Prompt Quality and LLM Memory: Soft Information in AI Advising*

Jing Huang Shumiao Ouyang

July 2026

Abstract

We provide a framework for human–AI communication that highlights prompt quality and LLM memory. We introduce preference uncertainty—capturing users' soft traits—into cheap talk (Crawford and Sobel, 1982), and model communication of user objective via optimal stopping with Brownian information flow. Prompt quality substitutes for memory; memory unambiguously helps only when the LLM's prior is aligned. Under limited memory, the investor prefers an LLM trained to be more "opinionated" than herself. We propose a novel empirical method to test these predictions. We simulate investor profiles from the Survey of Consumer Finances and conduct role-structured LLM advising experiments, benchmarked against portfolio questionnaires.

JEL Classification: D81, D82, D83, C61, C73, G11, C45, C63
Keywords: AI, LLMs, Dynamic Learning, Real Options, Financial Advising

*This paper was previously circulated under the title "Soft Information, Hard Decisions: AI Advising." Huang: Mays Business School, Texas A&M University, College Station, TX 77843, USA; email: jing.huang@tamu.edu. Ouyang: Saïd Business School, University of Oxford, Oxford OX1 1HP, United Kingdom; email: shumiao.ouyang@sbs.ox.ac.uk. For helpful comments, we thank Snehal Banerjee, Bo Bian, Juan Chen, Lin William Cong, Brandon Han, Peter Cziraki, Zhiguo He, Ziang Li, Dan Luo, Shane Miller, Hongwei Mo, Cecilia Parlatore, Jian Sun, Ruslan Sverchkov, Yizhou Xiao, Yao Zeng, Leifu Zhang, and seminar and conference participants at Texas A&M University (Mays), University of Michigan (Ross), the 2026 PHBS Finance Symposium, the OMI-SBS Conference on FinTech and Blockchain Economics, and the NBER-SAIF Conference on AI and Financial Markets. Shumiao Ouyang thanks Oxford RAST for their support, particularly Andreas Charisiadis and Shang Wang for their excellent research assistance. All errors are our own.

1 Introduction

As artificial intelligence becomes an economically pervasive technology, how effectively people interact with it is increasingly consequential. A neglected but central obstacle is that users often struggle to communicate what they want, especially when the task involves their own *soft traits*—preferences that are subjective and difficult to put into words, sometimes not even fully known to the user herself. However powerful the AI model, its performance can only be as good as the user's ability to articulate the problem, and much of the value of large language models (LLMs) lies in helping her do so through the back-and-forth of natural-language prompts.

This communication problem is governed by three features of the technology. The first is prompt quality. Digitizing soft traits is inherently lossy, since tone, hesitation, and context disappear in text, and non-expert users often do not know what information is relevant to provide in the first place, a "garbage-in, garbage-out" problem. The second is memory, the AI's capacity to synthesize information across an extended conversation. Memory determines the value of back-and-forth clarification, and despite rapid progress it remains costly.[1] The third is the prior the AI brings into the conversation, instilled by its training, which may or may not match the user it faces; how opinionated an AI's default stance should be is itself a design question.[2]

This paper develops a framework highlighting these three features, together with a new empirical method to test it. We introduce preference uncertainty, which captures soft information, into the canonical cheap talk setting, and model the user's prompt-based clarification of her objective as an optimal stopping problem with Brownian information flow. To fix ideas, we use financial advising as the running example, but the framework applies broadly when non-experts seek personalized guidance. While the LLM is free of the misaligned incentives common to human advisors (e.g. commission fees), it may misunderstand the objective of the investor, who then debiases the recommendation. We characterize the equilibrium in closed form and analyze how prompt quality, memory, and the AI's prior jointly determine the user's value. We then test the model's predictions where they can be evaluated against ground truth: role-structured, multi-turn LLM simulations of financial advising, with investor profiles drawn from the Survey of Consumer Finances and a rule-based allocation benchmark.

In the model, the investor has a quadratic loss utility function and faces two layers

[1] Inference over a longer context is computationally expensive, and agentic memory systems economize on context by consolidating, and thereby losing, information. See, e.g., Sam Altman, "people want memory."

[2] In practice this prior can be set at any layer of the stack: the training data mix, post-training, or a product-level system prompt; our simulations in Section 5 implement it through the system prompt.

of uncertainty. The first, uncertainty about the realizations of fundamental states of the world, is standard in the cheap talk literature. The second, which is novel in this paper, is uncertainty about the investor's own objective, which we call *preference uncertainty*, and captures soft information. For example, the investor does not know whether to target equity or fixed income (preference uncertainty); for each asset class, she does not observe the true fundamental realizations.

This preference can only be learned through prompt-based communication prior to receiving the AI advisor's nonverifiable recommendation. We model this communication as the investor's optimal stopping problem with Brownian information flow. We assume that the investor does not observe her true preference at prior: she knows the relevant information, such as income, tax status, and general risk attitudes in daily life, but does not know how these characteristics map to the investment objective. As she discusses her situation, the conversation generates public signals about her underlying preference that both she and the LLM learn from. She chooses when to stop and seek a final recommendation, weighing the informational value of continued learning against communication costs.

Prompt quality corresponds to signal precision and governs how much information each exchange conveys. AI memory, on the other hand, governs how much of that information the LLM absorbs. We model memory with a single parameter κ: for each signal, the LLM absorbs it with probability κ—reflecting the memory capacity—and otherwise misses it. This parameter indexes the rapid evolution of AI memory in practice, from stateless context windows to built-in within-session memory and agentic memory systems.[3] Our specification yields a simple relationship between the LLM's and the investor's beliefs: the update in the LLM's belief is exactly κ fraction of the investor's own update. A distinctive feature of our human–machine communication is that equilibrium inference exists only on the investor side; the LLM, by contrast, is a nonstrategic machine that does not infer the investor's belief from the gap its limited memory creates. We thus highlight a novel source that sustains asymmetric beliefs in dynamic settings, distinct from canonical sources such as heterogeneous priors (e.g. Adrian and Westerfield, 2009).

When the investor stops communication, she receives a recommendation based on the LLM's posterior and debiases it against her own belief—for instance, tilting the portfolio toward equities if she is more aggressive than the LLM perceives. The investor's payoff is high when she is more certain about her preference and the AI's belief is aligned. In addition, when the LLM's belief is more extreme, its recommendation is more informative about the asset fundamental realizations, thereby improving the investor's payoff.

[3]Yet memory remains costly: inference over a longer context is computationally expensive, and agentic memory systems economize on context only by consolidating, and thereby losing, information.

The investor's optimal communication policy follows a threshold rule: she continues interacting with the LLM until she becomes sufficiently confident about her preference, either above an upper threshold or below a lower one. For general LLM memory, we solve for the value function in closed form. For the corner cases of no memory and full memory plus aligned priors, the full equilibrium is characterized in closed form up to one unknown constant.

When does memory help? First, prompt quality can substitute for memory. When prompts are sufficiently informative, the investor's belief converges rapidly toward certainty, where even minimal memory lets the LLM reach certainty too, achieving the first best. Conversely, under worse prompt quality, the investor stops communication before certainty, incurring a loss even when LLM memory capacity is large. This matters practically: since memory is costly, precise prompting offers a low-cost alternative. Second, memory is unambiguously beneficial only when the LLM is trained with a prior aligned with the investor's. Under misalignment, a more retentive LLM may overshoot, amplifying rather than correcting disagreement, which makes a smaller memory a useful commitment device to improve the investor's value.

Alignment is a central concern in AI design and far cheaper to adjust than expanding memory; in our model, it corresponds to designing the LLM's prior. Under full memory, aligned training is optimal, since any initial misalignment is pure friction. However, this no longer holds under limited LLM memory. We show that with no LLM memory, the investor prefers a deliberately more "opinionated" AI that leans in the direction of her prior. The option value of learning tilts her preference away from exact alignment. Although her belief is a martingale so that the misalignment is as likely to shrink as to widen, the two directions differ in value: when her belief moves toward the AI's more extreme side, its recommendation becomes more informative about asset fundamentals, and this informational gain dominates.

We derive a set of testable implications that connect the model's primitives to observable advising behavior and outcomes. We focus on four hypotheses that our controlled LLM simulations can test directly. These are the predictions that turn on the advisor's information set and the dynamics of the conversation, which the simulations let us manipulate and observe in full. H1 posits that the primary value of interacting with a short-memory LLM advisor is that investors learn about their own initially uncertain preferences; the LLM's recommendation may remain generic, but dialogue helps investors reduce "preference uncertainty" and choose portfolios closer to their true objectives. H2 predicts that higher impatience, a higher intensity λ of the exogenous "impatience shock," induces earlier termination of the conversation, yielding less tailored portfolios. H3 predicts that giving the LLM persistent access to memory improves recommendation quality and narrows the gap with a human ad-

visor. H4 tests the model's prediction about optimal AI training: an investor is best served not by an advisor that mirrors her own beliefs, but by a deliberately "opinionated" one whose trained prior is more extreme than her own leaning—as discussed above, the more informative recommendation about fundamentals dominates the misalignment cost.

