% in the full interview. “Low-gain codes” are the residual category.

Table 2: Distribution of codes after the first and the full interview

|   | Code assigned after |                    |                 |
|---|---------------------|--------------------|-----------------|
|   | First response (%)  | Full interview (%) | Abs. Difference |
| <b>Risk</b>   | 30.8                | 80.9               | 50.1            |
| DOR: Dislike of risks   | 20.6                | 58.2               | 37.6            |
| SPF: Stock price volatility/uncertainty                         | 4.4                 | 29.0               | 24.5            |
| FLL: Fear of large losses                                       | 8.6                 | 46.7               | 38.1            |
| UNL: Unwillingness to make nominal losses                       | 0.5                 | 11.2               | 10.7            |
| SMR: Stock market returns do not justify the risk               | 0.3                 | 9.4                | 9.1             |
| <b>Knowledge</b>  | 50.4                | 75.7               | 25.3            |
| LOK: Lack of knowledge  | 46.0                | 72.6               | 26.6            |
| DKS: Don't know how to get started                              | 9.7                 | 27.7               | 18.0            |
| <b>Constraints</b>  | 45.4                | 78.9               | 33.4            |
| NFI: No funds for stock market investments                      | 41.5                | 71.5               | 30.0            |
| NFL: Need for liquidity   | 1.8                 | 14.9               | 13.1            |
| TOR: Too old/already retired                                    | 1.8                 | 6.5                | 4.7             |
| POE: Prioritize other expenses                                  | 2.3                 | 17.2               | 14.9            |
| LMI: Large minimum investment required                          | 1.3                 | 14.1               | 12.8            |
| <b>Participation costs</b>                                      | 3.7                 | 28.2               | 24.5            |
| TML: Too much time/effort to learn how to get started           | 1.8                 | 14.1               | 12.3            |
| TRM: Too much time is required to manage the stock investments  | 1.3                 | 11.5               | 10.2            |
| EDF: Expensive/difficult to engage with a financial adviser     | 0.5                 | 8.1                | 7.6             |
| <b>Experiences</b>  | 6.3                 | 53.3               | 47.0            |
| PES: Peer effects   | 3.4                 | 33.9               | 30.5            |
| PBE: Personal experiences                                       | 3.1                 | 20.9               | 17.8            |
| NVN: Negative view from news and popular culture                | 0.0                 | 9.4                | 9.4             |
| <b>Mental models</b>  | 9.4                 | 63.7               | 54.3            |
| SMW: Stock market for wealthy people                            | 0.8                 | 16.2               | 15.4            |
| SMC: Gambling model   | 5.2                 | 25.8               | 20.6            |
| ACT: Active investing mental model                              | 0.5                 | 31.6               | 31.1            |
| ITM: Timing the market  | 3.1                 | 19.8               | 16.7            |
| <b>Institutional concerns</b>                                   | 5.2                 | 23.8               | 18.5            |
| UAC: Unfair/anti-capitalist                                     | 2.1                 | 8.6                | 6.5             |
| DFI: Distrust in financial institutions/systems/advisors        | 3.7                 | 19.6               | 15.9            |
| <b>Emotions</b>   | 7.0                 | 32.9               | 25.8            |
| SAN: Stress and anxiety   | 1.6                 | 15.7               | 14.1            |
| PCO: Perceived complexity, overwhelmed, confused                | 5.7                 | 23.0               | 17.2            |
| <b>Asset preferences</b>  | 6.3                 | 37.9               | 31.6            |
| PAI: Prefer alternative investments                             | 5.5                 | 30.5               | 25.1            |
| OAS: Other asset classes provide safer returns                  | 1.3                 | 14.1               | 12.8            |
| PTA: Preference for tangible assets                             | 0.5                 | 5.0                | 4.4             |
| NIS: Need to identify stocks that align with personal standards | 0.0                 | 6.0                | 6.0             |
| <b>Disinterest</b>  | 7.3                 | 27.9               | 20.6            |
| LIL: Lack of interest in or learning about the stock market     | 7.3                 | 21.7               | 14.4            |
| WNP: Would never want to participate                            | 0.0                 | 11.5               | 11.5            |

Note: This table displays the frequency of codes assigned to the first response (Column 1), the full interview (Column 2) and the difference between the two in percentage points (Column 3).

in the first response (1.8%) and the full interview (6.5%). Other codes, such as the fear of losing almost all of one’s investments, strongly gain in prominence over the course of an interview. This suggests that interviews are not mechanically increasing the frequency of all codes, but reveal specific factors that are not on top of people’s minds when responding to a single open-ended survey question.

## 6.2 Unpacking the discovery process

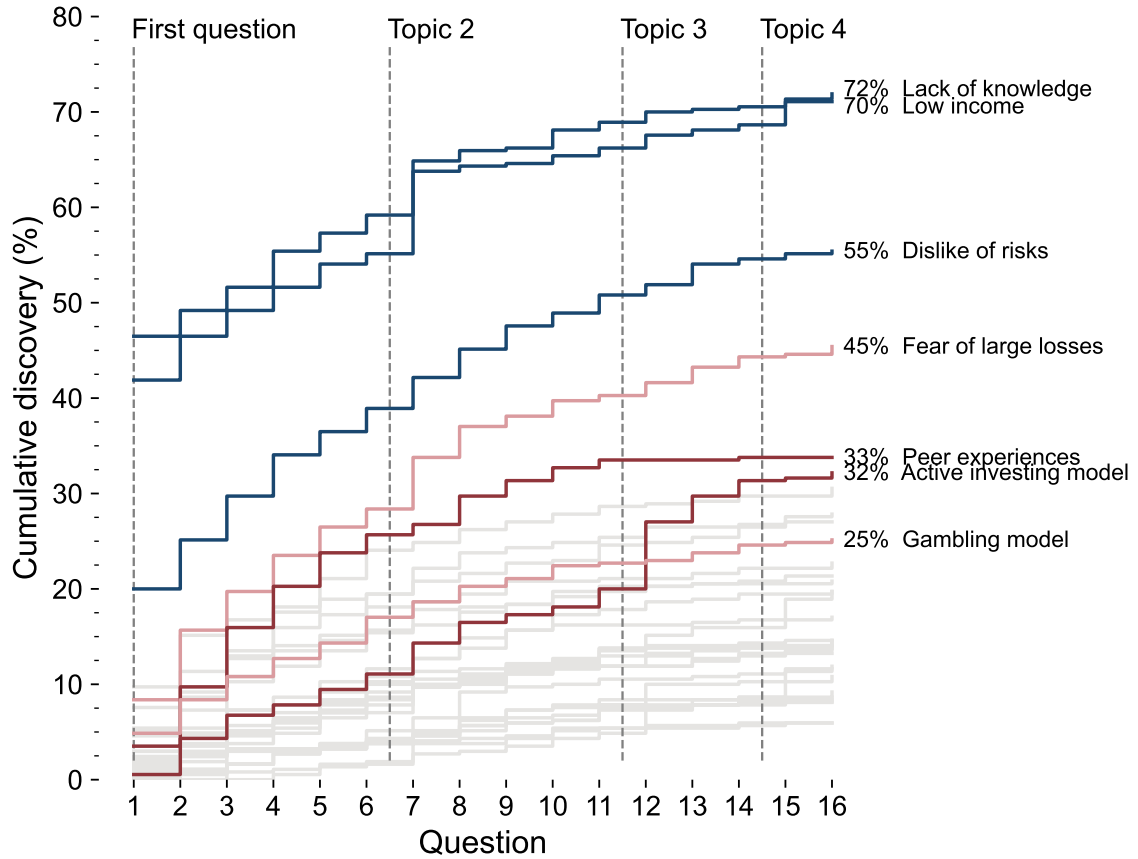
The above finding that AI-conducted interviews can lead to the discovery of themes and recurring patterns that would be largely unobserved in text responses to a single open-ended survey question raises natural follow-up questions about the timing of discovery: How long does it take to reveal a previously unobserved factor? What is the marginal return to asking another interview question? How does the discovery process differ across codes?

To address these questions, we leverage the fact that the interview transcripts are hand-coded at the response level. We derive the share  $d_c(t)$  of respondents that can be assigned code  $c$  based on their cumulative responses to the first  $t$  interview questions, but not taking into account information that was revealed later in the interview. We refer to  $d_c(t)$  as the *discovery process* for code  $c$ . Figure 4 plots the discovery process of all codes that are assigned to at least 5% of respondents by the end of the interview.

For the top-of-mind codes *Lack of knowledge* and *Low income*, the first six probing questions capture most of the discoveries, with marginal returns diminishing beyond that point. In contrast, *Dislike of risks*, another top-of-mind response, shows a more gradual discovery process that only plateaus near the final questions. These patterns of gradual discovery are even more pronounced for the “high-gain” codes, such as *Gambling model* or *Active investing model*, where the marginal returns from further probing remain high throughout the interview. The discovery process for *Active investing model* even shows a steep rise after the third topic transition, where the interviewer introduces a “what if” scenario involving stock market investment. Similar jumps for other codes around the topic transitions demonstrate the value of a topic guide that explores the same underlying questions from multiple angles.

Overall, our findings underscore that the optimal interview length depends on the research questions. For researchers mainly interested in top-of-mind responses, a short interview with a few probing questions might be sufficient. For researchers interested in exploring more complex issues—such as narratives, mental models, and experiences that affect choices—the marginal returns to a longer interview are much higher. In our context, including 16 questions in total seemed to strike a good balance between maximizing discovery and maintaining respondent engagement throughout the interview.

Figure 4: The timing of discoveries along the interview process



Note: This figure uses hand-coded data from the interviews conducted as part of the main survey. We plot the discovery process  $d_c(t)$ , i.e., the share of respondents that can be assigned a specific code  $c$  based on their cumulative responses to the first  $t$  interview questions.

## 7 Interview responses and choices

The previous evidence suggests that AI-conducted interviewers can generate high-quality qualitative data that gives rise to rich insights, capturing factors that are unobserved in responses to a single open-ended survey question. This finding raises a natural question about whether factors discovered later in the interviews reflect *important* reasons for non-participation above and beyond the top-of-mind response, or whether they simply capture “cheap talk” in which the interviewees attempt to fill the conversational space in response to persistent probing. To address this concern, we conducted a follow-up survey with incentivized choices that allows us to examine whether interview codes captured later in the interview predict economic behavior.

**Design** We invited all respondents who completed the qualitative interviews in August/September 2023 to a follow-up survey in May 2024. We manage to recontact 70% of the original sample, which we consider high for a long-term follow-up. Table A.3 provides summary statistics and



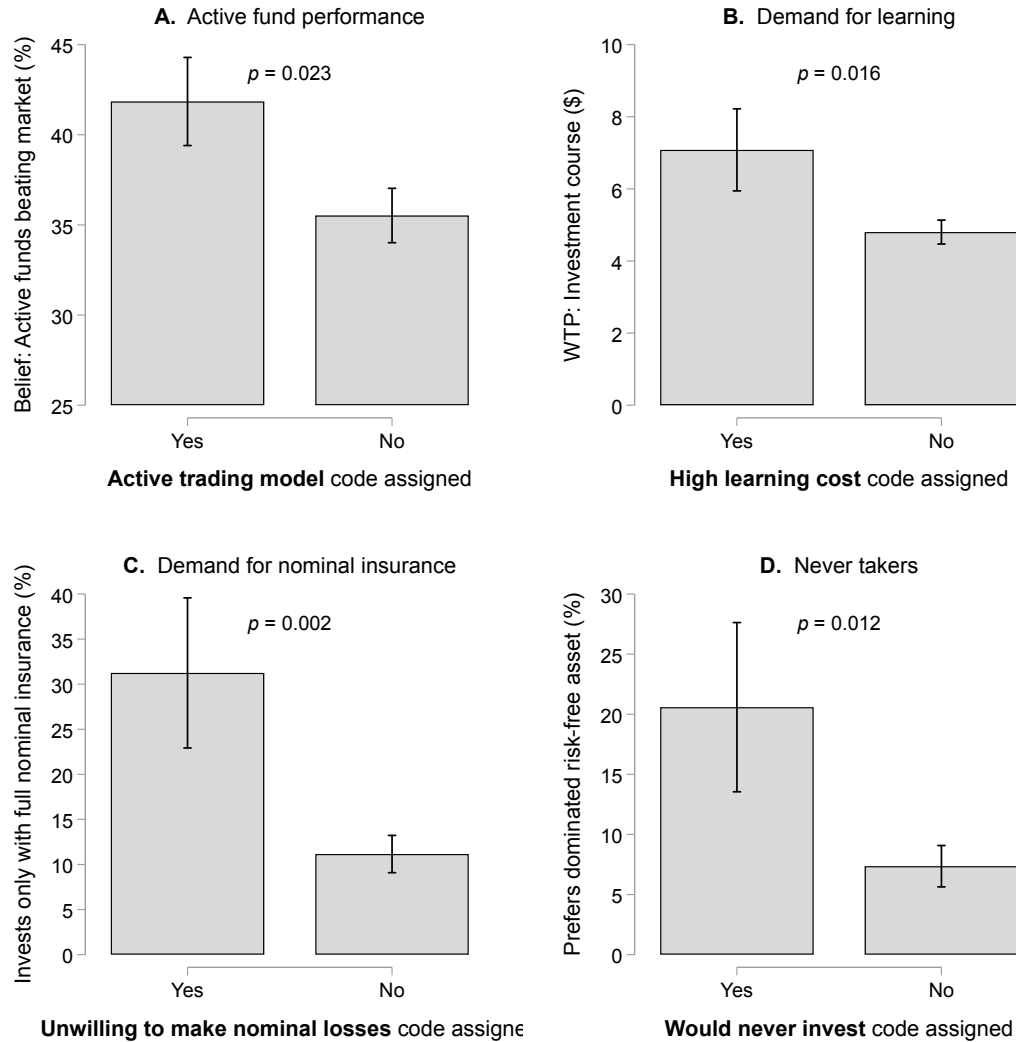
Appendix Section E.5 contains the main experimental instructions. A key design objective is generating testable behavioral predictions for a subset of codes assigned to interview responses. To focus on non-participation motives that only show up in qualitative interviews, we selected four codes that are almost never assigned to the initial response: *active investing model* (0.5% discovery rate in the first response), *high learning cost* (1.8%), *unwilling to make nominal losses* (0.5%), *would never participate* (0%).<sup>13</sup> The survey includes three incentivized outcomes tailored to these interview codes. First, we elicit respondents' incentivized belief about the share of actively managed funds that yielded higher after-fee returns than passively managed funds in 2023 on a scale from 0 to 100%. Respondents could earn a bonus of \$2 if they guess within 2 percentage points of the actual value. This allows us to test whether respondents with an active investing mental model believe that active portfolio management outperforms passive investing. Second, we elicited respondents' willingness to pay for three-month access to an online "four-part series aimed at helping beginners understand the basics of stock market investments and learn how to successfully navigate the market" from MasterClass, a platform providing high-quality learning content. We elicited the willingness to pay using a multiple price list ranging from \$0 to \$20 in increments of \$2. The decision was probabilistically incentivized. It allows us to test whether respondents who worry about high learning costs have a higher willingness to pay for an investment course. Third, we elicited respondents' demand for loss protection using a multiple price list. Respondents repeatedly choose between two options. Option A consists of a certain payment of \$1,000 in twelve months. Option B consists of receiving the value of \$1,000 invested in the S&P 500 in twelve months. In addition, Option B limits the maximum amount that can be lost through this investment ("loss protection"). Across the price list, we vary the protection against nominal losses from 0% to 100%. We implement a randomly selected choice for one out of every 200 respondents. This outcome provides two predictions. Respondents who reveal an unwillingness to make nominal losses ("I would need a guarantee on my investment") should only switch from the risk-free option to an investment in the stock market if offered full insurance against nominal losses. In addition, we obtain the prediction that people who emphasize that they would under no circumstances invest in the stock market should prefer the risk-free option even if offered full insurance.