We complement the theory with prompt-based LLM simulations that generate multi-turn, role-structured conversations approximating real-world advisory exchanges. We draw a subsample from the 2022 public-use Consumer Finances Survey (SCF) to generate $n = 500$ hypothetical investor profiles for simulation (fixed seed). Investor "ground truth" comes from the 11-question Investor Questionnaire provided by Vanguard; we obtain each investor profile's rule-based "optimal" stock/bond allocation, and use these as benchmarks in evaluation. The advisor is implemented with OpenAI's GPT-5 under strict prompting. We compare a memoryless advisor, fed only the most recent Q&A and thus approximating fixed priors due to context-length limits; a memory-augmented advisor that uses the full chat history; and a full-information counterfactual that observes the complete profile. Conversations can end in two ways: an exogenous "impatience shock" terminates the conversation with probability 0.10 in each round, or the investor endogenously chooses to stop, weighing the costs and benefits of continuing in each round, which implements the model's optimal stopping rule.

Our findings, drawn from 2,500 conversations (500 profiles each undergoing 5 interactions), reveal several key patterns. First, supporting H1, the act of interacting itself accounts for most of the observed improvements: accuracy rises by 14.4 percentage points even before any specific recommendations, and by 17.6 points afterward. Each additional round of exchange adds about 0.77 points to accuracy, while every extra word contributes roughly 0.012 points. Second, in line with H2, ending conversations prematurely, outside the advisor's control, reduces recommendation accuracy by about 0.99 points, indicating a real penalty for impatience. Third, as H3 predicts, access to memory significantly boosts recommendation quality: the system performs best with full access to information, next best with full memory, and worst with no memory. Fourth, we confirm H4. On a subsample of 50 profiles from the tails of the allocation distribution, chosen to match the model's binary preference (30,250 observations), we assign the advisor's prior exogenously and trace out the empirical optimal-prior curve. It qualitatively matches the theory: the investor is best served by an advisor more opinionated than herself.

Literature. This paper contributes to several strands of literature.

AI in Economics and Finance. A rapidly growing theoretical literature studies AI adoption across economic settings. Ide and Talamàs (2025) embed AI into a knowledge hierarchy, where autonomous AI benefits the most knowledgeable individuals and co-pilot AI the least

knowledgeable. Zhong (2025) analyzes the optimal allocation of authority between humans and AI in multilayered decisions. Chen and Han (2026) show that AI adoption can exacerbate agency conflicts due to finer information production. Ferreira and Li (2025) study how CEOs' private access to AI advice substitutes for board advice.

In asset markets, recent contributions study AI algorithms as autonomous market participants. For example, Dou, Goldstein, Li, and Yang (2025) show that reinforcement-learning traders endogenously herd on common rules, with profits limited by rational investors' private-information advantage, and Banerjee and Szydlowski (2025) find that trading by Q-learning algorithms generates stochastic volatility and predictable returns.

Cheap talk and advising. Our paper complements the above applications by focusing on AI's role in providing advice to non-expert users, emphasizing communication frictions and AI-specific features: prompt quality, memory, and training. We build on the classic cheap talk model (Crawford and Sobel, 1982) and introduce preference uncertainty to capture soft information.[4] Soft information is traditionally less transferable than hard information (e.g., Liberti and Petersen, 2019; Stein, 2002), and recent advances in Big Data and AI have increasingly digitized it into hard data (He, Huang, and Parlatore, 2024). Our paper microfounds the process of digitizing soft information as dynamic learning.

Dynamic learning and real option stopping. We model prompt-based communication as continuous-time learning under Brownian information flow, akin to Daley and Green (2012). The investor's stopping problem is a real option on continued communication, connecting to option exercise under Bayesian learning (Grenadier and Malenko, 2010). In finance applications, agency and information frictions shape the timing of investment and intervention (Cong, Grenadier, and Hu, 2020; Gryglewicz and Hartman-Glaser, 2020), and Daley, Geelen, and Green (2024) study the dynamic interplay between information acquisition and bargaining in acquisitions. Closer to our setting, Grenadier, Malenko, and Malenko (2016) combine cheap talk with real options: an informed but biased agent advises the principal on the timing of option exercise; Malenko and Tsoy (2019) study cheap talk advising in auctions. Relatedly, Orlov, Skrzypacz, and Zryumov (2020) analyze the persuasion counterpart. The distinctive feature of our human–machine communication is one-sided inference: the non-strategic LLM does not infer the investor's belief, which sustains asymmetric beliefs in dynamic settings (unlike canonical sources such as heterogeneous priors, see Adrian and Westerfield, 2009).

Empirical contribution via LLM simulations. We complement the theory with prompt-based

[4]Cheap talk is also applied in corporate-governance settings where boards or proxy advisors offer non-binding recommendations (e.g., Levit and Malenko, 2011; Malenko and Malenko, 2019).

LLM simulations that generate multi-turn, role-structured conversations—a methodology for testing theoretical predictions about soft-information transmission and optimal stopping that would be difficult to examine in traditional laboratory settings. Recent work uses LLMs as economic agents to replicate behavioral regularities (Horton, 2023; Ouyang, Yun, and Zheng, 2024) and to simulate agents in strategic and market settings (Anand, Duarte, and Veitch, 2025; Lopez-Lira, 2025). Our two-sided LLM framework, in which both the advisor and the investor are simulated, is instead designed explicitly as a tool for theory testing.

The remainder of the paper proceeds as follows. Section 2 introduces the model. Section 3 constructs the equilibrium of AI advising and solves it in closed form. Section 4 analyzes prompt quality, AI memory, and optimal AI training. Section 5 presents our empirical analysis. Section 6 concludes.

2 The Model

Our key innovation is to formalize the friction of pinning down the user's objective with the LLM: this friction becomes critical when the task depends on her soft traits, so that effective advising requires eliciting such needs through prompts. We model the prompt-based communication of the user's objective as continuous-time learning about a binary state, followed by the cheap talk framework (Crawford and Sobel, 1982) for the LLM's recommendation and the user's choice. Unlike in canonical cheap talk, the LLM is not strategically misaligned; the user nonetheless debiases its recommendation, since the LLM misunderstands her objective.

While the model applies to a broad range of advisory applications, to fix ideas we use the context of financial advising throughout the paper.

2.1 Agents

This subsection introduces the model environment—agents' preferences and actions, taking their information as given. Section 2.2 then describes the communication process that determines the agents' beliefs.

2.1.1 Investor

Following the cheap talk framework, the investor has a quadratic loss utility function and chooses an action a to match the realization of a fundamental state θ—for example, aggressive buying when the fundamental is high. She does not observe θ and so seeks a recommendation from an advisor.

On top of this standard uncertainty about the asset fundamental, the investor is inexperienced and does not know how to formulate the investment problem: she knows her income, tax status, and general risk attitudes in daily life, but not how these characteristics map to investment decisions. To capture this uncertainty in the simplest way, we introduce a binary investment target $\omega \in \{1, 0\}$—$\omega = 1$ for equity and $\omega = 0$ for fixed income—to denote her underlying preference, and we write $\tilde{\theta}_\omega$ for the relevant fundamental as a random variable and θ_ω for its realization. When $\omega = 1$, for instance, the investor should target equity and invest more when the equity fundamental realization θ_1 is high. The investor's underlying objective is then

$$-\mathbb{E}(a - \tilde{\theta}_\omega)^2.$$

To summarize, there are two layers of uncertainty for the investor. Uncertainty about the fundamental realizations θ_1 and θ_0 is standard in the literature. Uncertainty about ω is new: we call it *preference uncertainty*, and it captures *soft information*; we use the two terms interchangeably.

We assume that ω is unobservable to all agents, including the investor, and can be learned only through communication conveying the investor's soft traits. Section 2.2 formalizes this communication as dynamic learning with Brownian information flows. We introduce $p,\ \hat{p} \in [0, 1]$ to denote, respectively, the investor's and the LLM's belief about ω:

$$p \equiv \mathbb{P}^i(\omega = 1), \quad \hat{p} \equiv \mathbb{P}^{LLM}(\omega = 1). \tag{1}$$

As standard in the literature, the fundamental realizations θ_1 and θ_0 are unobservable to the investor but observable to the advisor, and she seeks the advisor's recommendation which conveys information about them, as introduced in Section 2.1.2. We assume that $\tilde{\theta}_1$ and $\tilde{\theta}_0$ are independent normal random variables,

$$\tilde{\theta}_1 \sim N(\mu_1, 1), \quad \tilde{\theta}_0 \sim N(\mu_0, 1),$$

where μ_i is the mean of $\tilde{\theta}_i$ for $i \in \{1, 0\}$, and the common variance is normalized to 1.[5] We further denote $\Delta_\mu^2 \equiv (\mu_1 - \mu_0)^2$.

Given the investor's belief p about her objective, her expected utility from choosing action a is

$$U(a, p, \theta_1, \theta_0) = -p\,(a - \theta_1)^2 - (1 - p)\,(a - \theta_0)^2\,. \tag{2}$$

[5]The unit-variance normalization is without loss of generality: if variance is σ_θ^2, one can show that the agents' choices depend only on the ratio $\frac{(\mu_1 - \mu_0)^2}{\sigma_\theta^2}$ rather than the level σ_θ^2.