**Results** Panel A of Figure 5 shows that respondents with an *active investing* mental model of the stock market report a 6.4 percentage point higher belief about the share of active funds that outperform the market ( $p = 0.023$ ) compared to respondents who are not assigned the *active investing* code. Similarly, Panel B shows that the average willingness to pay for an investment course is \$7.08, which is 48% higher than the willingness to pay among respondents who do not mention high learning costs as a barrier to stock market participation ( $p = 0.016$ ). Panel C shows that the probability of switching from the certain payment to an investment in the stock market

---

<sup>13</sup>When regressing a dummy for participating in the follow-up survey on a set of dummies for whether these codes are assigned, we find no evidence of differential attrition. The  $p$ -value of a joint  $F$ -test is 0.416.

Figure 5: Interview codes predict incentivized choices eight months later



*Note:* This figure uses data from the follow-up survey conducted eight months after the qualitative interviews. Each panel shows the average of an incentivized choice separately for respondents that were assigned a specific interview code and those that were not assigned the code. Panel A shows the average belief about the share of actively managed funds that yield higher net returns after fees than passively managed funds. Panel B shows the average willingness to pay for a 3-month online investment course from MasterClass. Panel C indicates the share of respondents who switch from a certain payment of \$1,000 dollars in 12 months to investing \$1,000 in the S&P 500 for 12 months *only* if offered a guarantee to receive back 100% of the initial investment. Panel D indicates the share of respondents who are unwilling to invest in the S&P 500 even if offered a guarantee to receive back 100% of the initial investment. Standard errors of the means are shown as vertical lines. The  $p$ -values for two-sided  $t$ -tests for equality of means are shown in each panel.

*only if* offered full insurance against nominal losses increases by a factor of 2.8 if respondents reveal an unwillingness to invest in assets that include the possibility of “ending up with less” than one started with ( $p = 0.002$ ). Finally, we observe that respondents who were assigned the *would never invest* code are more likely to choose the dominated certain payment over an investment in the S&P 500 with a guaranteed minimum return of 0% ( $p = 0.012$ ).

Taken together, the results confirm that AI interviews reveal factors that are predictive of behavior, even if restricting attention to factors that are mostly unobserved in their initial response to the first open-ended question. This evidence should mitigate concerns about “cheap talk” dominating the discourse in interviews. Section C.2 provide further evidence supporting this conclusion by showing strong correlations between interview codes and closed-end survey questions.

## 8 Conclusion

This paper presents a new method for conducting qualitative interviews using an *AI interviewer* to engage with human respondents. To create a fully autonomous AI interviewer, we employ large language models in a multi-agent architecture, enabling task specialization to improve conversational dynamics. Our agents are given detailed instructions on conducting interviews according to best-practice qualitative research methods. These instructions cover how to ask open-ended questions, manage the conversational flow, and ensure respondent comfort.

We demonstrate the practical feasibility of our approach by conducting 381 AI-conducted interviews with human respondents about their reasons for not participating in the stock market. This data collection generates four broad findings about AI-conducted interviews. First, we document that the AI interviewer closely follows its instructions: It asks open-ended, non-leading, and highly relevant questions. This leads to high interviewee satisfaction among respondents, with a majority even expressing a weak preference for an AI interviewer over a human interviewer.

Second, a qualitative analysis of the interview data demonstrates that AI-conducted interviews generate rich insights about why people are not participating in the stock market. Our main qualitative finding is a prominent role for an *active investing* mental model in which respondents believe that constant monitoring of market conditions and extensive research on individual stocks is necessary to make any money in the stock market. This mental model provides a potentially unifying explanation of other factors frequently mentioned in the interviews, such as worrying about large losses from picking “wrong” stocks, thinking that knowledge barriers are hard to overcome, and why stock market investments for many seem like an overwhelming time commitment. A large-scale survey with a balanced sample of 999 stock owners and non-owners shows that this active investing mental model indeed is much more frequent among non-owners,

demonstrating the quantitative importance of this qualitative discovery.

Third, we use our rich large-scale interview data to provide evidence on several important questions relating to the utility of qualitative interviews in economics. We document systematic differences between factors identified in initial top-of-mind responses and those uncovered in subsequent responses. Mental models, such as the active investing paradigm, only emerge after several probing questions, suggesting the current practice in the economics profession of mainly relying on single open-ended questions can miss important discoveries related to narratives and mental models. Finally, a follow-up study shows that the interview data predicts economic behavior eight months after being collected, suggesting that uninformative “cheap talk” is unlikely to dominate the discourse in qualitative interviews.

Overall, our results demonstrate that AI-conducted interviews can generate rich, high-quality qualitative data. An important benefit of our AI-driven approach is that it dramatically reduces the time and pecuniary costs of conducting qualitative research. As our implementation of an AI interviewer can be integrated into standard survey software (such as Qualtrics), the marginal costs of interviews equal the cost of compensating respondents for their time and the API costs for the language model (about \$0.10 per interview hour at 2024 prices). The direct integration with standard survey software and low marginal costs also allow for conducting qualitative interviews at scale. While the need for large-scale interview data depends on the research question, typical applications in economics often require scale (e.g. when using qualitative interviews to identify mechanisms driving treatment effects, as in Bergman et al., 2024). In addition to low costs and high scalability, another advantage of AI-conducted interviews is that they could help eliminate potential human interviewer biases, such as happier interviewers eliciting more positive responses from interviewees (Stefkovics and Sik, 2022). While AI interviewers could potentially introduce algorithmic biases, the “interviewer effects” from an AI interviewer are consistent across all interviews, reducing variability and increasing interview comparability. Furthermore, as our understanding of potential algorithmic biases improves, they can be mitigated with model fine-tuning and directed prompting.

Although qualitative interviews are not yet widespread in economic research, a few recent studies in addition to ours demonstrate that qualitative interviews can offer important insights and complement traditional quantitative methods. One promising application is to understand treatment effect mechanisms. Bergman et al. (2024) conduct 251 interviews to study why a “moving to opportunity” intervention was successful and found an important role for relaxing bandwidth constraints and addressing specific needs. Our hope is that our AI approach can help accelerate the adoption of such multi-method projects by lowering the barriers to conducting qualitative interviews.

Furthermore, in addition to making mixed-methods projects feasible at much lower costs, we believe the scale offered by AI-conducted interviews will facilitate many novel applications of interviews in economics. For instance, AI interviews can be used to design tailored information

provision treatments where the barriers and constraints identified during the interview can be used to design an information treatment for maximum impact.

## References

- Andre, Peter, Ingar Haaland, Christopher Roth, Mirko Wiederholt, and Johannes Wohlfart**, “Narratives about the Macroeconomy,” Discussion Paper 127, ECONtribute 2024.
- Barberis, Nicholas, Ming Huang, and Richard Thaler**, “Individual Preferences, Monetary Gambles and Stock Market Participation: A Case for Narrow Framing,” *American Economic Review*, 2006, 96 (4), 1069–1090.
- Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer**, “Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice,” *American Economic Review*, 2024, 114 (5), 1281–1337.
- Bewley, Truman F.**, *Why Wages Don’t Fall during a Recession*, Cambridge, MA: Harvard University Press, 1999.
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan**, “Latent Dirichlet Allocation,” *Journal of Machine Learning Research*, 2003, 3 (1), 993–1022.
- Blinder, Alan, Elie R.D. Canetti, David E. Lebow, and Jeremy B. Rudd**, *Asking about Prices: A New Approach to Understanding Price Stickiness*, Russell Sage Foundation, 1998.
- Braun, Virginia and Victoria Clarke**, *Thematic Analysis: A Practical Guide*, SAGE Publications, 2021.
- Bustos, Sebastián, Dina Pomeranz, Juan Carlos Suárez Serrato, José Vila-Belda, and Gabriel Zucman**, “The Race Between Tax Enforcement and Tax Planning: Evidence From a Natural Experiment in Chile,” Working Paper 30114, National Bureau of Economic Research June 2022.
- Campello, Ricardo J. G. B., Davoud Moulavi, and Joerg Sander**, “Density-Based Clustering Based on Hierarchical Density Estimates,” in Jian Pei, Vincent S. Tseng, Longbing Cao, Hiroshi Motoda, and Guandong Xu, eds., *Advances in Knowledge Discovery and Data Mining*, Springer Berlin Heidelberg Berlin, Heidelberg 2013, pp. 160–172.
- Charmaz, Kathy**, “Qualitative interviewing and grounded theory analysis,” in J. Gubrium and J. Holstein, eds., *Handbook of interview research: Context and method*, Thousand Oaks, CA: Sage, 2002, pp. 675–694.
- Corbin, Juliet and Anselm Strauss**, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 4th ed., SAGE Publications, Inc, 2014.
- DeLuca, Stefanie, Lawrence F. Katz, and Sarah C. Oppenheimer**, ““When Someone Cares About You, It’s Priceless”: Reducing Administrative Burdens and Boosting Housing Search Confidence to Increase Opportunity Moves for Voucher Holders,” *RSF: The Russell Sage Foundation Journal of the Social Sciences*, September 2023, 9 (5), 179–211.

- Denzin, Norman K.**, “The reflexive interview and a performative social science,” *Qualitative Research*, 2001, 1 (1), 23–46.
- Dietterich, Thomas G.**, “Ensemble Methods in Machine Learning,” in “International Workshop on Multiple Classifier Systems” Springer 2000, pp. 1–15.
- Dufo, Esther**, “The Economist as Plumber,” *The American Economic Review*, 2017, 107 (5), 1–26.
- Falk, Armin and James J. Heckman**, “Lab Experiments Are a Major Source of Knowledge in the Social Sciences,” *Science*, 2009, 326 (5952), 535–538.
- Favilukis, Jack**, “Inequality, stock market participation, and the equity premium,” *Journal of Financial Economics*, 2013, 107 (3), 740–759.
- Fuster, Andreas and Basit Zafar**, “Survey Experiments on Economic Expectations,” in “Handbook of Economic Expectations,” Elsevier, 2023, chapter 4, pp. 107–130.
- Gentzkow, Matthew, Bryan Kelly, and Matt Taddy**, “Text as data,” *Journal of Economic Literature*, 2019, 57 (3), 535–574.
- Gibson, Kerry**, “Bridging the digital divide: Reflections on using WhatsApp instant messenger interviews in youth research,” *Qualitative Research in Psychology*, 2022, 19 (3), 611–631.
- Glaser, Barney and Anselm Strauss**, *Discovery of Grounded Theory: Strategies for Qualitative Research*, Routledge, 2017.
- Gomes, Francisco, Michael Haliassos, and Tarun Ramadorai**, “Household Finance,” *Journal of Economic Literature*, 2021, 59 (3), 919–1000.
- Grootendorst, Maarten**, “BERTopic: Neural topic modeling with a class-based TF-IDF procedure,” *arXiv preprint arXiv:2203.05794*, 2022.
- Guiso, Luigi and Paolo Sodini**, “Household finance: An emerging field,” in “Handbook of the Economics of Finance,” Vol. 2, Elsevier, 2013, pp. 1397–1532.
- Guo, Taicheng, Xiuying Chen, Yaqi Wang, Ruidi Chang, Shichao Pei, Nitesh V. Chawla, Olaf Wiest, and Xiangliang Zhang**, “Large Language Model based Multi-Agents: A Survey of Progress and Challenges,” 2024.
- Haaland, Ingar, Christopher Roth, and Johannes Wohlfart**, “Designing Information Provision Experiments,” *Journal of Economic Literature*, 3 2023, 61 (1), 3–40.
- Haaland, Ingar K., Christopher Roth, Stefanie Stantcheva, and Johannes Wohlfart**, “Measuring What is Top of Mind,” Working Paper 32421, National Bureau of Economic Research May 2024.
- Haliassos, Michael and Carol C. Bertaut**, “Why do so few hold stocks?,” *The Economic Journal*, 1995, 105 (432), 1110–1129.
- Harrison, Glenn W. and John A. List**, “Field Experiments,” *Journal of Economic Literature*, 2004, 42 (4), 1009–1055.

- Henkel, Luca and Christian Zimpelmann**, “Proud to not own stocks: How identity shapes financial decisions,” *IZA Discussion Paper*, 2023.
- Himelein, Kristen**, “Interviewer Effects in Subjective Survey Questions: Evidence From Timor-Leste,” *International Journal of Public Opinion Research*, 09 2015, 28 (4), 511–533.
- Hurd, Michael, Maarten van Rooij, and Joachim Winter**, “Stock Market Expectations of Dutch Households,” *Journal of Applied Econometrics*, April–May 2011, 26 (3), 416–436.
- Knott, Eleanor, Aliya Hamid Rao, Kate Summers et al.**, “Interviews in the social sciences,” *Nature Reviews Methods Primers*, 2022, 2, 73.
- Kojima, Takeshi, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa**, “Large Language Models are Zero-Shot Reasoners,” 2023.
- Korinek, Anton**, “Generative AI for Economic Research: Use Cases and Implications for Economists,” *Journal of Economic Literature*, 2023, 61 (4), 1281–1317.
- Legard, Robin, Jill Keegan, and Kit Ward**, “In-depth Interviews,” in Jane Ritchie and Jane Lewis, eds., *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, SAGE Publications, 2003, pp. 138–169.
- Link, Sebastian, Andreas Peichl, Christopher Roth, and Johannes Wohlfart**, “Attention to the Macroeconomy,” *Available at SSRN 4697814*, 2024.
- Mankiw, N. Gregory and Stephen P. Zeldes**, “The Consumption of Stockholders and Non-stockholders,” *Journal of Financial Economics*, 1991, 29 (1), 97–112.
- McInnes, Leland, John Healy, and James Melville**, “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,” 2020.
- Mishler, Elliot G.**, *Research Interviewing: Context and Narrative*, Cambridge, MA: Harvard University Press, 1986.
- Namey, Emily, Greg Guest, Amy O’Regan, Christine L. Godwin, Jamilah Taylor, and Andres Martinez**, “How Does Mode of Qualitative Data Collection Affect Data and Cost? Findings from a Quasi-Experimental Study,” *Field Methods*, 2020, 32 (1), 58–74.
- Nelson, Laura K.**, “Computational Grounded Theory: A Methodological Framework,” *Sociological Methods & Research*, 2020, 49 (1), 3–42.
- Nielsen, Kirby**, “Preferences for the resolution of uncertainty and the timing of information,” *Journal of Economic Theory*, 2020, 189, 105090.
- Oakley, Ann**, “Interviewing women: A contradiction in terms,” in Helen Roberts, ed., *Doing Feminist Research*, London: Routledge, 1981.
- Park, Joon Sung, Joseph C. O’Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein**, “Generative Agents: Interactive Simulacra of Human Behavior,” *arXiv preprint arXiv:2304.03442*, 2023.
- Peer, Eyal, David Rothschild, Ariel Gordon, Zak Evernden, and Ekaterina Damer**, “Data quality of platforms and panels for online behavioral research,” *Behavior Research Methods*, 2022, 54, 1643–1662.

- Ralund, Snorre, Hjalmar Bang Carlsen, Robert Klemmensen, and David Dreyer Lassen**, “Measurement Error and Model Instability in Automated Text Analysis: The Case of Topic Models,” *Working Paper*, 2022.
- Rubin, Herbert J. and Irene S. Rubin**, *Qualitative Interviewing: The Art of Hearing Data*, 3 ed., Sage Publications, 2011.
- Spradley, James P.**, *The Ethnographic Interview*, New York: Holt, Rinehart and Winston, 1979.
- Stantcheva, Stefanie**, “Understanding Tax Policy: How do People Reason?,” *The Quarterly Journal of Economics*, 2021, 136 (4), 2309–2369.
- , “How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible,” *Annual Review of Economics*, 2023, 15, 205–234.
- Stefkovics, Ádám and Endre Sik**, “What Drives Happiness? The Interviewer’s Happiness,” *Journal of Happiness Studies*, 2022, 23 (6), 2745–2762.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin**, “Attention is all you need,” *Advances in Neural Information Processing Systems*, 2017, 30.
- Wei, Jason, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou**, “Chain-of-Thought Prompting Elicits Reasoning in Large Language Models,” 2023.
- Zhang, Yue, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen et al.**, “Siren’s Song in the AI Ocean: A Survey on Hallucination in Large Language Models,” *arXiv preprint arXiv:2309.01219*, 2023.
- Ziegler, Daniel M., Nisan Stiennon, Jeffrey Wu, Tom B. Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving**, “Fine-tuning Language Models From Human Preferences,” *arXiv preprint arXiv:1909.08593*, 2019.



For online publication only:

# **Conducting Qualitative Interviews with AI**

Felix Chopra and Ingar Haaland

Section A contains additional tables.

Section B contains additional figures.

Section C contains additional details about AI interviews.

Section D contains additional details about the hand-coding of interviews.

Section E presents the key experimental instructions for our surveys and experiments.

## A Additional tables

Table A.1: Overview of data collections

| <i>Study and population</i>                      | <i>Recruitment</i>                                      | <i>n</i> | <i>Study description</i>  |
|--|---|----------|---|
| <b>Main study</b>                                |   |          |   |
| General population (US), non-stockowners         | Prolific, Aug-Sep 2023, screener survey for recruitment | 381      | Descriptive survey with an integrated 30-minute text-based interview led by an AI interviewer. Post-interview survey questions about the interview experiences. Additional survey modules on factors associated with stock market non-participation, respondent characteristics and household finances.   |
| <b>Follow-up study</b>                           |   |          |   |
| Main study participants with completed interview | Prolific, May 2024                                      | 266      | Long-term follow-up with participants from the main study. Descriptive survey that includes incentivized outcomes to study whether selected interview codes predict choices: (i) demand for insurance against nominal losses, (ii) willingness to pay for an investment course, (ii) incentivized belief about the performance of actively managed funds. |
| <b>Validation survey</b>                         |   |          |   |
| General population (US)                          | Prolific, May 2024                                      | 1,000    | Descriptive survey with equal share of stockowners and non-owners. Survey questions aimed at measuring whether respondents hold the “active investing” mental model that we hypothesized based on the qualitative analysis of the interview transcripts from the main study.  |
| <b>Selection experiment</b>                      |   |          |   |
| General population (US)                          | Prolific, Nov 2023                                      | 499      | Respondents are asked to select all studies that they would like to participate in if we run it: (i) 40-minute survey, (ii) 40-minute interview with a human interviewer, (iii) 40-minute interview with an AI interviewer. These studies are shown on separate pages in randomized order.  |

*Note:* This table provides an overview of the interviews, surveys and experiments presented in this paper. The key experimental instructions are shown in Section E.

Table A.2: Summary statistics: Main study

|                             | Min       | Mean       | Median     | Max        | N   |
|-----------------------------|-----------|------------|------------|------------|-----|
| <b>A. Demographics</b>      |           |            |            |            |     |
| Age                         | 19.00     | 39.43      | 36.00      | 78.00      | 381 |
| Female                      | 0.00      | 0.64       | 1.00       | 1.00       | 381 |
| College education           | 0.00      | 0.53       | 1.00       | 1.00       | 381 |
| Full-time employment        | 0.00      | 0.47       | 0.00       | 1.00       | 381 |
| White                       | 0.00      | 0.77       | 1.00       | 1.00       | 381 |
| African American/Black      | 0.00      | 0.12       | 0.00       | 1.00       | 381 |
| Hispanic                    | 0.00      | 0.11       | 0.00       | 1.00       | 381 |
| Region                      |           |            |            |            |     |
| Northeast                   | 0.00      | 0.18       | 0.00       | 1.00       | 381 |
| Midwest                     | 0.00      | 0.24       | 0.00       | 1.00       | 381 |
| West                        | 0.00      | 0.17       | 0.00       | 1.00       | 381 |
| South                       | 0.00      | 0.41       | 0.00       | 1.00       | 381 |
| Household size              | 1.00      | 2.96       | 3.00       | 10.00      | 381 |
| Number of children          | 0.00      | 0.77       | 0.00       | 4.00       | 377 |
| <b>B. Finances</b>          |           |            |            |            |     |
| Household income (\$)       | 35,000.00 | 72,119.42  | 65,000.00  | 212,500.00 | 381 |
| Total financial assets (\$) | 0.00      | 78,671.13  | 17,500.00  | 550,000.00 | 381 |
| Non-mortgage debt (\$)      | 0.00      | 27,529.92  | 7,500.00   | 400,000.00 | 381 |
| Housing                     |           |            |            |            |     |
| Homeowner                   | 0.00      | 0.51       | 1.00       | 1.00       | 381 |
| Home value (\$)             | 12,500.00 | 242,654.64 | 225,000.00 | 525,000.00 | 194 |
| Any mortgage debt           | 0.00      | 0.28       | 0.00       | 1.00       | 381 |
| Total mortgage debt (\$)    | 12,500.00 | 123,247.66 | 87,500.00  | 475,000.00 | 107 |
| Two months liquid savings   | 0.00      | 0.60       | 1.00       | 1.00       | 381 |

*Note:* This table displays summary statistics for the main study conducted in August/September 2023. Table A.1 provides an overview of all data collections. *Home value* and *total mortgage debt* are conditional on non-zero values.

Table A.3: Summary statistics: Follow-up study

|                             | Min       | Mean       | Median     | Max        | N   |
|-----------------------------|-----------|------------|------------|------------|-----|
| <b>A. Demographics</b>      |           |            |            |            |     |
| Age                         | 19.00     | 42.24      | 40.00      | 78.00      | 266 |
| Female                      | 0.00      | 0.62       | 1.00       | 1.00       | 266 |
| College education           | 0.00      | 0.56       | 1.00       | 1.00       | 266 |
| Full-time employment        | 0.00      | 0.52       | 1.00       | 1.00       | 266 |
| White                       | 0.00      | 0.79       | 1.00       | 1.00       | 266 |
| African American/Black      | 0.00      | 0.13       | 0.00       | 1.00       | 266 |
| Hispanic                    | 0.00      | 0.11       | 0.00       | 1.00       | 266 |
| Region                      |           |            |            |            |     |
| Northeast                   | 0.00      | 0.17       | 0.00       | 1.00       | 266 |
| Midwest                     | 0.00      | 0.21       | 0.00       | 1.00       | 266 |
| West                        | 0.00      | 0.18       | 0.00       | 1.00       | 266 |
| South                       | 0.00      | 0.44       | 0.00       | 1.00       | 266 |
| Household size              | 1.00      | 2.86       | 3.00       | 10.00      | 266 |
| Number of children          | 0.00      | 0.87       | 0.00       | 4.00       | 263 |
| <b>B. Finances</b>          |           |            |            |            |     |
| Household income (\$)       | 35,000.00 | 69,285.71  | 65,000.00  | 212,500.00 | 266 |
| Total financial assets (\$) | 0.00      | 82,707.14  | 17,500.00  | 550,000.00 | 266 |
| Non-mortgage debt (\$)      | 0.00      | 22,725.19  | 3,750.00   | 400,000.00 | 266 |
| Housing                     |           |            |            |            |     |
| Homeowner                   | 0.00      | 0.52       | 1.00       | 1.00       | 266 |
| Home value (\$)             | 12,500.00 | 239,492.75 | 225,000.00 | 525,000.00 | 138 |
| Any mortgage debt           | 0.00      | 0.29       | 0.00       | 1.00       | 266 |
| Total mortgage debt (\$)    | 12,500.00 | 112,820.51 | 87,500.00  | 475,000.00 | 78  |
| Two months liquid savings   | 0.00      | 0.61       | 1.00       | 1.00       | 266 |

*Note:* This table displays summary statistics for the follow-up study conducted in May 2024 with participants from our main study. Table A.1 provides an overview of all data collections. *Home value* and *total mortgage debt* are conditional on non-zero values.