2.1.2 LLM

Seeking advice from LLMs is a central aspect of AI use: nearly half of ChatGPT usage involves practical guidance or seeking information (Chatterji, Cunningham, Deming, Hitzig, Ong, Shan, and Wadman, 2025). Unlike human advisors, LLMs are not strategically biased: developers optimize the model for general-purpose usage, so any strategic bias they introduce is likely orthogonal to the particular advisory task the user brings. However, the LLM serves the investor's interest only under its own understanding of what she wants, and efficient communication is challenging when the investor's soft traits are involved, for two reasons detailed below.

First, digitizing soft information inherently incurs information loss (Liberti and Petersen, 2019). This loss is exacerbated when the investor is a non-expert: she may include irrelevant or misleading details, yielding a "garbage in, garbage out" outcome.

Second, memory governs how much soft information the LLM can synthesize across the conversation—the capacity to integrate information over an extended exchange, not merely raw context length. The fundamental cost of memory is architectural: the Transformer underlying mainstream LLMs scales quadratically with input length, so maintaining long context is costly at inference (Liu et al., 2023; Wang and Sun, 2025). Memory has improved rapidly, through expanded context windows and built-in session memory,[6] but frictions remain: tokens are costly in practice; and in our setting of non-expert users, noisy prompt input further degrades the LLM's summarization of the chat history.

Recall that $\hat{p} \equiv \mathbb{P}^{LLM}(\omega = 1) \in [0, 1]$ is the LLM's understanding of the investor's latent preference ω, which is determined in the dynamic learning in Section 2.2 subject to the frictions discussed above. The LLM observes the fundamental realizations θ_1 and θ_0 (through its model training). Its payoff U^{LLM} is the investor's payoff but evaluated at the LLM's own belief $\hat{p}$:

$$U^{LLM}(a, \hat{p}, \theta_1, \theta_0) = U(a, \hat{p}, \theta_1, \theta_0) = -\hat{p}\,(a - \theta_1)^2 - (1 - \hat{p})\,(a - \theta_0)^2\,. \tag{3}$$

Maximizing this payoff, the LLM recommends

$$m = \hat{p}\theta_1 + (1 - \hat{p})\theta_0. \tag{4}$$

Therefore, the LLM is not strategically misaligned, but its recommendation is based on its own understanding of what the investor wants. Importantly, it does not observe the investor's belief p, which is also her soft information.

[6] Even when long contexts are available, retrieval and synthesis degrade as interactions accumulate.

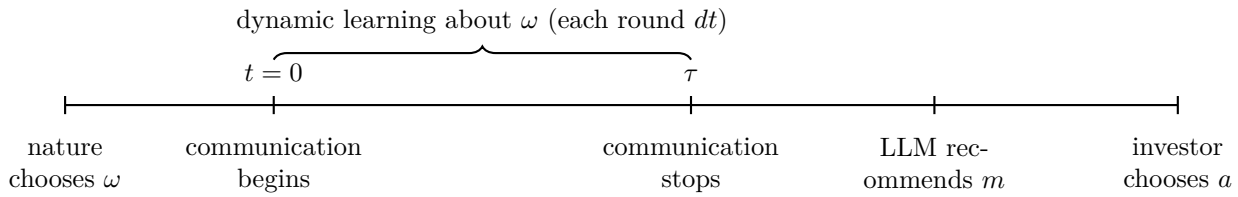


Figure 1: **Timeline.**

The investor understands that m is based on the LLM's belief $\hat{p}$, and therefore debiases it and learns about θ_1 and θ_0 from m at the same time. She chooses her optimal action $a(m, \hat{p})$ to maximize her expected utility (2) under her own belief p:

$$\begin{aligned} g(p, \hat{p}|m) &\equiv \max_a \ \mathbb{E}[U(a, p, \theta_1, \theta_0)|\, m, \hat{p}] \\ &= \max_a \left\{ -p\, \mathbb{E}[(a - \theta_1)^2 | m, \hat{p}] - (1 - p)\mathbb{E}[(a - \theta_0)^2 | m, \hat{p}] \right\}. \end{aligned} \tag{5}$$

For example, suppose $p = 0.5$ and $\hat{p} = 1$. The LLM recommends $m = \theta_1$—aggressive investment when the equity fundamental θ_1 is high. The investor understands that she is more conservative and adjusts her investment a toward the fixed-income distribution. At the same time, she perfectly learns the equity realization θ_1 from m, but obtains no information about the fixed-income realization θ_0.

Let $a^*(m, \hat{p})$ denote the optimal action which is the solution to (5). Given any pair of beliefs $(p, \hat{p})$, the investor's expected utility from seeking a recommendation is then

$$g(p, \hat{p}) \equiv \mathbb{E}[g(p, \hat{p}|m)] = \mathbb{E}[U(a^*(m, \hat{p}), p, \theta_1, \theta_0)]. \tag{6}$$

Because the investor first communicates about ω and seeks recommendation m when she stops, Eq. (6) is her stopping value in the communication process introduced next.

Remark 1. A natural question is why the LLM does not simply report θ_1 and θ_0 directly, which would allow the investor to achieve the first best. In practice, such reporting is infeasible given the "black-box" nature of AI outputs: "reporting θ_1 and θ_0" would correspond to disclosing intermediate reasoning steps, from which a typical user would be unable to infer θ_1 and θ_0 in a meaningful way.

2.1.3 Timeline

The timing of the model is summarized in Figure 1.

2.2 Communication with the LLM

Before seeking a recommendation m, the investor can initiate multiple rounds of prompt exchanges with the LLM to discuss her needs. We model this communication in continuous time over an infinite horizon starting at $t = 0$. The dynamic learning about ω is akin to Daley and Green (2012), and the LLM's memory is a distinctive feature of our model. Each round lasts an interval dt, over which a public signal ds_t about her true preference ω is released,

$$ds_t = \omega dt + \sigma dB_t, \tag{7}$$

where $B = \{B_t, \mathcal{F}_t, 0 \leq t \leq \infty\}$ is a standard Brownian motion on the canonical probability space $\{\Omega, \mathcal{F}, \mathbb{Q}\}$.

We use $\frac{1}{\sigma}$ to capture *prompt quality*: when it is high (low σ), the signal is more informative. Both parties observe the same signal history $\{s_\tau\}_{\tau \leq t}$, though how much the LLM absorbs from it depends on its *memory*, as discussed below. Importantly, because of this common history, the two state variables, the investor's belief p_t and the LLM's belief $\hat{p}_t$, can be reduced to one.

Communication may end exogenously through a Poisson shock of intensity λ, capturing the possibility that the investor becomes distracted. At each time t, as long as the shock has not yet arrived, she chooses whether to continue, paying a flow cost $c\,dt > 0$ while she does. If communication stops at time t, endogenously or exogenously, the investor receives recommendation m defined in (4) and her payoff is

$$-ct + g(p_t, \hat{p}_t),$$

where $g(p, \hat{p})$ is the stopping value in (6).

The cost of communication can be modeled in several ways. We adopt a Poisson shock to capture the practical reality that AI consultation is less engaging and easily disrupted. The flow cost c is also needed: Poisson stopping alone does not penalize continued communication, so without c the value function is convex and the investor never stops endogenously. Alternative formulations, such as a discount rate, yield qualitatively similar results.

2.2.1 Belief Updating

Investor's belief p. The investor holds a prior belief p_0 that $\omega = 1$. At time t, the investor's belief p_t is conditioned on the history of past communication, or the filtration $\mathcal{F}_t^s$ generated from $\{s_\tau, 0 \leq \tau \leq t\}$. Let f_t^ω denote the density of s_t conditional on ω, which is normally distributed with mean ωt and variance $\sigma^2 t$, i.e., $s_t \sim N(\omega t, \sigma^2 t)$. The investor's posterior

belief p_t satisfies the Bayes' rule

$$p_t = \frac{p_0 f_t^1(s_t)}{p_0 f_t^1(s_t) + (1-p_0) f_t^0(s_t)}. \tag{8}$$

This implies that the log-likelihood ratio $z_t \equiv \frac{p_t}{1-p_t}$ evolves according to a simple process, from which we derive the investor's belief process $\{p_t\}$. From Eq. (8):

$$z_t = z_0 + \ln \frac{f^1(s_t)}{f^0(s_t)} = z_0 + \frac{1}{\sigma^2}(s_t - \frac{t}{2}),$$

or

$$dz_t = \frac{1}{\sigma^2}\left(ds_t - \frac{1}{2}dt\right). \tag{9}$$

From the investor's perspective, the signal is released according to $ds_t = p_t dt + \sigma dB_t$, where $p_t = \frac{e^{z_t}}{1+e^{z_t}}$, so Eq. (9) becomes

$$dz_t = \frac{1}{\sigma^2}\left(\frac{e^{z_t}}{1+e^{z_t}} - \frac{1}{2}\right)dt + \frac{1}{\sigma}dB_t. \tag{10}$$

Using Ito's Lemma and $z_t(p_t) = \ln \frac{p_t}{1-p_t}$, we show that p_t evolves according to the standard binary state Brownian signal formula (for details, see Appendix OA.1),

$$dp_t = \frac{p_t(1-p_t)}{\sigma}dB_t. \tag{11}$$

There are a few things worth noting in (11). First, the belief process $\{p_t\}$ is a martingale without a drift term, reflecting prior consistency that the expectation of posterior beliefs equals the prior. Second, $\frac{1}{\sigma}$ is the signal-to-noise ratio and captures *prompt quality*, governing the speed at which the investor's preference ω is learned. Finally, the belief p_t is absorbing at $p_t = 0$ or 1, where the investor becomes certain about ω.