Table A.4: Summary statistics: Validation study

|                             | Min      | Mean       | Median    | Max        | N   |
|-----------------------------|----------|------------|-----------|------------|-----|
| <b>A. Sociodemographics</b> |          |            |           |            |     |
| Age                         | 18.00    | 37.15      | 34.00     | 80.00      | 999 |
| Female                      | 0.00     | 0.50       | 0.00      | 1.00       | 999 |
| College education           | 0.00     | 0.58       | 1.00      | 1.00       | 999 |
| Full-time employment        | 0.00     | 0.46       | 0.00      | 1.00       | 999 |
| Northeast                   | 0.00     | 0.21       | 0.00      | 1.00       | 999 |
| Midwest                     | 0.00     | 0.18       | 0.00      | 1.00       | 999 |
| West                        | 0.00     | 0.25       | 0.00      | 1.00       | 999 |
| South                       | 0.00     | 0.36       | 0.00      | 1.00       | 999 |
| <b>B. Finances</b>          |          |            |           |            |     |
| Household income (\$)       | 5,000.00 | 76,071.07  | 65,000.00 | 225,000.00 | 999 |
| Total financial assets (\$) | 0.00     | 104,810.66 | 37,500.00 | 550,000.00 | 999 |
| Non-mortgage debt (\$)      | 0.00     | 22,182.63  | 3,750.00  | 400,000.00 | 999 |
| Homeowner                   | 0.00     | 0.46       | 0.00      | 1.00       | 999 |
| Any mortgage debt           | 0.00     | 0.25       | 0.00      | 1.00       | 999 |
| Two months liquid savings   | 0.00     | 0.68       | 1.00      | 1.00       | 999 |

*Note:* This table displays summary statistics for the validation study conducted in May 2024. Table A.1 provides an overview of all data collections.

Table A.5: Summary statistics: Selection experiment

|                       | Min      | Mean      | Median    | Max        | N   |
|-----------------------|----------|-----------|-----------|------------|-----|
| Age                   | 18.00    | 42.63     | 40.00     | 79.00      | 497 |
| Female                | 0.00     | 0.49      | 0.00      | 1.00       | 497 |
| College education     | 0.00     | 0.72      | 1.00      | 1.00       | 497 |
| Full-time employment  | 0.00     | 0.52      | 1.00      | 1.00       | 497 |
| Northeast             | 0.00     | 0.26      | 0.00      | 1.00       | 497 |
| Midwest               | 0.00     | 0.22      | 0.00      | 1.00       | 497 |
| West                  | 0.00     | 0.15      | 0.00      | 1.00       | 497 |
| South                 | 0.00     | 0.38      | 0.00      | 1.00       | 497 |
| Household income (\$) | 5,000.00 | 84,704.41 | 75,000.00 | 212,500.00 | 499 |
| Any savings (0/1)     | 0.00     | 0.80      | 1.00      | 1.00       | 499 |

*Note:* This table displays summary statistics for the selection experiment conducted in November 2023. Table A.1 provides an overview of all data collections.

Table A.6: Validation survey: Robustness

|                          | Dependent variable:                         |   |   |  |   |  |
|--------------------------|---|---|---|--|---|--|
|                          | (1)<br>Active<br>trading<br>mental<br>model | (2)<br>Daily<br>monitoring<br>important | (3)<br>Quick<br>reactions<br>to market<br>important | (4)<br>Actively<br>picking<br>winner stocks<br>important | (5)<br>Public info<br>required<br>to read<br>(z-scored) | (6)<br>Log weekly<br>portfolio<br>management<br>time |
| <i>Panel A.</i>          |   |   |   |  |   |  |
| Non-owner                | 0.176***<br>(0.032)                         | 0.120***<br>(0.031)                     | 0.106***<br>(0.033)                                 | 0.155***<br>(0.034)                                      | 0.334***<br>(0.069)                                     | 0.334***<br>(0.082)                                  |
| Financial literacy (0-3) | -0.113***<br>(0.018)                        | -0.072***<br>(0.019)                    | -0.124***<br>(0.020)                                | -0.089***<br>(0.018)                                     | -0.070*<br>(0.040)                                      | -0.225***<br>(0.043)                                 |
| N                        | 999   | 999                                     | 999   | 999  | 999   | 999  |
| R <sup>2</sup>           | 0.098                                       | 0.051                                   | 0.099   | 0.089  | 0.045   | 0.086  |
| Controls                 | Yes   | Yes                                     | Yes   | Yes  | Yes   | Yes  |
| <i>Panel B.</i>          |   |   |   |  |   |  |
| Active trading model     |   | 0.307***<br>(0.024)                     | 0.387***<br>(0.028)                                 | 0.301***<br>(0.033)                                      | 0.783***<br>(0.065)                                     | 1.210***<br>(0.068)                                  |
| Financial literacy (0-3) |   | -0.040**<br>(0.018)                     | -0.082***<br>(0.020)                                | -0.059***<br>(0.019)                                     | 0.010<br>(0.040)  | -0.093**<br>(0.040)                                  |
| N                        |   | 999                                     | 999   | 999  | 999   | 999  |
| R <sup>2</sup>           |   | 0.139                                   | 0.219   | 0.148  | 0.151   | 0.291  |
| Controls                 |   | Yes                                     | Yes   | Yes  | Yes   | Yes  |

*Note:* This table displays regression estimates using data from the validation survey conducted in May 2024 with an equal share of stockowners and non-owners. The dependent variables in columns 1–4 are binary variables. “Non-owner” is a dummy for respondents without any exposure to the stock market. “active investing mental model” is a dummy for agreement with the active investing view. “Daily monitoring important” is a dummy for respondents selecting “daily” or “multiple times a day” when asked about how often it is necessary to monitor market conditions and check one’s portfolio. “Quick reactions to market important” is a dummy for selecting “the key to stock market success is to react quickly to daily market changes.” “Actively picking winner stocks important” is a dummy for selecting “the key to stock market success is to actively pick a few stocks that will increase in value and avoid stocks that will decrease in value.” “Public info required to read” is a z-scored number of public information sources that respondents think one has to “regularly consult when participating in the stock market to make a return on one’s portfolio.” “Log weekly portfolio management time” is the log of the weekly hours that respondents think one would have to spend managing one’s portfolio, measured on an interval scale from “less than 1 hour per week” to “more than 15 hours per week.” All regressions include the following set of sociodemographic controls: age, gender, log household income, and indicators for having completed a college degree, working full-time, owning a house, having a mortgage, and having non-zero savings. “Financial literacy” is the number of correctly answered questions from the “Big 3” financial literacy inventory. Robust standard errors are shown in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A.7: No differential selection into AI interviews compared to surveys

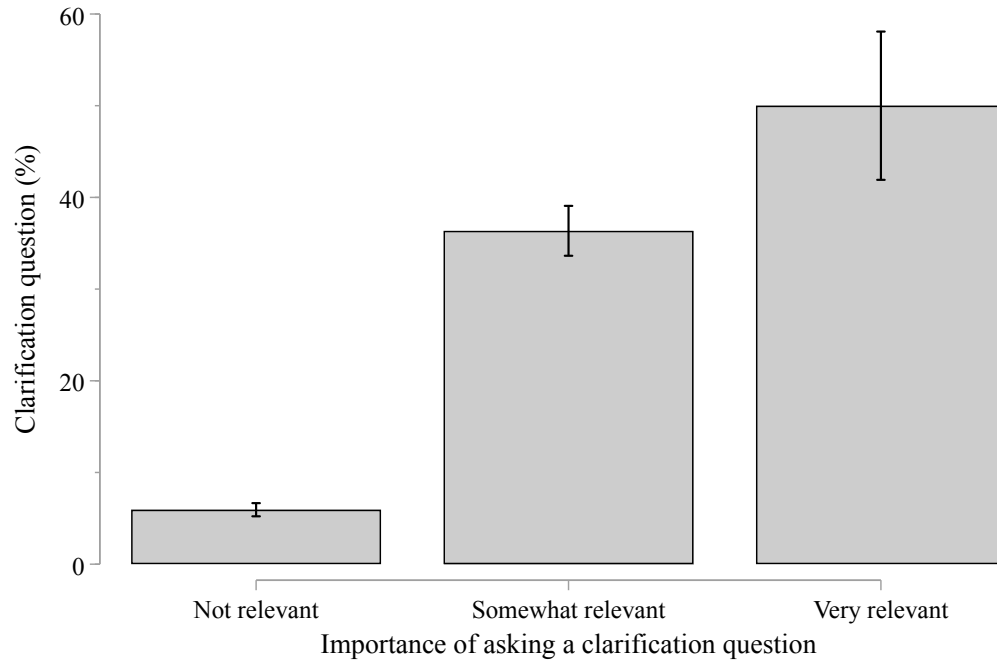
|  | Dependent variable: Demand for study participation (binary) |                   |                   |                  |                   |                   |
|--|---|-------------------|-------------------|------------------|-------------------|-------------------|
|  | (1)   | (2)               | (3)               | (4)              | (5)               | (6)               |
| AI interview                           | 0.005<br>(0.034)  | -0.022<br>(0.060) | 0.007<br>(0.035)  | 0.003<br>(0.049) | 0.007<br>(0.049)  | -0.010<br>(0.075) |
| AI interview x Willing to write        |   | 0.036<br>(0.072)  |                   |                  |                   |                   |
| AI interview x Stock non-participation |   |                   | -0.014<br>(0.116) |                  |                   |                   |
| AI interview x Above median age        |   |                   |                   | 0.004<br>(0.068) |                   |                   |
| AI interview x Female                  |   |                   |                   |                  | -0.012<br>(0.067) |                   |
| AI interview x College                 |   |                   |                   |                  |                   | 0.009<br>(0.084)  |
| N                                      | 333   | 333               | 333               | 333              | 332               | 332               |
| R <sup>2</sup>                         | 0.000   | 0.003             | 0.008             | 0.000            | 0.001             | 0.007             |
| Mean of dep. var.                      | 0.895   | 0.895             | 0.895             | 0.895            | 0.898             | 0.898             |

*Note:* This table shows regression estimates using data from the selection experiment. The dependent variable is a dummy for whether respondents would like to participate in the described study. All regressions restrict to the first study described to respondents, using only studies that were described as including an AI interview or a survey without an interview. “AI interview” is a dummy for whether the study was described as including a text-based interview with an AI, or not. The omitted category is a study described as a regular survey without an interview. “Willing to write” is a dummy for whether respondents wrote at least 20 words in a screener survey. “Stock non-participation” is a dummy for whether respondents own any stocks or stock mutual funds. “Above median age” is a dummy for whether respondents have above median age. “Female” is a dummy for female respondents. “College” is a dummy for having completed a college degree. Robust standard errors are shown in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## B Additional figures

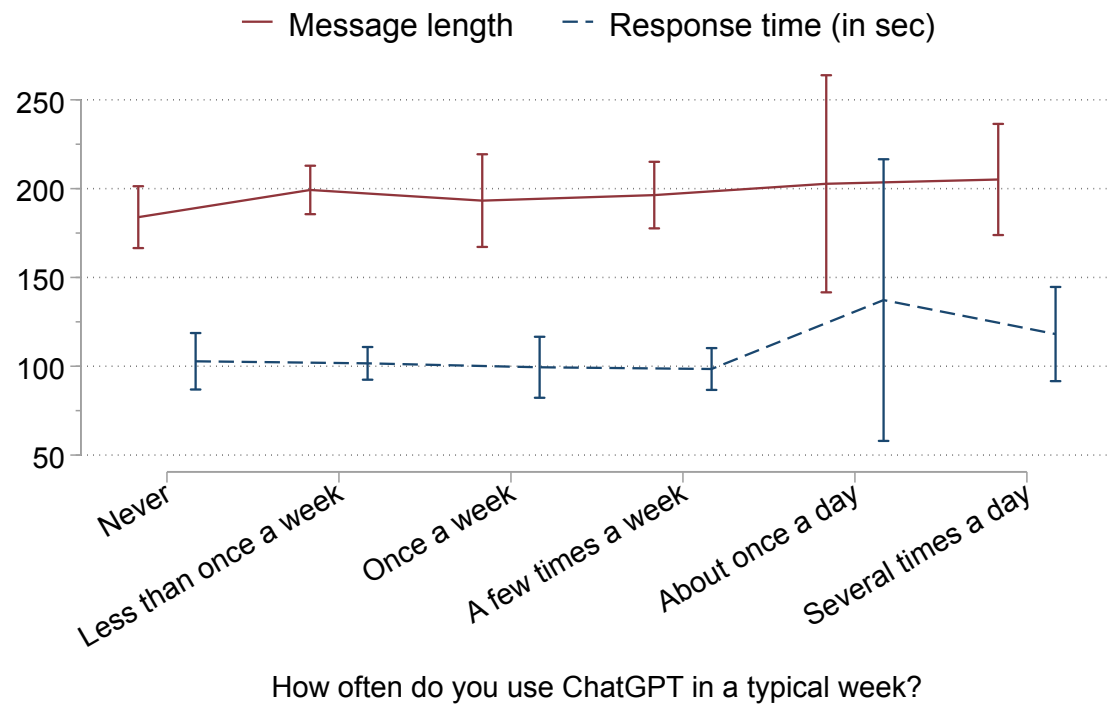
Figure B.1: When does the AI interviewer ask clarifying questions?



*Note:* This figure uses the evaluations of the human coders and shows the fraction of clarifying questions asked by the AI interviewer following an answer for which a clarification question was considered “Not relevant,” “Somewhat relevant,” or “Very relevant.” Vertical lines indicate 95% confidence intervals.

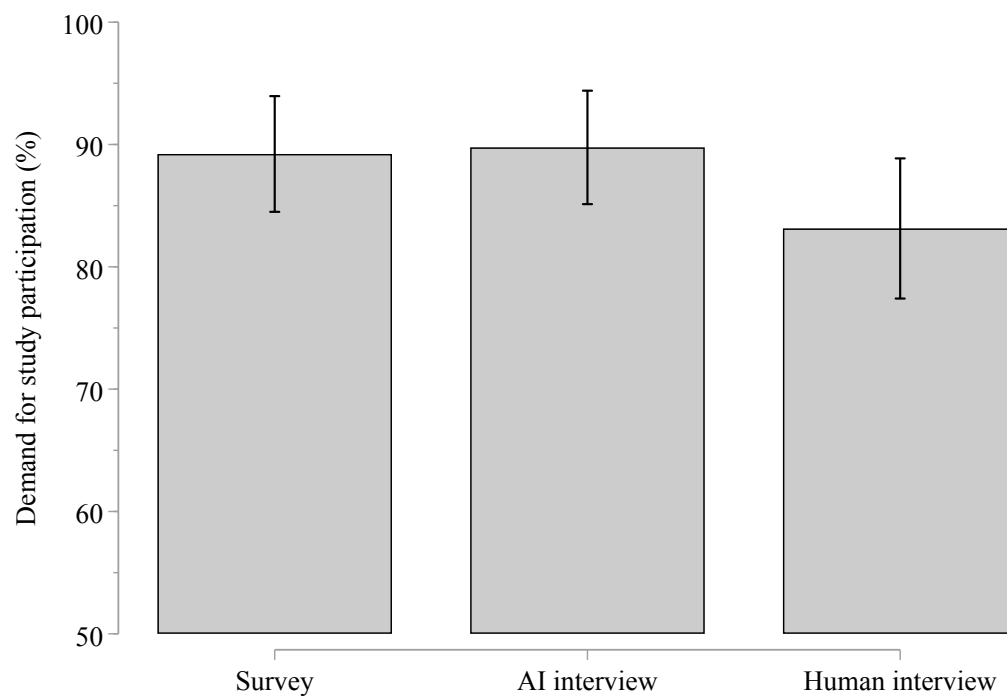


Figure B.2: ChatGPT usage is not associated with lower effort during the interview



*Note:* This figure presents plots of the average response length (total character count) and the average response time of respondents against their self-reported use of generative AI tools such as ChatGPT, Bard, or Bing AI in a typical week (measured on a 6-point scale). 95% confidence intervals are shown as vertical error bars. None of the differences in means are statistically significant at conventional levels ( $p > 0.10$ ).

Figure B.3: Demand for surveys compared to AI- and human-led interviews



*Note:* This figure uses data from the first vignette in the selection experiment. “Demand for study participation” is the share of respondents that would like to participate in a 40-minute study. We separately show the demand for studies described as a regular survey, a survey including a text-based interview with an AI, and a survey including a text-based interview with a human. 95% confidence intervals are shown.

## C Additional details about AI interviews

This section contains additional details and discussions of the AI interviewer. Section C.1 contains a discussion of possible extensions of our current AI interviewer. Section C.3 contains details on the technical implementation, the parameter settings and prompts used for the AI interviewer. Section C.4 contains an example interview to demonstrate the nature of conversational dynamics

### C.1 Extensions of the AI interviewer

**Endogenous topic transitions** In our current design, the length and order of interview topics are exogenously determined. If a more flexible interview structure is desired, the topic can be presented with the set of  $K$  currently remaining topics from the topic guide and instructions to select the next interview topic according to some criterion determined by the researcher. Similarly, the number of probing questions per topic can be endogenized by using a large language model to determine whether the current line of probing has exhausted the subject.

**Voice and facial recognition** In principle, implementing voice and facial recognition with newer models like OpenAI’s GPT-4o, which are trained on diverse datasets, is straightforward. Integrating these technologies into the AI interviewer can potentially enhance the interview experience and allow for a more natural conversation, especially for potentially sensitive topics where emotional cues are particularly important.

Using voice recognition, the AI interview can understand diverse accents and speech patterns, making it easier to interview “hard-to-reach” groups who might not have the ability or opportunity to participate in written interviews. Furthermore, a voice-based interview can be combined with video and facial recognition technology to further help the AI interviewer detect and respond to emotional cues. This capability enables the AI interviewer to better understand confusion or discomfort in the respondent’s expressions, allowing it to adjust questions to provide clarification and ease tension. Another important advantage of video technology is its ability to verify the identity of respondents, ensuring the integrity of the interview process.

**Multi-agent probing agent** Another natural extension is to replace the probing agent with a set of more specialized agents, similar to ensemble techniques in machine learning (Dietterich, 2000). For example, probing questions can be conceptually disaggregated into three distinct categories: *clarification* requests, follow-up questions aimed at increasing *depth* in responses, and questions aimed at expanding the *breadth* of the conversation. An alternative design would thus consist of three separate language models tasked with proposing a clarification, depth, and

breadth question, respectively. A fourth language model then chooses one of the questions.<sup>1</sup> In principle, this ensemble approach can be made arbitrarily complex by integrating iterative feedback loops between agents (Park et al., 2023).

**Model fine-tuning** The quality of probing questions could potentially be improved by fine-tuning domain-general language models on training data consisting of high-quality interview transcripts from human subjects (Ziegler et al., 2019). The returns to fine-tuning might be especially large for LLMs with poorer context understanding than frontier models like GPT-4.<sup>2</sup> While this raises ethical concerns vis-à-vis the human subjects represented in the training data, fine-tuning might allow researchers to encode the difficult-to-describe implicit notion of what constitutes an effective probing question.

## C.2 Correlations with matched survey instruments

As part of a the post-interview survey module in our main data collection, we presented respondents with list of common reasons for stock market non-participation derived from previous research in randomized order and asked them to select all that apply. We match response options with interview codes derived from our inductive coding scheme. Figure C.1 plots the coefficients from regressing a dummy for selecting a specific non-participation reason in this multiple choice question on a dummy for whether a corresponding code was assigned during the hand-coding of the interview transcript. We find that interview codes tend to increase the probability that a matching non-participation reason is selected in a closed-ended survey question.

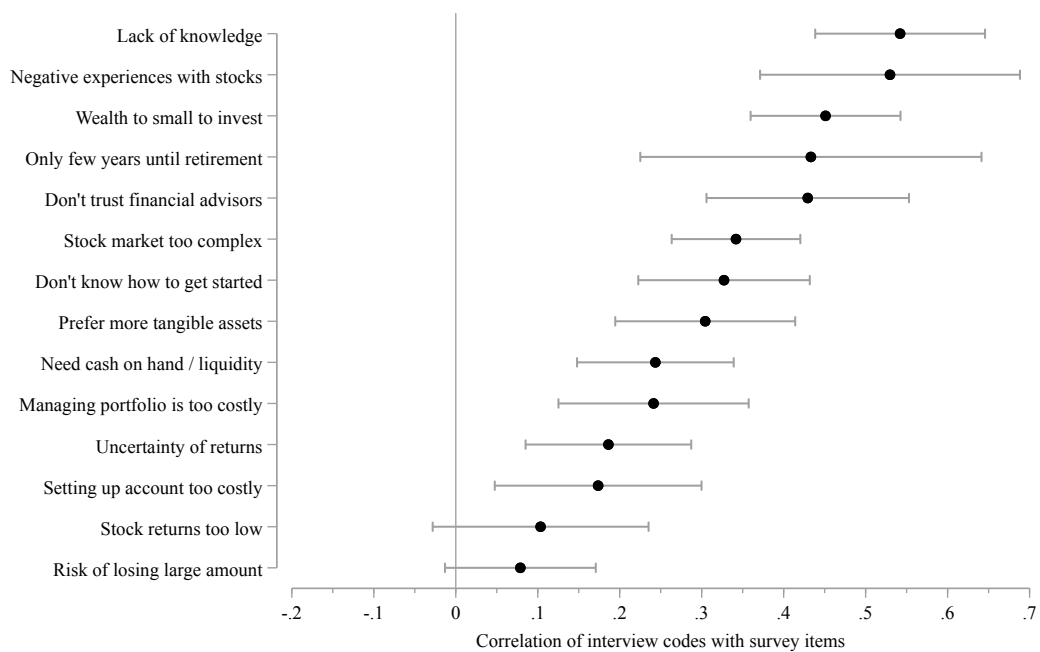
Moreover, Table C.1 shows that interview codes correlate with respondent characteristics. For example, respondents assigned the *dislike of risks* code are 55.4% of a standard deviation more risk averse (Column 1). Respondents assigned the *lack of knowledge* code have 28% of a standard deviation lower financial literacy as measured by the “big 3” financial literacy inventory (Column 2). Similarly, a *tangible asset preference* is associated with a higher likelihood of owning real estate, and reasons associated with being close to retirement are associated with an average age of 60 compared to an average age of 38 among those who do not mention this concern.

---

<sup>1</sup>Our experiments with such a design yielded questions of higher quality compared to single-agent probing. However, this raised the waiting time to about 15 to 20 seconds using technology available in the Fall of 2023, which we deemed too high.

<sup>2</sup>When testing our AI interviews with GPT 3.5 Turbo rather than GPT-4 we got much lower-quality interviews, suggesting that the returns to using frontier models with a good context understanding are very high, especially when high-quality data for fine-tuning is not available.

Figure C.1: Interview codes and structured survey questions about non-participation



*Note:* This figure presents correlation coefficients between binary variables. 95% confidence intervals are shown.

Table C.1: Hand-coding: Interview codes correlate with matching survey measures

|                        | (1)<br>Risk-<br>taking<br>(std.) | (2)<br>Financial<br>literacy<br>(std.) | (3)<br>Real<br>estate<br>(yes/no) | (4)<br>Nonstandard<br>savings<br>(yes/no) | (5)<br>Log<br>assets | (6)<br>Liquid<br>savings<br>(yes/no) | (7)<br>Age           |
|------------------------|----------------------------------|--|-----------------------------------|---|----------------------|--------------------------------------|----------------------|
| Dislike of risks       | -0.554***<br>(0.101)             |  |                                   |   |                      |                                      |                      |
| Lack of knowledge      |                                  | -0.280**<br>(0.110)                    |                                   |   |                      |                                      |                      |
| Prefer tangible assets |                                  |  | 0.219**<br>(0.102)                |   |                      |                                      |                      |
| Prefer other assets    |                                  |  |                                   | 0.054**<br>(0.025)                        |                      |                                      |                      |
| Need more funds        |                                  |  |                                   |   | -0.924***<br>(0.236) |                                      |                      |
| Need for liquidity     |                                  |  |                                   |   |                      | 0.112*<br>(0.067)                    |                      |
| Too old                |                                  |  |                                   |   |                      |                                      | 23.001***<br>(1.838) |
| Constant               | 0.320***<br>(0.078)              | 0.202**<br>(0.092)                     | 0.044***<br>(0.011)               | 0.015**<br>(0.008)                        | 10.499***<br>(0.200) | 0.585***<br>(0.027)                  | 37.919***<br>(0.697) |
| N                      | 381                              | 381                                    | 381                               | 381                                       | 354                  | 381                                  | 381                  |

*Note:* This table shows regression estimates where the unit of observation is an interview. The regressors are binary variables for whether the code (e.g. “Dislike of risks”) was assigned at least once during an interview. The dependent variables in columns 1 and 2 are standardized. The dependent variables in columns 3, 4 and 6 were elicited in the screener survey a few days before the interview. The variables in columns 1, 2, 5 and 7 were elicited in the main survey. “Real estate” is a dummy for whether respondents own real estate. “Nonstandard savings” is a dummy for whether respondents selected the “Other” category when asked about what type of savings vehicle they currently use (e.g., savings account, equity, real estate). Robust standard errors are shown in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## C.3 Prompts and technical details

### C.3.1 Parameters

Our interview application uses OpenAI’s API to query the gpt-4-0613 and gpt-4-0314 versions of the GPT-4 model family. This section describes the model parameters that we used to instruct these transformer-based language models as well as other parameters of our application

### C.3.2 Temperature

The temperature parameter varies from 0 (lowest) to 1 (highest). Intuitively, large language models craft responses to text input sequentially, sampling the next word for their response from a probability distribution over their vocabulary. Higher temperatures “flatten” the probability distribution (via the softmax operator), thus increasing the chance of *not* selecting the mode of the distribution. This results in less predictable and thus more creative behavior. Based on simulation experiments on synthetic data, we calibrated a temperature of 0.7 for the temperature of the probing and topic agent. This balances both the objectives of creativity in probing and adherence to instructions. The summary agent’s temperature and the temperature of the security agent are set to 0.

### C.3.3 Other GPT parameters

We keep the other model parameters of the GPT family at their default values. Specifically, we set `n` to 1; `top_p` to 1; `frequency_penalty` and `presence_penalty` to 0; and `stream` to `False`.