LLM's memory and belief $\hat{p}$. How much the LLM can learn from communication depends on its memory, a key capability of LLMs discussed in Section 2.1.2. As noted there, the fundamental cost of memory is architectural as the Transformer scales quadratically with input length. Hence, although memory has improved rapidly, a cost still manifests in practice as token fees and lossy summarization. In this paper, we use a simple parameter κ to index the ongoing development and the cost of memory.

Specifically, we assume that for each signal ds_t, the LLM absorbs the signal ds_t with probability $\kappa \in [0,1]$ and updates its belief $\hat{p}_t$; otherwise, with probability $1-\kappa$, it misses the signal and does not update. These missing events are independent across signals.

We show that the LLM's learning takes a simple form when expressed in the log-likelihood ratio, $\hat{z}_t \equiv \ln \frac{\hat{p}_t}{1-\hat{p}_t}$. If the LLM misses the signal ds_t, this is equivalent to receiving an alternative signal $d\tilde{s}_t$ with infinite noise; by Eq. (10), the belief update is then

$$d\hat{z}_t = \lim_{\tilde{\sigma}\to\infty} \frac{1}{\tilde{\sigma}^2}\left(d\tilde{s}_t - \frac{1}{2}dt\right) = 0.$$

If instead the LLM absorbs ds_t, it shares the same belief update as the investor, $d\hat{z}_t = dz_t$. Taken together, the LLM's belief update in the log-likelihood ratio is exactly a κ fraction of the investor's,[7]

$$d\hat{z}_t = \kappa dz_t. \tag{12}$$

Equivalently,

$$\hat{z}_t = \hat{z}_0 + \int_0^t d\hat{z}_s = \hat{z}_0 + \kappa(z_t - z_0). \tag{13}$$

Using $p(z) = \frac{e^z}{1+e^z}$ and $\hat{p}(\hat{z}) = \frac{e^{\hat{z}}}{1+e^{\hat{z}}}$, Eq. (13) implies that the LLM's belief $\hat{p}_t$ is a function of the investor's belief p_t,

$$\hat{p}_t\left(p_t\right) = \frac{\frac{\hat{p}_0}{1-\hat{p}_0}\left[\frac{(p_t/p_0)}{(1-p_t)/(1-p_0)}\right]^\kappa}{1+\frac{\hat{p}_0}{1-\hat{p}_0}\left[\frac{(p_t/p_0)}{(1-p_t)/(1-p_0)}\right]^\kappa}, \tag{14}$$

thereby reducing the two state variables to one. To summarize, the LLM *partially* learns from communication, and κ indexes its memory capacity.

Consider two corner cases as examples. When $\kappa = 0$, the LLM misses all signals and its belief stays fixed at its trained prior $\hat{p}_0$, so it recommends based on a "typical" customer profile that it is trained for. Under the continuous-time limit, no memory is equivalent to absorbing only the most recent prompt exchange ds_{t^-} (see Appendix OA.1). In practice, an LLM's knowledge resets between interactions; this case therefore maps to an LLM with no built-in session memory carrying the chat history into its working context. At the other corner, $\kappa = 1$, the LLM absorbs every signal, so its belief perfectly tracks the investor's whenever their priors match ($\hat{p}_0 = p_0$).

A distinctive feature of our model is that rational expectations hold only on the investor's side, not the LLM's, so equilibrium inference is one-sided. In standard dynamic communica-

[7] One can consider the following microfoundation. Let $\{X_i\}$ be iid Bernoulli random variables with $\mathbb{P}(X_i = 1) = \kappa$. Then

$$\hat{z}_t - \hat{z}_0 = \int_0^t \mathbf{1}_{X_s} dz_t = \lim_{n\to\infty} \sum_{s=0}^{tn-1}\left[\lim_{K\to\infty}\sum_{k=1}^{K}\mathbf{1}_{X_{sk}}\frac{1}{K}\left(z_{\frac{s+1}{n}} - z_{\frac{s}{n}}\right)\right] = \lim_{n\to\infty}\sum_{s=0}^{tn-1}\kappa\left(z_{\frac{s+1}{n}} - z_{\frac{s}{n}}\right) = \kappa\int_0^t dz_t,$$

where the second last equation applies the Law of Large Numbers on $\{X_{sk}\}$ for $k = 1, 2, ..., K$.

tion between rational agents, a human advisor would back out the investor's belief from the observed history and the structure of (14). The LLM, by contrast, is a *non-strategic machine*: it updates its belief mechanically through (14) and never infers p_t, so a gap between p_t and $\hat{p}_t$ persists throughout the interaction. We further discuss this point in Remark 2.

Remark 2 (One-sided inference)**.** The investor is rational and infers the LLM's belief, whereas the LLM, a non-strategic machine, does not internalize how its belief is affected by memory. This one-sided inference is a novel feature of our human–machine communication, and will be more relevant as AI's role in economic interactions expands. It also provides a new source of belief asymmetry in dynamic settings, distinct from canonical sources such as heterogeneous priors (e.g. Adrian and Westerfield, 2009).

We also rule out the investor directly communicating her belief p_t to the LLM: we interpret p_t as soft information that cannot be fully conveyed, and were she able to convey it partially, the setting would resemble our current model where the LLM partially learns.

2.3 Equilibrium Definition

Given the evolution of her belief $\{p_t\}$, the investor faces an optimal stopping problem

$$(SP) \quad \sup_{\tau \geq 0} \mathbb{E}^{\omega} \left\{ \int_0^{\tau} \lambda e^{-\lambda t} \left[-ct + g(p_t, \hat{p}_t(p_t)) \right] dt + e^{-\lambda \tau} \left[-c\tau + g(p_\tau, \hat{p}_\tau(p_\tau)) \right] \right\},$$

where her belief p_t and the LLM's belief $\hat{p}_t$ evolve according to Eq. (11) and (14) respectively. The investor chooses a stopping time τ, and the conversation continues at flow cost c until the earlier of τ and the arrival of an exogenous Poisson shock with intensity λ. If the shock arrives first at some time $t < \tau$, which occurs with density $\lambda e^{-\lambda t}$, the game ends exogenously, and the investor receives the stopping value $g(p_t, \hat{p}_t)$ net of the cost ct accrued up to that point; otherwise, with probability $e^{-\lambda \tau}$, no shock arrives by τ, and she stops voluntarily and receives $g(p_\tau, \hat{p}_\tau)$ net of the cost $c\tau$. The stopping value g is given by Eq. (6).

Definition 1 (Equilibrium)**.** When the investor consults the LLM, an equilibrium consists of the investor's stopping rule τ, the LLM's recommendation $m(\hat{p})$, and the investor's action $a(m, \hat{p})$ such that:

(i) Stopping time τ solves the investor's communication problem (SP);

(ii) The process $\{p_t\}$ is given by (11), and the process $\{\hat{p}_t\}$ is implicitly given by (13);

(iii) Given any belief $\hat{p} \in [0, 1]$, the LLM's recommendation $m(\hat{p})$ is given by Eq. (4), which solves $\max_m U^{LLM}(m, \hat{p}, \theta_1, \theta_0)$;

(iv) Given any pair of beliefs $(p, \hat{p})$, the investor's action $a(m, \hat{p})$ solves $\max_a \mathbb{E}[U(a, p, \theta_1, \theta_0)|m, \hat{p}]$.

3 Equilibrium and Model Implications

In this section, we construct the equilibrium and solve it in closed form.

3.1 Equilibrium Construction

We first take the agents' beliefs $(p, \hat{p})$ as given and solve for the investor's stopping value $g(p, \hat{p})$. Then we endogenize the agents' beliefs from the communication process, construct the equilibrium of interest and provide a closed-form value function.

3.1.1 Stopping Value

Recall that the investor's expected payoff when seeking a recommendation is defined in (6). Under the normal distributions of $\tilde{\theta}_1$ and $\tilde{\theta}_0$, it has a simple form as characterized in the following lemma.

Lemma 1. *(Stopping value) Given any pair of the investor's belief p and the LLM's belief $\hat{p}$, the investor's expected payoff when seeking a recommendation is*

$$g(p, \hat{p}) = -\frac{(p-\hat{p})^2}{\hat{p}^2 + (1-\hat{p})^2} - p(1-p)\Big(\Delta_\mu^2 + 2\Big). \tag{15}$$

The friction of AI advising arises from inefficient communication of the soft information ω. When the investor and the LLM are both certain and agree on ω, that is, $p = \hat{p} = 1$ or $p = \hat{p} = 0$, the investor attains the first-best payoff, $g(p, \hat{p}) = 0$.

More generally, Equation (15) decomposes her payoff into two costs. The first term captures the cost of *belief misalignment* between the investor and the LLM; notably, the LLM's belief enters the investor's payoff only through this channel. The LLM maximizes the investor's expected payoff, albeit under its own belief $\hat{p}$, so the investor still debiases. A more aligned LLM belief (a smaller $(p-\hat{p})^2$ in the numerator) reduces the cost of debiasing, and she most prefers a perfectly aligned LLM belief. Given a level of disagreement, as shown by the denominator $\hat{p}^2+(1-\hat{p})^2$ of this term, debiasing is more efficient when the LLM's belief $\hat{p}$ is more extreme because the recommendation is more informative about fundamentals θ_1, θ_0.

The second term of (15) captures the cost of *residual preference uncertainty*. This cost is larger when the investor herself is more uncertain about her preference (p closer to $\frac{1}{2}$) and when the stakes of misidentifying her objective are high, as captured by a large gap

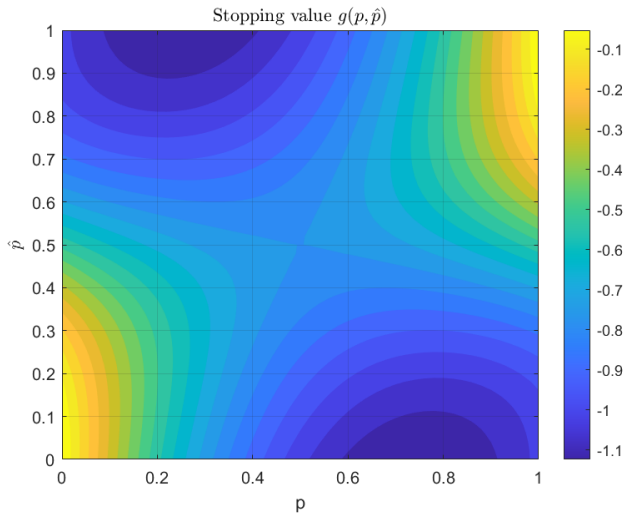


Figure 2: **Investor's stopping value $g(p, \hat{p})$.** Figure shows the investor's expected utility when seeking recommendation, $g(p, \hat{p})$, as a function of the investor's belief p in the horizontal axis and the LLM's belief $\hat{p}$ in the vertical axis. Parameters: $\Delta_\mu \equiv \mu_1 - \mu_0 = 1$.

Δ_μ between stock ($\tilde{\theta}_1$) and fixed income ($\tilde{\theta}_0$). Hence, the investor has an incentive to continue communication to become more certain about ω, trading off this benefit against the communication cost.

Figure 2 provides an illustration. The investor's payoff is higher when she is more certain about her preference, that is, as p approaches 0 or 1, and when the LLM's belief is more aligned with hers, along the 45-degree line of the figure. Notably, the investor's value is not lowest along the anti-diagonal $\hat{p} = 1 - p$, because the recommendation from an LLM with more extreme beliefs is more informative about fundamental realizations.

So far, we have taken the investor's and LLM's beliefs $(p, \hat{p})$ as given, but they are endogenous to the communication process that we examine next. Because agents share the same communication history, the LLM's belief $\hat{p}$ is a function of the investor's, allowing us to track only one state variable, p. With a slight abuse of notation, we write $g(p) \equiv g(p, \hat{p}(p))$ for the investor's stopping value expressed in the single argument p.

3.1.2 Value Function and Optimal Policy

The investor's belief p evolves according to (11). The LLM learns only partially, with the extent governed by its memory κ; its belief is a function of the investor's, $\hat{p} = \hat{p}(p; \kappa)$ in (14). The dynamic communication problem is stationary, and the investor's value function depends only on her belief p; let $V(p)$ denote her value. We conjecture that there exists a pair of thresholds $0 \leq \underline{p} < \overline{p} \leq 1$ such that the investor continues communicating with the LLM if and only if $p \in (\underline{p}, \overline{p})$.

Outside the continuation region ($p \notin (\underline{p}, \overline{p})$), the investor stops and seeks a recommendation from the LLM immediately. The value function then coincides with the stopping value $g(p, \hat{p})$ from Lemma 1. We retain both state variables here because, in the stopping region, the value of $\hat{p}$ depends on how $\hat{p}$ is reached. If agents start from within the continuation region through communication, then $\hat{p} = \hat{p}(p; \kappa)$ in (14) applies given the common history. If instead the investor starts in the stopping region, $p_0 \notin (\underline{p}, \overline{p})$, no communication takes place and the LLM's belief stays at its prior, $\hat{p} = \hat{p}_0$.

Inside the continuation region, $p \in (\underline{p}, \overline{p})$, the investor continues communication with the LLM over the infinitesimal interval $[t, t + dt]$, paying a flow cost $c\,dt$. Over this interval, with probability $\lambda\,dt$, an exogenous Poisson shock (e.g., the investor is distracted) ends communication and she receives the recommendation, yielding the stopping value $g(p)$. With the complementary probability $1-\lambda\,dt$, communication continues, her belief evolves to $p+dp$, and her continuation payoff is $\mathbb{E}_p[V(p + dp)]$. Combining these terms yields

$$V(p) = -c\,dt + \lambda\,dt\,g(p) + (1 - \lambda\,dt)\mathbb{E}_p[V(p + dp)]. \tag{16}$$

Expanding $V(p + dp)$ about p yields

$$V(p) \approx -cdt + \lambda dt g(p) + (1 - \lambda dt)\mathbb{E}_p\Big[V(p) + V'(p)\,dp + \tfrac{1}{2}V''(p)(dp)^2\Big].$$

Applying Itô's lemma along the diffusion in (11), and taking the limit $dt \to 0$ yields the linear second-order Hamilton–Jacobi–Bellman equation for the investor's value function in the continuation region:

$$-c + \lambda\left(g\left(p\right) - V\left(p\right)\right) + \frac{p^2\left(1-p\right)^2}{2\sigma^2}V''\left(p\right) = 0. \tag{17}$$

Without discounting, the investor requires a zero net return (the right-hand side of (17)) from continuing: the flow communication cost $-c$ and the Poisson loss $\lambda(g(p) - V(p))$, which replaces her continuation value with the stopping value at intensity λ, are offset by the option value of further belief updating in the last term.

We show in Appendix A.1 that any solution of (17) takes the form

$$V(p) = Q(p) + C_1 p^{\frac{1}{2}+\gamma}(1-p)^{\frac{1}{2}-\gamma} + C_2 p^{\frac{1}{2}-\gamma}(1-p)^{\frac{1}{2}+\gamma}, \tag{18}$$

where $\gamma \equiv \sqrt{2\sigma^2\lambda + \frac{1}{4}}$. $Q(p)$ is a particular solution to (17). The constants are pinned down

by the value-matching conditions,

$$V(\underline{p}) = g(\underline{p}), \tag{19}$$

$$V(\overline{p}) = g(\overline{p}). \tag{20}$$

Finally, the optimal stopping thresholds $\underline{p}$ and $\overline{p}$ satisfy two smooth pasting conditions, which are required such that the investor's strategy solves (SP). Specifically,

$$V'(\underline{p}) = g'(\underline{p}), \tag{21}$$

$$V'(\overline{p}) = g'(\overline{p}). \tag{22}$$

Recall that writing the stopping value in the single argument p, $g(p) \equiv g(p, \hat{p}(p;\kappa))$ requires a common communication history. In Online Appendix OA.2 we show that this holds in the neighborhood of the stopping thresholds, so that (21), (22) apply.

3.2 Equilibrium Characterization

We first characterize the equilibrium for a general memory capacity $\kappa \in [0,1]$ and express the value function in closed form. Then we show that for corner cases of κ, the full equilibrium can be characterized in closed form up to one unknown constant.

Proposition 1. *Given LLM memory $\kappa \in (0,1]$, the LLM's belief $\hat{p}_t(p;\kappa)$ is given in (14). When c is sufficiently small, there exist thresholds $0 \le \underline{p} < \overline{p} \le 1$ such that*

1. *When $p \notin (\underline{p}, \overline{p})$, the investor immediately seeks a recommendation and receives $V(p) = g(p, \hat{p})$ given in (15);*[8]

2. *When $p \in (\underline{p}, \overline{p})$, the investor continues to communicate, and her value function has the form*

$$V(p) = Q\Big(p, \hat{p}(p;\kappa)\Big) + C_1 U_1(p) + C_2 U_2(p), \tag{23}$$

where $U_1(p) \equiv p^{\frac{1}{2}+\gamma}(1-p)^{\frac{1}{2}-\gamma}$ and $U_2(p) \equiv p^{\frac{1}{2}-\gamma}(1-p)^{\frac{1}{2}+\gamma}$ with $\gamma \equiv \sqrt{2\sigma^2\lambda + \frac{1}{4}}$, and $Q\Big(p, \hat{p}(p;\kappa)\Big)$ is a particular solution constructed in Appendix A.1. The four unknowns $(C_1, C_2, \underline{p}, \overline{p})$ are jointly pinned down by the value-matching and smooth-pasting conditions (19), (20), (21), and (22).

[8] As discussed in Section 3.1.2, $\hat{p}(p;\kappa)$ in (14) applies only when agents communicate. The two-argument notation also covers the case where the investor starts in the stopping region $p_0 \notin [\underline{p}, \overline{p}]$.