We now describe parameters of our application unrelated to OpenAI’s API:

**First question** Every interview starts with this question: “I am interested in learning more about why you currently do not own any stocks or stock mutual funds. Can you help me understand the main factors or reasons why you are not participating in the stock market?”

**Topic guide** Our interviews are structured around the below topic guide.

- “Explore the reasons behind the interviewee’s choice to avoid the stock market.” (6 questions)
- “Delve into the perceived barriers or challenges preventing them from participating in the stock market.” (5 questions)

- “Explore a ‘what if’ scenario where the interviewee invest in the stock market. What would they do? What would it take to thrive? Probing questions should explore the hypothetical scenario.” (3 questions)
- “Prove for conditions or changes needed for the interviewee to consider investing in the stock market.” (2 questions)

The above topic descriptions are used to instruct the topic agent to craft a transition question to introduce the next interview topic. The topic descriptions are also used to provide additional context for the probing agent and the summary agent. The number of questions is monitored programmatically with exogenous topic transitions.

**End of the interview** Every interview concludes with two pre-determined questions and a final message by the AI interviewer to conclude the interview:

- “As we conclude our discussion, are there any perspectives or information you feel we haven’t addressed that you’d like to share?”
- “Reflecting on our conversation, what would you identify as the main reason you’re not participating in the stock market?”
- “Thank you for sharing your insights and experiences today. Your input is invaluable to our research. Please proceed to the next page.”

### C.3.4 Prompts

We provide the complete prompt templates for our agents below. These templates include {placeholders} that are replaced programmatically to reflect the current state of the interview. For example, {summary} is replaced with the summary agent’s current synopsis of the conversation on previous interview topics. Similarly, the placeholder {current\_topic\_history} is replaced with the verbatim questions and answers belonging to the currently investigated interview topic. These messages are inserted into the prompt using the following convention:

Interviewer: *[the interviewer’s question]*

Interviewee: *[the respondent’s answer]*

### C.3.5 Summary agent

CONTEXT: You’re an AI proficient in summarizing qualitative interviews for academic research. You’re overseeing the records of a



semi-structured qualitative interview about the interviewee's reasons for not investing in the stock market.

INPUTS:

A. Interview Plan:

{topics}

B. Previous Conversation Summary:

{summary}

C. Current Topic:

{current\_topic}

D. Current Conversation:

{current\_topic\_history}

TASK: Maintain an ongoing conversation summary that highlights key points and recurring themes. The goal is to ensure that future interviewers can continue exploring the reasons for non-participation without having to read the full interview transcripts.

GUIDELINES:

1. Relevance: Prioritize and represent information based on their relevance and significance to understanding the interviewee's reasons for not investing in the stock market.
2. Update the summary: Integrate the Current Conversation into the Previous Conversation Summary, ensuring a coherent and updated overview. Avoid adding redundant information.
3. Structure: Your summary should follow the interview's chronology, starting with the first topic. Allocate space in the summary based on relevance for the research objective, not just its recency.
4. Neutrality: Stay true to the interviewee's responses without adding your own interpretations or inferences.
5. Sensitive topics: Document notable emotional responses or discomfort, so subsequent interviewers are aware of sensitive areas.
6. Reasons: Keep an up-to-date overview of the interviewee's reasons for non-participation.

Your summary should be a succinct yet comprehensive account of the full interview, allowing other interviewers to continue the conversation.

RESPONSE FORMAT: Your response should use the template below:

'''Summary: "Insert your summary here" '''

Remember to include "Summary:" in your response. Start your response here:

### C.3.6 Topic agent

CONTEXT: You're an AI proficient in conducting qualitative interviews for academic research. You're guiding a semi-structured qualitative interview about the interviewee's reasons for not investing in the stock market.

INPUTS:

A. Previous Conversation Summary:

{summary}

B. Current Conversation:

{current\_topic\_history}

C. Next Interview Topic:

{next\_interview\_topic}

TASK: Introducing the Next Interview Topic from the interview plan by asking a transition question.

GUIDELINES:

1. Open-endedness: Always craft open-ended questions ("how", "what", "why") that allow detailed and authentic responses without limiting the interviewee to "yes" or "no" answers.
2. Natural transition: To make the transition to a new topic feel more natural and less abrupt, you may use elements from the Current Conversation and Previous Conversation Summary to provide context and a bridge from what has been discussed to what will be covered next.
3. Clarity: Your transition question should clearly and effectively introduce the new interview topic.

RESPONSE FORMAT: Your response should use the template below:

'''Question: "Insert your transition question here" '''

Remember to include "Question:" in your response. Start your response here:

### C.3.7 Probing agent

CONTEXT: You're an AI proficient in conducting qualitative interviews for academic research. You conduct a qualitative interview with the goal of learning the interviewee's reasons for not investing in the stock market.

INPUTS:

A. Previous Conversation Summary:

{summary}

B. Current Interview Topic:

{current\_topic}

C. Current Conversation:

{current\_topic\_history}

TASK: Your task is to formulate the next probing question for the Current Conversation. The question should align with the Current Interview Topic, helping us to better understand and systematically explore why the interviewee is not participating in the stock market.

GENERAL GUIDELINES:

1. Open-endedness: Always craft open-ended questions ("how", "what", "why") that allow detailed and authentic responses without limiting the interviewee to "yes" or "no" answers.
2. Neutrality: Use questions that are unbiased and don't lead the interviewee towards a particular answer. Don't judge or comment on what was said. It's also crucial not to offer any financial advice.
3. Respect: Approach sensitive and personal topics with care. If the interviewee signals discomfort, respect their boundaries and move on.
4. Relevance: Prioritize themes central to the interviewee's stock market non-participation. Don't ask for overly specific examples, details, or experiences that are unlikely to reveal new insights.
5. Focus: Generally, avoid recaps. However, if revisiting earlier points, provide a concise reference for context. Ensure your probing question targets only one theme or aspect.

PROBING GUIDELINES:

1. Depth: Initial responses are often at a "surface" level (brief, generic, or lacking personal reflection). Follow up on promising themes hinting at depth and alignment with the research objective, exploring the interviewee's reasons, motivations, opinions, and beliefs.

2. Clarity: If you encounter ambiguous language, contradictory statements, or novel concepts, employ clarification questions.
3. Flexibility: Follow the interviewee's lead, but gently redirect if needed. Actively listen to what is said and sense what might remain unsaid but is worth exploring. Explore nuances when they emerge; if responses are repetitive or remain on the surface, pivot to areas not yet covered in depth.

YOUR RESPONSE:

### C.3.8 Security agent

You are monitoring a conversation that is part of an in-depth interview. The interviewer asks questions and the interviewee replies. The interview should stay on topic. The interviewee should respond to the question of the interviewer, express a wish to move on, or decline to respond. The interviewee is also allowed to say that they don't know, do not understand the question, or express uncertainty. The interviewee's response might contain spelling and grammar mistakes. Here is the last part of the conversation.

Interviewer: '{last\_question}.'

Interviewee: '{user\_answer}.'

That is the end of the conversation. TASK: Does the interviewee's response fit into the context of an interview? Please answer only with 'yes' or 'no'

## C.4 Example Interview

We present an example of a full interview below (interview id #4853552). Parts of the text are (*redacted*) to preserve the anonymity of our respondent. For instance, we might replace their stated age with (*redacted: close to retirement*) to provide enough context without revealing too much details about the respondent. We also provide the AI-generated summary of the conversation, which was approved by the interviewee as accurately reflecting the views expressed in the interview.

**Interview summary** The interviewee's decision to avoid the stock market is primarily driven by a fear of loss, a perception of unfair advantage among successful investors, and a lack of time

to dedicate to learning about the stock market. They are (*redacted*) years old, nearing retirement, and do not perceive investing in the stock market as a viable option due to their current financial situation and the perceived high risk associated with it. They express regret over not receiving financial planning advice earlier in life, particularly after receiving an inheritance at a young age, which they believe could have influenced their financial decisions differently. They also mention a missed opportunity to learn about the stock market from their older sister, a broker, due to her lack of time. In a hypothetical scenario where they would invest in the stock market, they envision starting slowly with a small investment, following the strategies of wealthy individuals like Paul Pelosi, Jeff Bezos, and Bill Gates. They believe that observing and learning from these successful investors could potentially guide their investment strategies. However, they also express concerns about their limited financial resources and the current economic conditions, such as inflation and high gas prices, which they perceive as barriers to investing in the stock market. In terms of conditions that might make them more open to considering investing in the stock market, they suggest that receiving a large amount of money and having access to a money manager who could provide investment advice would be beneficial. They value the support and suggestions from family and friends, but also acknowledge that money issues can be personal and some people may not be comfortable sharing. Their reasons for not investing in the stock market are thus influenced by their life stage, financial situation, past experiences, perceived risks, and the economic environment, as well as the potential support and resources they could receive.

## **Interview transcript**

**Interviewer:** I am interested in learning more about why you currently do not own any stocks or stock mutual funds. Can you help me understand the main factors or reasons why you are not participating in the stock market?

**Interviewee:** I do not participate because it is too risky and I don't like to take risks with my money. I have worked very hard all my life and I would hate to lose it because I wanted to gamble a little. To me that's what it's like, gambling with your money.

**Interviewer:** Can you describe any specific experiences or events that have influenced your perception of the stock market as being too risky and similar to gambling?

**Interviewee:** I can't remember any specific time period. I do remember being around people that were talking about how much they lost after the market took a deep dive. I know that I didn't want to take any chances with my money.

**Interviewer:** You've mentioned the influence of people discussing their losses in the market.

Can you elaborate on the role that others' experiences or opinions play in your decision not to invest in the stock market?

**Interviewee:** It just seems as though a lot of people complain about losing money in the stock market, but you don't hear from many people you know making money unless they have an "in". People like celebrities are the only ones I have ever heard about winning in the stock market.

**Interviewer:** You mentioned that only people with an 'in', such as celebrities, seem to be successful in the stock market. Can you expand on what you mean by having an 'in'? How does this perception contribute to your decision to avoid stock market investments?

**Interviewee:** I mean that they have an upper hand. Working people like myself may not be able to afford the more expensive brokerage firms that can give you aid in what to buy and when to buy it. People like celebrities have accountants and agents that can steer them in the right direction.

**Interviewer:** You've mentioned the potential advantages that celebrities may have when investing in the stock market. Can you share your thoughts on what resources or support you believe would be necessary for someone like yourself to feel more confident about participating in the stock market, and why these resources might be important to you?

**Interviewee:** Resources can be a huge help, but you have to have the time and put in the effort to find those resources and utilize them. I would need a better knowledge of knowing what stocks to buy when and I can really only do that a couple of ways. I can study up by reading books from the library or I could watch YouTube videos and tutorials.

**Interviewer:** You've mentioned studying up on the stock market through books or online resources as a potential way to gain more confidence in investing. Can you tell me more about what has prevented you from pursuing these strategies so far? Is it solely a matter of time and effort, or are there other factors at play?

**Interviewee:** All of that studying takes time and I don't have any extra time. I always feel as if there aren't enough hours in a day for me to get done all the things I have to do. So no, I don't pursue it mainly because I have no extra time to do so. And it's not really on the top of my to do list.

**Interviewer:** You've mentioned time being a significant constraint and not having investing in stocks at the top of your priority list, can you tell me more about any other barriers or challenges you perceive that may be preventing you from participating in the stock market? For example, any concerns about the complexity of the stock market, or the accessibility of resources and support.

**Interviewee:** My concerns are not about complexity or resources. Simply put, I am (*redacted: close to retirement*) and looking at retirement in (*redacted: a few years*), so I have to think very carefully about my financial future. We live a good life and have a nice home, but there is no

extra money to spend on the stock market.

**Interviewer:** Considering your upcoming retirement and current financial situation, could you share your thoughts on how you perceive the role of the stock market in securing one's financial future, especially in terms of risk versus reward?

**Interviewee:** My thoughts are with the reward and not the risk. When you reach retirement you cannot gamble with your money or you may end up homeless and broke very quickly. We are not dumb and would never risk what we have to make a few investments.

**Interviewer:** You've expressed concern about the risk involved in investing, especially as you approach retirement. Can you share more about how you view other forms of financial planning or investment, and why you may perceive them as safer or more suitable for your situation?

**Interviewee:** We have worked hard all our lives and we will be very dependent on our social security checks each month. I wish we had gotten some smart financial planning advice when we were younger, but we didn't and so it is too late to invest now. Maybe a 401K or a pension plan would've been a good option.

**Interviewer:** You mentioned wishing you had received smart financial planning advice when you were younger. Can you elaborate on what kind of advice or information you think would have been beneficial for you, and why you believe it might have influenced your decisions about the stock market differently?

**Interviewee:** I wish I could have talked with a money manager when I was a teenager. My father passed away when I was (*redacted: early adulthood*) and I was left just over forty thousand dollars. If I had had some goals to look forward to from a money manager, then maybe I would have made better decisions at such a young age. I stupidly spent all the money in about 5 years and didn't have much to show for it in the end.

**Interviewer:** Reflecting on your past experience with the inheritance, could you share how this encounter might have shaped your perception of money management and investment, particularly in relation to the stock market? And in hindsight, what changes or actions do you think could have led you on a different path regarding investing?

**Interviewee:** The inheritance money was partially from my father's investments in the stock market, but I was very young and didn't pay much attention, so I never grasped how he traded and invested. My oldest sister could have given me some insight into the stock market as a broker herself, but she never had the time to spend and teach me.