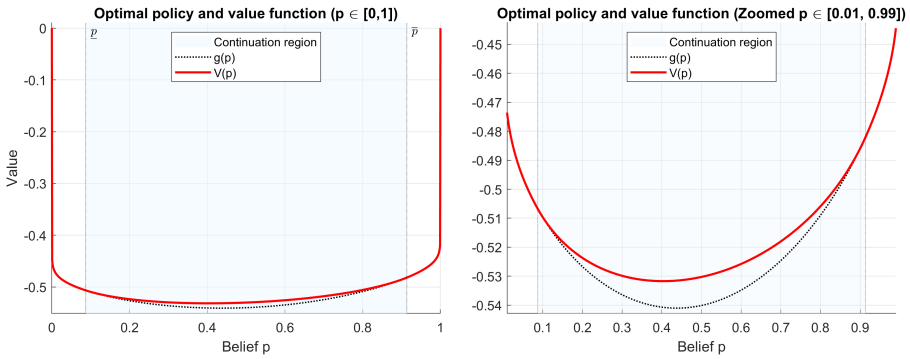


Figure 3: **Optimal policy and value function (extension $\kappa > 0$).** The left panel shows the full plot for $p \in [0, 1]$ and the right panel shows the zoomed plot $p \in (0.01, 0.99)$. The figure shows the continuation region $p \in (\underline{p}, \overline{p})$ (shaded area), the investor's value function $V(p)$ (red line) and stopping value $g(p) \equiv g(p, \hat{p}(p; \kappa))$ (black dotted). Parameters: $\Delta_\mu = 0.4,\ \sigma = 0.3,\ \lambda = 0.4,\ c = 0.1,\ \hat{z}_0 = \ln \frac{\hat{p}_0}{1-\hat{p}_0} = 0.03,\ \kappa = 0.019$.

The equilibrium requires the communication cost to be small relative to the option value of learning. The investor communicates when her belief is intermediate, where she is less certain about the underlying ω and the option value of learning is high.

Figure 3 illustrates the equilibrium in both panels; the right panel zooms in on the region $p \in (0.01, 0.99)$ because of the singularity of V near $p = 0$ and $p = 1$, which we explain below. The horizontal axis is the investor's belief p; the black dotted curve is her immediate stopping value $g(p) \equiv g(p, \hat{p}(p; \kappa))$,[9] and the red curve is her value function $V(p)$. The shaded region $(\underline{p}, \overline{p})$ is the continuation region: there $V(p)$ lies strictly above $g(p)$, so the investor keeps communicating. Outside it, the two curves coincide, $V(p) = g(p)$, and she seeks a recommendation immediately. The curves meet tangentially at $\underline{p}$ and $\overline{p}$, reflecting the value-matching and smooth-pasting conditions (19)–(22). In general, the continuation region is not symmetric around $\frac{1}{2}$ because the LLM's belief is tilted toward its trained prior $\hat{p}_0$.

A notable feature of memory ($\kappa > 0$) is the singularity of V near $p = 0$ and $p = 1$ in the left panel. As the investor becomes certain of her preference ($p \to 0$ or 1), communication becomes efficient even with minimal LLM memory, so the LLM's belief reaches certainty as well and she attains the first-best payoff $V = 0$. We show below that without memory ($\kappa = 0$) the investor's value never reaches the first best.

Corner cases of AI memory. In two corner cases of κ—the no-memory case $\kappa = 0$ and the full-memory case with an aligned prior, $\kappa = 1, \hat{p}_0 = p_0$—we characterize the full equilibrium in closed form up to a single unknown constant. Tractability here comes from a quadratic $g(p)$ and the resulting symmetry of the stopping thresholds around $\frac{1}{2}$.

[9] Here we specify $\hat{p} = \hat{p}(p; \kappa)$ in (14), assuming that p is reached from communication.

To see the symmetry, let $v(p) \equiv V(p) - g(p)$ denote the option value of waiting. Rearranging the HJB equation (17) gives

$$\lambda v\left(p\right) = \frac{p^2\left(1-p\right)^2}{2\sigma^2} \underbrace{g''\left(p\right)}_{\text{constant}} - c + \frac{p^2\left(1-p\right)^2}{2\sigma^2} v''\left(p\right).$$

Under Poisson impatience, her decision is independent of the level of $g(p)$, which is generally asymmetric around $\frac{1}{2}$ because of the LLM's tilted belief. Instead, it is governed solely by the marginal value of learning—the curvature $g''(p)$, which is constant in these corner cases. Together with the symmetric belief volatility $\frac{p^2(1-p)^2}{2\sigma^2}$, this renders the option value $v(p)$ symmetric, and hence the stopping thresholds symmetric around $\frac{1}{2}$.

Proposition 2. *Consider the two corner cases $\kappa = 0$ and $\{\kappa = 1, \hat{p}_0 = p_0\}$, and let*

$$q_2(\kappa) \equiv g''(p) = \begin{cases} \Delta_\mu^2 + 2 - \dfrac{1}{\hat{p}_0^2 + (1-\hat{p}_0)^2}, & \kappa = 0, \\ \Delta_\mu^2 + 2, & \kappa = 1,\ \hat{p}_0 = p_0 \end{cases} \tag{24}$$

denote the constant curvature of the stopping value $g(p) \equiv g(p, \hat{p}(p;\kappa))$. If $q_2(\kappa) > 16c\sigma^2$, there exists a unique pair of stopping thresholds $\underline{p}$ and $\overline{p}$ with $\underline{p} + \overline{p} = 1$, such that

1. *When $p \notin (\underline{p}, \overline{p})$, the investor seeks recommendation and receives $g(p) \equiv g(p, \hat{p}(p;\kappa))$ in (15);*

2. *When $p \in (\underline{p}, \overline{p})$, the investor communicates and her value is*

$$V(p) = g(p) - \frac{c}{\lambda} - \frac{q_2}{\gamma}\left[U_1(p)I_2(p) - U_2(p)I_1(p)\right] + C(\underline{p})\left[U_1(p) + U_2(p)\right], \tag{25}$$

where $U_1(p) \equiv p^{\frac{1}{2}+\gamma}(1-p)^{\frac{1}{2}-\gamma}$, $U_2(p) \equiv p^{\frac{1}{2}-\gamma}(1-p)^{\frac{1}{2}+\gamma}$, $I_1(p) \equiv \int_{\frac{1}{2}}^{p} U_1(s)\,ds$, $I_2(p) \equiv \int_{\frac{1}{2}}^{p} U_2(s)\,ds$, and $\gamma \equiv \sqrt{2\sigma^2\lambda + \frac{1}{4}}$. The constant $C(\underline{p})$ is given by

$$C(\underline{p}) = \frac{\frac{c}{\lambda} + \frac{q_2}{\gamma}\left[U_1\left(\underline{p}\right) I_2\left(\underline{p}\right) - U_2\left(\underline{p}\right) I_1\left(\underline{p}\right)\right]}{U_1\left(\underline{p}\right) + U_2\left(\underline{p}\right)}. \tag{26}$$

The optimal stopping threshold $\overline{p} = 1 - \underline{p}$ and $\underline{p}$ is uniquely determined by

$$\frac{d}{dp}\left\{\frac{\frac{c}{\lambda} + \frac{q_2}{\gamma}\left[U_1\left(p\right) I_2\left(p\right) - U_2\left(p\right) I_1\left(p\right)\right]}{U_1\left(p\right) + U_2\left(p\right)}\right\}\Bigg|_{p=\underline{p}} = 0, \quad \text{where } \underline{p} \in (0, \tfrac{1}{2}). \tag{27}$$

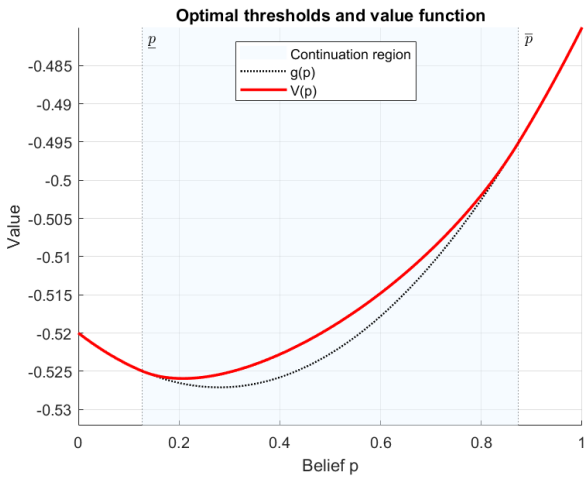


Figure 4: **Optimal policy and value function (no memory, $\kappa = 0$).** The figure shows the continuation region $p \in (\underline{p}, \overline{p})$ (shaded area), the investor's value function $V(p)$ (red line) and stopping value $g(p) \equiv g(p, \hat{p}_0)$ (black dotted). Parameters: $\Delta_\mu = 0.3,\ \sigma = 0.3,\ \lambda = 0.3,\ c = 0.035,\ \hat{p}_0 = 0.51$.

The condition $g''(p) = q_2 > 16c\sigma^2$ requires that the option value of learning is large relative to the communication cost, so that the continuation region exists. The quadratic form of $g(p)$ that drives the tractability arises broadly in economic applications: when an agent's payoff is linear in her information and she chooses an action based on it, the resulting value function is quadratic in that information (e.g., Weitzman, 1974).

Figure 4 illustrates the no-memory case $\kappa = 0$. Since the LLM's belief is fixed at $\hat{p}_t = \hat{p}_0 > \frac{1}{2}$, the investor's stopping value $g(p)$ and value function $V(p)$ are tilted upwards for $p > \frac{1}{2}$, where her belief p is more aligned with the LLM's. The optimal thresholds, however, remain symmetric around $p = \frac{1}{2}$, as established above. In the shaded continuation region, the investor's value function $V(p)$ lies above the immediate stopping payoff $g(p)$. Notably, even at $p = 0$ or $p = 1$ where the investor becomes certain, she still incurs a loss from debiasing the LLM's recommendation, because the LLM never learns under $\kappa = 0$.