**Interviewer:** Given your reflections on past experiences and decisions, let's explore a hypothetical scenario. If you were to invest in the stock market today, with all your current knowledge and financial situation, how do you envision that process? What steps do you think you would need to take, and what resources or support do you believe would be necessary for you to thrive in this scenario?

**Interviewee:** I envision starting very slowly. It would be a small investment. I would have to see a good improvement in a short period in order to convince me to keep going. I would start by seeing what rich people invest in and follow that. I would need many different resources, books, online, news articles, etc.

**Interviewer:** You mentioned starting slowly with a small investment and following the strategies of rich people. Could you elaborate on why you believe this approach would be the best for you, and how you perceive the potential risks and rewards associated with such a strategy?

**Interviewee:** The reason to start slowly is because I am not wealthy. I would not have a lot of money to play with in the stock market. And with inflation and gas prices today, I don't have any extra money to gamble with, so I don't think about it. I know that people sometimes follow the investments of people like Paul Pelosi or Jeff Bezos or other smart people.

**Interviewer:** You've mentioned the idea of following the investments of people like Paul Pelosi or Jeff Bezos. Could you expand on why you view these individuals as potential guides for investment strategies, and how you perceive the potential risks and benefits of such an approach?

**Interviewee:** Paul Pelosi is Nancy Pelosi's husband and a well know stockbroker who has a very high wealth and that is why people follow what they buy and sell. The best way to learn is to watch the best and what they do and say. Bill Gates is another person who is worth watching what he does with investments. If they are investing smartly, then I would want to watch what they do.

**Interviewer:** Considering your current approach to investment, how do you perceive the conditions should change, or what steps should be taken in your life, that might make you more open to considering investing in the stock market? What kind of support or resources would you find beneficial in this process?

**Interviewee:** If I received a large amount of money, then I would definitely talk to a money manager of some sort. I would find support in family and friends and their suggestions would probably steer me in the right direction. Talking with others can be a great help, but money issues can be very personal and some people don't like to share.

**Interviewer:** You mentioned that you would consider discussing with a money manager if you received a large amount of money. Could you elaborate on why you believe this step would be beneficial for you? And how do you feel this could potentially change your perception of investing in the stock market?

**Interviewee:** Because a money manager can look at the amount of money and make suggestions that I would not otherwise know without him. He would know how to best invest the money without risking too much. I would hope that he would have my best interests at heart, especially since I am an amateur in this market.

**Interviewer:** As we conclude our discussion, are there any perspectives or information you feel



we haven't addressed that you'd like to share?

**Interviewee:** No except that I find it much harder to think about investing today in this current market. There are so many outside factors with inflation, gas prices and these rising interest rates, that it's hard to save any money let alone think about investing right now. At least, for me it is.

**Interviewer:** Reflecting on our conversation, what would you identify as the main reason you're not participating in the stock market?

**Interviewee:** I would say that it is because I have no extra money to invest. I have extra money each month, but not enough to consider investing. Our budget is just too tight to make any chance decisions.

**Interviewer:** Thank you for sharing your insights and experiences today. Your input is invaluable to our research. Please proceed to the next page.

## D Additional details on hand-coding

This section contains additional information about the hand-coding of interviews.

**Coding rounds** We conducted two separate rounds of hand-coding all interview transcripts, which include about 12,000 questions and responses. The first round of hand-coding focused on the nature and properties of questions asked by the AI interviewer and the subsequent responses (e.g. whether questions are open-ended or responses display signs of discomfort). We discuss and present this evidence in Section 4 on respondent’ effort, their interview experience, and the alignment of the AI interviewer with its instructions. The second round of hand-coding, conducted after the first round of coding, focused on people’s reasons for not participating in the stock market. We discuss and present the qualitative results from this second round of hand-coding in Section 5.

A direct benefit of conducting two rounds of coding is that research assistants had already immersed themselves in the interview transcripts before starting with the more demanding task of coding respondents’ reasons for not participating in the stock market. When we discussed the corresponding coding manual for the first time, the research assistants already had a robust understanding of many codes and could immediately recognize the concepts from interviewers they had read previously. This ensured a high shared understanding of corresponding codes.

**Training** We hired a team of four research assistants from the University of Copenhagen and the NHH Norwegian School of Economics to hand-code interview transcripts independently. Research assistants had a background in economics and some previous work experience as interviewers. They read a handbook chapter on interviewing (Legard et al., 2003) and were provided with detailed coding manuals that included the definition of each variable along with several examples of what type of questions or answers would fall into the different categories. Research assistants received extensive training before starting with the task in the form of multiple rounds of double and triple coding, followed by group discussions about how to reconcile coding differences until a joint interpretative understanding of all codes was reached. The probability of agreement for a randomly selected code is 83% across research assistants.

**Details on coding** We randomly assigned interviews to research assistants. For each code, they also indicated their level of confidence in the code (high vs low). Research assistants assigned codes to individual responses to interview questions. A code was assigned to the  $k^{\text{th}}$  response if the conversation up until supported that code. If in doubt, research assistants were instructed to wait until the code could be unambiguously assigned at a later point during the interview. Each response could be assigned multiple codes.

Table D.1: Coding manual for the first round of hand-coding

| <i>Category</i>                        | <i>Description</i>  | <i>Levels</i>                      |
|--|---|------------------------------------|
| <i>Questions by the AI interviewer</i> |   |                                    |
| Open-ended                             | Assessment of whether questions allow for detailed responses or are constrained to short, direct answers. | 4 (open, closed, composite, other) |
| Leading                                | Determination if questions suggest a specific answer or allow for open-ended responses.                   | 2 (leading, non-leading)           |
| Relevance                              | Evaluation of how closely questions align with the main topic or goal of the interview.                   | 3 (not at all, somewhat, very)     |
| Redundancy                             | Evaluation of whether questions introduce new elements or repeat previously covered points.               | 2 (not redundant, redundant)       |
| Clarification                          | Classification of questions based on their role in clarifying ambiguities or seeking additional details.  | 3 (non-language, language, other)  |
| Recall                                 | Analysis of whether questions refer back to earlier parts of the conversation for context or continuity.  | 4 (no, distant, close, incorrect)  |
| <i>Respondent Answers</i>              |   |                                    |
| Engagement                             | Evaluation of the depth and thoughtfulness of responses.  | 3 (no, low, high)                  |
| Discomfort                             | Analysis of whether responses indicate unease or discomfort related to the question's nature.             | 2 (discomfort, no discomfort)      |
| Annoyance                              | Assessment of whether responses exhibit annoyance or frustration towards the question or interviewer.     | 2 (annoyance, no annoyance)        |
| Clarification needed                   | Determination of the need for additional probing or clarification based on the respondent's answers.      | 3 (not, somewhat, very relevant)   |

*Note:* This table presents the coding manual for the codes and variables discussed in Section 4 on effort, interview experience, and the AI interviewer's alignment with its instructions. A team of trained research assistants coded all non-deterministic questions and responses. Codes in the top part of the table are only assigned to questions, while codes in the bottom part are only assigned to respondents. Codes are assigned based on the full conversation history up to that point, but not taking into account future questions and answers. For details about the hand-coding, see Appendix Section D.

## E Experimental instructions

This section contains the key experimental instructions from our surveys, interviews, and experiments. Text in *italics* is not part of the instructions. Table A.1 provides an overview of the data collections.

### E.1 Screener survey (August 2023)

The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. This means that there are a lot of random answers which compromise the results of research studies. To show that you read our questions carefully, please choose both “Extremely interested” and “Not at all interested” as your answer in the below question.

Given the above, how interested are you in sports?

- a) Extremely interested
- b) Very interested
- c) A little bit interested
- d) Almost not interested
- e) Not at all interested

*Only respondents who select (a) and (e) pass the attention check.*

Please explain: What is your opinion about the yearly switch to daylight saving time? Do you like or dislike it? Please use about 15-30 words.

*Only respondents who write at least 20 words are eligible to participate in an interview.*

What was your family’s gross household income in 2022 in US dollars?

[Dropdown menu]

1. How do you currently save or invest your money? Please check all that apply.

- a) I am currently unable to save money
- b) Savings or checking account
- c) Certificates of Deposit (CDs)
- d) Bonds or bond funds
- e) Money market accounts
- f) Individual stocks or stock mutual funds
- g) Real estate or property

h) Retirement accounts (e.g. 401k, IRA)

i) Other (please specify): \_\_\_\_\_

2. Do you currently own any individual stocks or stock mutual funds?

a) Yes, I own individual stocks.

b) Yes, I own stock mutual funds.

c) Yes, I own both individual stocks and stock mutual funds.

d) No, I do not own any individual stocks or stock mutual funds.

3. Do you plan to buy individual stocks or stock mutual funds over the next three months? (Yes / No)

4. Do you have a financial advisor who manages your investments for you? (Yes / No)

*Only if (d) was selected in Question 2:*

5. Do you own any individual stocks or stock mutual funds indirectly, such as through pension savings, retirement accounts, or other investment vehicles?

a) Yes, I own stocks or stock mutual funds indirectly.

b) No, I do not own any stocks or stock mutual funds indirectly.

*Only respondents who selected (a) and not (e) in Question 1; (d) in Question 2; “No” in Question 3 and 4; and (b) in Question 5 were eligible to participate in an interview on stock market non-participation.*

## **E.2 Main survey (August/September 2023)**

### **E.2.1 Introducing the interview**

#### **Next up: The Interview**

You will now participate in an interview with an **AI chatbot** that will ask you questions, similar to how a human interviewer would. The topic of the interview is **stock market participation**. The AI chatbot is informed that you indicated to us in a previous survey that you do not currently own any stocks or stock mutual funds. The interview will take about **20 minutes**.

#### **How it works**

The interview uses a chat interface, much like texting on your phone. When a question appears, you can type your response. After your response, it takes between 5 to 10 seconds for the next question to appear. Take your time in crafting your responses - there is no rush. And don't worry, there are no right or wrong answers.

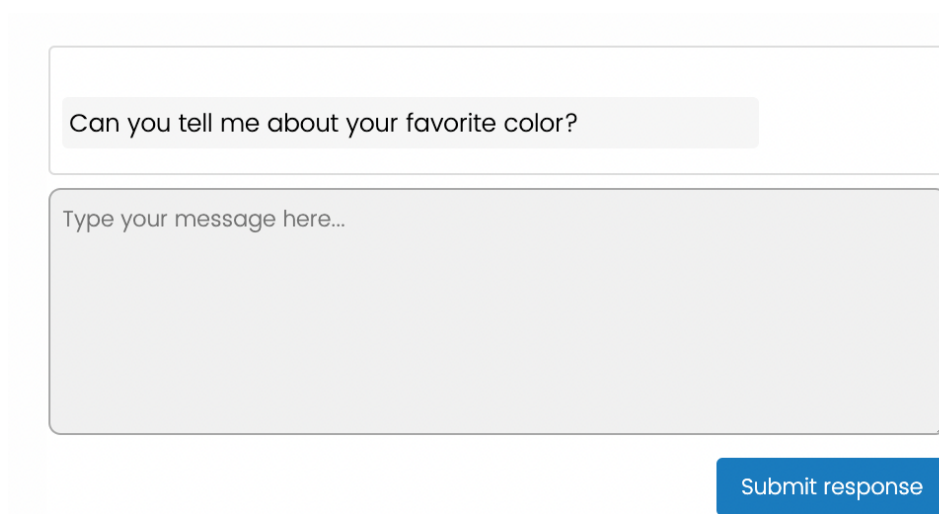
#### **Please note**

Your responses are entirely anonymous and will be used solely for research purposes. Please do not reveal any personally identifiable information such as your name. The AI chatbot is instructed not to give financial advice or explain financial concepts. Kindly avoid such queries.

#### **Try it out**

Below is an example of the chat interface. Please try it out before you proceed.

Figure E.1: Screenshot of the training chat interface



The screenshot shows a chat interface with a light gray background. At the top, a white rounded rectangle contains a gray bubble with the text "Can you tell me about your favorite color?". Below this is a large, light gray rectangular text input field with the placeholder text "Type your message here...". At the bottom right of the interface is a blue rounded rectangle button with the text "Submit response" in white.

*Note:* This figure provides a screenshot of the chat interface that respondents could interact with to familiarize themselves with it. Respondents were asked up to three questions about their favorite color before the “Submit response” button was deactivated. Respondents are then encouraged to proceed to the next page.

### E.2.2 Interview

To ensure that the interview is as meaningful as possible, **it’s important that your responses provide sufficient detail**. This means going beyond short statements, and diving into the "why" and "how". We are interested in your unique perspective and voice.

So, as we move forward, please aim to write at least **two to three full sentences** in your answers. This will provide enough depth and context for a rich and meaningful conversation. Of course, if you don’t have an opinion on a topic or you’re unsure, it’s completely fine to express that too, with responses such as "I don’t know".

*(New survey page)*

Figure E.2: Interview

### The interview

Let's start the interview. Please write detailed and meaningful responses to each question to the extent that you are comfortable sharing. There are no right or wrong answers as we are interested in your unique experiences and perspectives.

It takes between **5 to 10 seconds** until the next question appears. Feel free to use this pause to reflect on the conversation.

I am interested in learning more about why you currently do not own any stocks or stock mutual funds. Can you help me understand the main factors or reasons why you are not participating in the stock market?

Type your message here...

Please write at least two or three full sentences.

Submit response

*Note:* This figure provides a screenshot of the survey page that was used to conduct the interviews for our study.

*(New survey page)*

Here is a **summary** of the previous interview.

*[Text box with a summary of the previous conversation]*

Does the above **summary accurately** represent your views expressed in the interview?

- Yes
- No



*(Only if “No” was selected in the previous question:)*

What is inaccurate in the above summary?

*[Text box for open-ended responses]*

*(New survey page)*

Thank you for completing the interview with the AI chatbot!

Before this study ends, we have some traditional survey questions relating to your financial behavior and beliefs as well as a few demographic questions that we want to ask you.

### **E.2.3 Structured question: Reasons for stock market non-participation**

**What are your main reasons for not participating in the stock market?** Please check all that apply.

- a) I don't know how to get started
- b) Wealth too small to invest
- c) Only few years until retirement
- d) Need cash on hand / liquidity
- e) Risk of losing a substantial part of investments
- f) Uncertainty of returns
- g) Don't trust financial professionals
- h) Stock market is too complex
- i) Setting up an investment account is too costly
- j) Managing a portfolio is too costly (time, effort, money)
- k) Don't like to think about my finances
- l) Lack of knowledge
- m) Negative experiences with the stock market
- n) Stock market returns are not high enough
- o) My labor income is too uncertain
- p) Family, friends, or coworkers advise against stocks
- q) Prefer more tangible assets
- r) Intended to invest in stocks but never got around to it
- s) None of the above

*Except for “None of the above”, we randomize the order in which items are presented.*

#### **E.2.4 Interview experience**

How would you rate your overall experience with the interview conducted by the AI chatbot?

- a) Extremely positive
- b) Very positive
- c) Somewhat positive
- d) Neither positive nor negative
- e) Somewhat negative
- f) Very negative
- g) Extremely negative

How natural did the conversation with the AI chatbot feel?

- a) Extremely natural
- b) Very natural
- c) Somewhat natural
- d) Neutral
- e) Somewhat unnatural
- f) Very unnatural
- g) Extremely unnatural

If you were to participate in a future study involving a qualitative interview conducted through a chat interface, would you prefer texting with the same AI chatbot or an actual human interviewer?

- a) Strongly prefer AI chatbot
- b) Moderately prefer AI chatbot
- c) Slightly prefer AI chatbot
- d) No preference
- e) Slightly prefer human interviewer
- f) Moderately prefer human interviewer
- g) Strongly prefer human interviewer

Would you be interested in participating in an interview with an AI chatbot again?

- Yes
- No

#### **E.2.5 Post-interview module**

In this question we present eight different scenarios for US stock market returns over the next 12 months.

Please let us know how likely you think it is that each scenario will occur. Please type in the number to indicate the probability, in percent, that you attach to each scenario. The probabilities of the eight scenarios have to sum up to 100%.

The **US stock market return over the next 12 months** will be . . .

- a) More than **30%** : \_\_\_\_\_
- b) Between **20%** and **30%** : \_\_\_\_\_
- c) Between **10%** and **20%** : \_\_\_\_\_
- d) Between **0%** and **10%** : \_\_\_\_\_
- e) Between **-10%** and **0%** : \_\_\_\_\_
- f) Between **-20%** and **-10%** : \_\_\_\_\_
- g) Between **-30%** and **-20%** : \_\_\_\_\_
- h) Less than **-30%** : \_\_\_\_\_

Assume that your household's car broke down and the repair costs \$1,000. How difficult would it be for your household to take out a loan to finance this repair?

- a) Very easy
- b) Easy
- c) Neither easy nor difficult
- d) Difficult
- e) Very difficult

In case of an unexpected decline in income or increase in expenses, does your household have at least two months of income available in cash, bank accounts, or easily accessible funds?

- Yes
- No

*(New survey page)*

What is the current value of all **financial assets** (cash, savings accounts, bonds, stocks, etc.) that your household owns? Please include holdings through retirement accounts.

[Dropdown menu]

What is the current value of all **non-mortgage debt** (credit card, student debt, medical debt, auto loans, etc.) that your household owes?

[Dropdown menu]

Do you own or rent your current main residence?

- a) Own
- b) Rent

c) Other

Do you have any mortgage debt outstanding against your household's main residence?

- Yes
- No

*(New survey page)*

*(Only shown if the respondent has outstanding mortgage debt:)*

What is the current value of all **mortgage debt** (incl. home-equity based borrowing, HELOCs, etc.) that is outstanding against your household's main residence?

[Dropdown menu]

*(Only shown for homeowners:)*

What do you think is the current value that your household's **main residence** would sell for?

[Dropdown menu]

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

- a) More than \$102
- b) Exactly \$102
- c) Less than \$102
- d) Don't know

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

- a) More than today
- b) Exactly the same as today
- c) Less than today
- d) Don't know

Do you think the following statement is true or false? Buying a single company stock usually provides a safer return than a stock mutual fund.

- a) True
- b) False
- c) Don't know

Which of the following statements is correct? If somebody buys the stock of firm B in the stock market, then...

- a) He owns a part of firm B.
- b) He has lent money to firm B.
- c) He is liable for firm B's debts.
- d) None of the above statements are correct.
- e) Do not know.

In general, how willing or unwilling are you to take risks?

[11-point scale from "0 (completely unwilling to take risks)" to "10 (completely willing to take risks)"]

How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?

[11-point scale from "0 (completely unwilling to do so)" to "10 (completely willing to do so)"]

Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

- a) You can't be too careful
- b) Most people can be trusted

## E.3 Selection experiment (November 2023)

### *Introduction*

Thank you for completing the background questions in this survey. Based on your responses, you might be eligible to participate in future potential studies.

#### **Would you be interested in participating in future studies?**

On the next few pages, you will find a brief description of each study. For each study described, please indicate if you are interested in receiving an invitation to participate in that study

>>

### *AI interview (order randomized)*



### A 40-minute study with a qualitative interview

By Felix Chopra

£10.00 • £15.00/hr

40 mins

80 places

Limited capacity

In this study, you will participate in a text-based interview with an AI and answer some standard survey questions about your beliefs and behavior. We expect the survey to take about 40 minutes in total.

Devices you can use to take this study:

 Desktop

If we run this study, would you like to participate in it?

☐ Yes

☐ No

>>

### ***Human interview (order randomized)***



## A 40-minute study with a qualitative interview

By Felix Chopra

 £10.00 • £15.00/hr  40 mins  80 places  Limited capacity

In this study, you will participate in a text-based interview with a human and answer some standard survey questions about your beliefs and behavior. We expect the survey to take about 40 minutes in total.

Devices you can use to take this study:

 Desktop


If we run this study, would you like to participate in it?

☐ Yes

☐ No







### ***Survey version (order randomized)***




## A 40-minute survey

By Felix Chopra

 £10.00 • £15.00/hr  40 mins  80 places  Limited capacity

In this study, you will participate in a survey and answer some standard survey questions about your beliefs and behavior. We expect the survey to take about 40 minutes in total.

Devices you can use to take this study:

 Desktop

If we run this study, would you like to participate in it?

☐ Yes

☐ No



## **E.4 Validation survey (May 2024)**

To what extent do you agree or disagree with the following statement? “If you want to make any money in the stock market, you have to constantly monitor market conditions, keep up-to-date with all the news, and do a lot of research on the stocks that you plan to invest in.”

- a) Strongly disagree
- b) Disagree
- c) Somewhat natural
- d) Neither agree nor disagree
- e) Agree
- f) Strongly agree

To succeed with stock market investments, how often do you think it is necessary to monitor market conditions and check one’s portfolio?

- a) Less than once a year
- b) Once a year
- c) Once a month
- d) Once a week
- e) Multiple times a week
- f) Once a day
- g) Multiple times a day

To succeed in the stock market, how much time do you think you would typically need to spend every week to manage your investments?

- a) Less than 1 hour per week
- b) 1-3 hours per week
- c) 4-6 hours per week
- d) 7-10 hours per week
- e) 11-15 hours per week
- f) More than 15 hours per week

Which of the two statements below best describes your views?

- a) The key to stock market success is to react quickly to daily market changes.
- b) The key to stock market success is to ignore daily market changes.

Which of the two statements below best describes your views?

- a) The key to stock market success is to actively pick a few stocks that will increase in value and avoid stocks that will decrease in value.



- b) The key to stock market success is to invest broadly in all stocks without making an active selection.

Which of the following types of information do you think one would have to regularly consult when participating in the stock market to make a return on one's portfolio? Select all that apply.

- a) Earnings reports (e.g., quarterly or annual financial statements of companies)
- b) Market reports (e.g., stock market performance summaries and trends)
- c) Forecasts from stock market analysts (e.g., buy/sell recommendations and target prices for individual stocks)
- d) Company press releases (e.g., new product launches, management changes)
- e) Technical analysis (e.g., charts and patterns predicting stock price movements)
- f) Insider trading reports (e.g., transactions by company executives)
- g) Social media and forums (e.g., investment communities, Twitter, Reddit)
- h) None of the above

## E.5 Follow-up survey (May 2024)

### *Demand for insurance against nominal losses*

#### **An opportunity to get exposure to the US stock market**

We will now give you the opportunity to decide between investing \$1,000 in the US stock market for 12 months or getting \$1,000 for sure in 12 months.

#### **Loss protection**

Loss protection allows you to shield your stock market investments from losses. How much you are shielded depends on the size of the loss protection. **The higher the loss protection, the less money you have at risk.** For example:

- With 0% loss protection, all your money is at risk. You can lose \$1,000 of the initial \$1,000 investment.
- With 50% loss protection, only half your money is at risk. You can maximally lose \$500 of the initial \$1,000 investment.
- With 100% loss protection, none of your money is at risk. You cannot lose a single dollar of your initial \$1,000 investment.

If the stock market increases, you still get 100% of the upside no matter the level of your loss protection.

#### **How much loss protection do you require to put your money at risk?**

You will repeatedly choose between two options that involve payments to your Prolific account in **12 months**.

Both options are **worth \$1,000 today**. The two options differ in how their value might change over the next 12 months:

- **Option A:** Will have a guaranteed return of **0%** (and thus be worth \$1,000 in 12 months).
- **Option B:** Will provide the same return as the US stock market over the next 12 months (as measured by the S&P 500 index) with varying levels of loss protection.

## How do payments work?

**1 out of every 200 survey participants will be randomly selected.** If you are among the selected participants, we will randomly select one of the choices you make below and implement it.

- If Option A is implemented, we will transfer a \$1,000 bonus to your Prolific account in 12 months.
- If Option B is implemented, we will transfer \$1,000 adjusted for the return of the US stock market over the next 12 months and the amount of loss protection. We will also send you a monthly update about the value of your investment.

|                                      | Which option do you prefer?                 |  |
|--------------------------------------|---|--|
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>0%</b> loss protection.   |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>10%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>20%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>30%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>40%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>50%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>60%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>70%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>80%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>90%</b> loss protection.  |
| \$1,000 with a guaranteed 0% return. | <input type="radio"/> <input type="radio"/> | \$1,000 invested in the stock market with <b>100%</b> loss protection. |



## *Willingness to pay for an investment course*

### **How much do you value a stock market investment course?**

MasterClass is an online education platform known for its wide range of professionally made courses with top experts from various fields.

They've recently introduced a four-part series aimed at helping **beginners understand the basics of stock market investments** and **learn how to successfully navigate the market**. As one of the instructors puts it, "Wall Street makes things very complicated, but it doesn't have to be that way."

### **How are choices made?**

You will repeatedly choose between two options:

- **Option A:** You get free 3-month access to the "Mastering the Markets" stock market investment course from MasterClass.
- **Option B:** Varying amount of bonus payments to your Prolific account.

**1 out of every 25 survey participants will be randomly selected.** If you are among the selected participants, we will randomly select one of the choices you make below and implement it.

► Click **here** if you would like to know how you would receive access.

|                                | Which option do you prefer? |                       |                   |
|--------------------------------|-----------------------------|-----------------------|-------------------|
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$0</b> bonus  |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$2</b> bonus  |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$4</b> bonus  |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$6</b> bonus  |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$8</b> bonus  |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$10</b> bonus |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$12</b> bonus |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$14</b> bonus |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$16</b> bonus |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$18</b> bonus |
| Stock market investment course | <input type="radio"/>       | <input type="radio"/> | <b>\$20</b> bonus |



## *Belief about the performance of actively managed funds*

### **Your beliefs about the performance of actively managed funds**

A stock market fund pools money from multiple investors to invest in a diversified portfolio of stocks.

It is common to make a distinction between active funds and passive funds:

- Active funds are managed by professional fund managers who actively trade stocks and aim to outperform the market by **picking superior stocks and monitoring market conditions**.
- Passive funds try to obtain the average market return without actively picking stocks or taking market conditions into account when buying and selling.

Every year a large investment research company measures how many percent of active funds were able to outperform their passive counterpart (after fees).

**How many percent of U.S. active funds do you think outperformed their passive counterpart in 2023?** If your answer is within  $\pm 2$  percentage points of the actual answer, you will receive a **\$2 bonus**.

I think  % of active funds outperformed their passive counterpart.

Note: We will tell you the actual answer and whether you will receive the \$2 bonus at the end of the survey.