In the other corner case of full memory with aligned prior, $\kappa = 1, \hat{p}_0 = p_0$, the LLM's belief tracks the investor's exactly, $\hat{p}(p; \kappa = 1) = p$, and there is no debiasing of AI recommendations. The remaining frictions reduce to standard communication costs: whether the investor is willing to wait at cost c for further learning at prompt informativeness $\frac{1}{\sigma}$.

4 AI Memory and Training

This section studies how AI's memory capacity and training affect the investor's value. We show that prompt quality can substitute for AI memory, and that more memory unambiguously benefits the investor only when the LLM's trained prior is aligned. Finally, we

characterize the optimal training of the LLM and show that, under limited memory, the investor prefers an LLM trained to be more "opinionated" than her own prior.

4.1 When Does AI Memory Help?

AI memory has improved rapidly in practice. The dominant form in current deployments is within-session memory, in which built-in functions grant explicit access to earlier exchanges (for example, in Anthropic's Claude and OpenAI's ChatGPT), and session capacity has grown significantly. In this part, we ask when more memory benefits the investor.

First, we argue that prompt quality, indexed by signal precision $\frac{1}{\sigma}$, can substitute for AI memory. Prompt precision governs how much information each exchange conveys, while AI memory governs how much of that information the LLM absorbs. The singularity of the investor's value at $p \to 0, 1$, as is visible in Figure 3, is the key behind this result. When prompts are sufficiently informative, the investor rapidly becomes certain about her objective, driving her belief toward the boundary $p \in \{0, 1\}$ and into the singularity region. Then even a small amount of LLM memory $\kappa > 0$ lets it reach the same certainty and attain the first best, as summarized by the following proposition.

Proposition 3. *For any AI memory $\kappa > 0$ and $\varepsilon > 0$, there exists $\underline{\sigma}(\kappa, \varepsilon) > 0$ such that if $\sigma < \underline{\sigma}(\kappa, \varepsilon)$, then $V(\underline{p}(\sigma, \kappa); \sigma, \kappa) > -\varepsilon$ and $V(\overline{p}(\sigma, \kappa); \sigma, \kappa) > -\varepsilon$. That is, the investor's value approaches the first best at the stopping thresholds.*

Even arbitrarily small AI memory is sufficient to approximate the first best when prompt quality is sufficiently high. Conversely, when prompts are imprecise (large σ), even full LLM memory fails to resolve the uncertainty about ω before the investor stops communication. Therefore, better prompt quality relaxes the requirement on AI memory. This result explains why AI succeeds in domains such as customer service, where tasks involve hard, verifiable information and short exchanges: prompts are effectively precise, so memory is barely needed.

Second, we ask whether better AI memory is always beneficial to the investor. The answer depends on the alignment between the LLM's and the investor's priors, as summarized by the following proposition.

Proposition 4. *Compare two memory capacities $\kappa_1 < \kappa_2$.*

1. *If the LLM and the investor share the same prior, $\hat{p}_0 = p_0$, then the investor's value is increasing in AI memory: $V(p; \kappa_1) < V(p; \kappa_2)$ for any p in the continuation region.*

2. *If $\hat{p}_0 \neq p_0$, stronger AI memory may hurt the investor: there exist $\hat{p}_0 \neq p_0$ such that $V(p_0; \kappa_1) > V(p_0; \kappa_2)$.*

Part 1 is intuitive: with aligned priors, a more retentive LLM tracks the investor's beliefs more closely and produces better-aligned recommendations.

When the LLM is trained with a misaligned prior, however, greater memory may hurt the investor, because its learning may overshoot. If the conversation moves the LLM's belief further from the investor's rather than toward it, a more retentive LLM amplifies the misalignment, whereas a shorter memory acts as a commitment device: by limiting how much the LLM learns, it keeps recommendations from drifting too far from her objective. Although belief movement is a martingale, making either direction equally likely, the wrong direction is more consequential when the investor's prior sits near the corresponding stopping threshold.

4.2 Optimal AI Training

Expanding AI memory is expensive, requiring costly advances in model architecture and infrastructure, and longer context is paid for at inference in every conversation. A cheaper and more readily adjustable lever is the AI's *trained prior*: how strong a view the model holds by default about the representative user it serves. This prior can be instilled in post-training or even a product-level system prompt, a one-time adjustment with negligible marginal cost. Indeed, how opinionated an AI assistant should be is itself a central question in AI design. In the model, this corresponds to designing the LLM's prior $\hat{p}_0$, given the representative user's prior p_0. Perhaps surprisingly, under limited memory, she may prefer an AI more opinionated than herself.

Aligning the LLM more closely with the investor's belief reduces the debiasing friction. Yet a more opinionated LLM, one whose belief is more extreme, issues a recommendation that is more informative about the underlying fundamentals. This trade-off is further complicated by the option value of learning: both agents' beliefs evolves endogenously as martingales during the conversation.

Under full memory, $\kappa = 1$, an aligned LLM prior is optimal to eliminate disagreement.

Proposition 5. *Suppose $\kappa = 1$. The investor's value is highest when the AI's prior is aligned,*

$$V(p; \hat{p}_0) \leq V(p; \hat{p}_0 = p_0), \quad \textit{for all } p \in (0, 1).$$

Now we consider the case of limited memory where disagreement can no longer be eliminated. For simplicity, we focus on the no-memory case $\kappa = 0$, where the LLM's belief is fixed at the training level $\hat{p}_0$. The optimal LLM belief is plotted in Figure 5 and Proposition 6 presents the formal results.

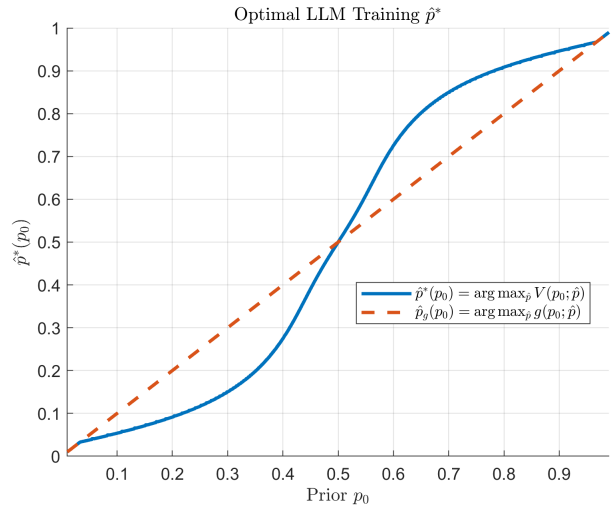


Figure 5: **Optimal LLM Training (no memory, $\kappa = 0$).** The figure plots the optimal training of the LLM, $\hat{p}^*$, for users with different prior beliefs p_0. The dashed red line plots a benchmark case of the optimal LLM training if the user does not communicate and immediately seeks a recommendation. Parameters: $\Delta_\mu = 0.7,\ \sigma = 0.45,\ \lambda = 0.3,\ c = 0.08,\ \kappa = 0.$

Proposition 6. *There exist threshold prior beliefs $\underline{p}_0$ and $\overline{p}_0$, satisfying $\underline{p}_0 + \overline{p}_0 = 1$, such that for a representative user with prior belief $p_0 \in [0,1]$, the optimal LLM training choice $\hat{p}^*(p_0) \equiv \arg\max_{\hat{p}} V(p_0;\hat{p}_0)$ satisfies*

$$\hat{p}^*(p_0) \begin{cases} > p_0, & \text{if } p_0 \in (0.5, \overline{p}_0), \\ < p_0, & \text{if } p_0 \in \left(\underline{p}_0, 0.5\right), \\ = p_0, & \text{if } p_0 = 0.5, \text{ or } p_0 \notin \left(\underline{p}_0, \overline{p}_0\right). \end{cases} \tag{28}$$

At these threshold priors, we have $\underline{p}(\hat{p} = \underline{p}_0) = \underline{p}_0$ and $\overline{p}(\hat{p} = \overline{p}_0) = \overline{p}_0$,

In the stopping region, the investor seeks a recommendation immediately and receives $g(p_0, \hat{p} = \hat{p}_0)$ given in (15). Without any option value of learning, her most preferred LLM training is exactly the aligned prior $\hat{p} = p_0$ that matches her belief, as illustrated by the red dashed line in Figure 5. There is no debiasing, and the benefit of a more opinionated LLM in giving more informative recommendations vanishes.

In the continuation region, the option value of learning tilts the investor's preference toward a more "opinionated" LLM, one that leans in the direction of her prior belief. For example, in Figure 5, when the investor's prior is $p_0 = 0.2$, she leans toward fixed income and her most preferred LLM prior is the *even more* conservative 0.1, rather than the aligned prior 0.2 she would choose were she to stop immediately. Because her belief is a martingale, it may drift either toward fixed income ($p \to 0$) or toward equity ($p \to 1$), shrinking or widening the misalignment with the LLM. The two directions are not symmetric, however:

when her belief drifts toward fixed income, the opinionated LLM delivers a more informative recommendation about the fixed-income fundamental θ_0, and this gain outweighs the loss when belief instead drifts toward equity. On net, she prefers the opinionated LLM. Last, when she is most uncertain, $p_0 = 0.5$, by symmetry she chooses an equally "confused" LLM, $\hat{p}^*(0.5) = 0.5$.

5 Empirical Analysis

In this section, we test model implications using prompt-based simulations in LLMs.

A remark on interpretation is in order. The simulations are controlled tests of the theory, not descriptions of human conversational behavior. The advisor side involves no simulation leap, since the object of interest is itself an LLM advisor. The investor side is simulated, which buys an advantage that no human-subject design can match: full control over each agent's information set, so that memory and information conditions vary one at a time while everything else is held fixed. Profile heterogeneity is disciplined by data from the Survey of Consumer Finances rather than generated freely by the model. The cost is that simulated investors need not capture all aspects of human responses; the results should be read as testing the model's information-based mechanisms, complementing rather than substituting for field evidence (Mo and Ouyang, 2025).

5.1 Testable Hypotheses

Building on the theoretical model, we derive testable hypotheses that connect the model's primitives to observable investor behavior and advice outcomes. Each hypothesis articulates a directional prediction, identifies the underlying mechanism from the theory, and outlines an empirical strategy for validation.

In a controlled setting, we can simulate the advising process with an LLM using prompt response logs and resulting portfolio recommendations to directly test several model predictions. In particular, these experiments allow us to control the agents' *information sets*. The theory also motivates predictions about advisor choice and adoption that require observational field data, such as records from platforms offering both human and AI advising; we leave these to future work.

H1 (Investor Learning and Decision Quality): The primary value of interacting with a memory-less LLM advisor stems from the investor clarifying their own initially uncertain preferences. A deeper interaction allows the investor to reduce their own "preference uncertainty," leading to a final investment decision that is better aligned with their true

objectives, even if the LLM's output remains generic.

H2 (Investor Impatience and Early Termination): Investors who face a higher opportunity cost of time break off LLM conversations sooner, which corresponds to a higher λ, and accept portfolios less tailored to their preferences.

To test this hypothesis, we examine the relationship between conversation length, which reflects the Poisson parameter λ, and the investor's final portfolio choice.

H3 (Memory Augmentation and Advisor Performance): Providing the LLM advisor with a form of persistent memory about past interactions (a larger κ) will improve its advice quality.

H4 (Optimal AI Training): An investor generally benefits from an AI advisor whose trained prior is more opinionated, meaning more extreme than her own prior belief, except when she is most uncertain about her type.

As developed in Section 4.2 (Proposition 6), a more extreme advisor issues recommendations that are more informative about the underlying fundamentals, and this gain dominates the cost of the greater misalignment.

5.2 Hypothesis Testing with LLM Simulations

Our prompt-based LLM simulations generate dynamic, multi-turn advisory conversations in which information acquisition and belief updating can be observed directly. This section describes how we construct investor profiles, implement an LLM-based advisor, and map the discrete simulation environment back to the continuous-time model.

5.2.1 Data and Investor Profiles

The baseline for "optimal" advice in our simulations comes from the Vanguard Investor Questionnaire, a widely used 11-question survey that generates personalized asset-allocation recommendations based on an individual's investment horizon, financial stability and risk tolerance.[10] For example, the questionnaire recommends investors with short time horizons to hold a smaller fraction of equity in their portfolio, while those with longer horizons to take on more risks. Similarly, it notes that a stable income stream allows investors to tolerate greater market volatility and therefore merits a higher equity weight.

In our empirical construction, we draw a subsample from the 2022 public-use Consumer Finances Survey (SCF) to generate $n = 500$ hypothetical investor profiles for simulation. To maintain representativeness, selection is conducted at the household level using probabilities

[10]The service is available at the website: https://investor.vanguard.com/tools-calculators/investor-questionnaire/questions. The questionnaires are listed in Online Appendix OA.9.

proportional to the final analysis weights of the survey, reflecting the dual-frame design of the SCF combining an area-probability sample and a list sample that oversamples high-income households. After selecting households, an imputed record per household is retained at random so that the final dataset contains exactly one observation per household while still reflecting imputation uncertainty. We fix the random seed prior to sampling to ensure that the draw is fully reproducible across runs. This design preserves fidelity to the underlying probability structure of the SCF while producing a compact and stable subsample suitable for repeated LLM experiments.

We then take each simulated profile and use it to complete the Vanguard questionnaire on the website, recording the recommended stock/bond allocation for each hypothetical investor. Each profile in our benchmark is thus defined by: (i) a list of question-answer pairs that can be expressed in natural language, which correspond to the investor's underlying preference ω, and (ii) the recommended allocation between equity vs fixed income, returned by the Vanguard algorithm, which maps to the ideal recommendation $\omega\theta_1 + (1-\omega)\theta_0$ without any frictions. In this way, we obtain a consistent set of simulated investors, grounded in established industry practice.

The questionnaire responses exhibit coherent internal correlations, strong among the risk-tolerance items and between the two time-horizon questions, consistent with realistic interdependencies among investor characteristics rather than random generation. Online Appendix OA.7 reports the full correlation matrix and discussion.

5.2.2 Conversation with the LLM Advisor

For each investor profile we simulate a conversation between the investor and an LLM-based advisor. The LLM we use is the then-state-of-the-art OpenAI GPT-5, accessed via an API at temperature 0.75. Conversations are orchestrated through system and developer messages that clearly specify each agent's role and information set. The investor's system prompt states that she is preparing to set up an investment portfolio and will interact with a financial advisor. It also includes the full text of her questionnaire profile. As explained below, the investor chats with the LLM based on this questionnaire but makes portfolio decisions without it. The advisor's system prompt specifies that it is a financial advisor whose objective is to determine the client's optimal allocation between equities and bonds. It is instructed to remain in information-gathering mode and to ask diagnostic questions about financial situation, investment experience, time horizon and risk preferences until the conversation ends. Importantly, the system prompt explicitly forbids the advisor from providing any recommendations before being told to do so, ensuring that all intermediate messages consist only of questions.

The interaction unfolds in discrete rounds:

1. **Eliciting the prior.** Before any questions are asked, the investor reports her prior preferred equity allocation. She is prompted to return her intended equity percentage as `JSON`, and this value is logged. This step parallels the initial belief p_0 in our theoretical model.

2. **Question–answer loop.** A developer message instructs the advisor to "ask your client one question that will help you identify their optimal split between equities and bonds." The advisor generates a question, which we append to the conversation. The investor's developer prompt tells her to answer concisely. She answers truthfully *according to her profile questionnaires*, and the answer is appended to both the investor and advisor message lists. In subsequent rounds the developer instructs the advisor to "ask your client another question," so that exactly one question is asked per round. The conversation thus alternates between one question from the advisor and one answer from the investor.

3. **Termination decision.** After each round, the simulation consults the pre-drawn termination schedule (described below) to determine whether the conversation should stop. If the investor terminates, the question–answer loop ends; otherwise the advisor is prompted to ask another question.

4. **Recommendations.** Upon termination the advisor is asked to provide equity-allocation recommendations in two ways. First, to simulate the memoryless case, the advisor receives only the most recent question and answer and is instructed to return the optimal equity allocation for the individual. Second, to simulate memory, the advisor receives the entire chat history (including all simulated questions and answers) and is instructed to produce an optimal equity allocation. In both cases, the advisor's output must be a `JSON` snippet specifying the recommended equity percentage; no narrative explanation is allowed. For the full-information advisor, the advisor is given the complete investor profile up front and asked once for the optimal equity allocation.

5. **Final decision.** Finally, the investor is told that "the advisor recommends an equity allocation of $x\,\%$" and is prompted to choose her final allocation. Like the prior, the final allocation is returned as `JSON`. This step captures the investor's own action in the theoretical model after observing the advisor's message.

Throughout the conversation, we log every prompt and response. The strict formatting (e.g., `JSON` outputs, one question per round, no unsolicited recommendations) reflects an

attempt to control the LLM's behavior and reduce hallucinations. The advisor's inability to update its belief in the memoryless condition comes from receiving only a single question–answer pair as input; this effectively resets its context window at every termination event.

We implement three variants of the LLM advisor to isolate the effect of memory.

The *advisor without memory*, motivated by the architecture of current LLMs, cannot recall previous signals. After each prompt, the LLM uses only the most recent question–answer pair as context to recommend an allocation. This design captures the short context windows of Transformer models, whose self-attention complexity scales quadratically in the input length and cannot reliably capture the entire query history. Wang and Sun (2025) show that even when longer contexts are available, retrieval accuracy can deteriorate rapidly due to interference from earlier inputs; the probability of recalling the most recent key–value pair declines log-linearly as similar distractors accumulate. Our "no-memory" LLM therefore approximates our baseline model that the advisor updates its prior based on only the last signal, $\hat{p}_t = \mathbb{E}[\omega = 1 | ds_{t^-}]$. (In the model, the advisor's belief remains fixed at its trained prior under the continuous time limit.)

The *advisor with memory* is instead fed the entire conversation history as context. It has access to previous answers and can refine its recommendation as it learns more about the investor. This scenario is our best approximation of a human advisor free of conflicts of interest with a transcript of the conversation. This aligns with our model extension with $\kappa = 1$: both the investor and the advisor update their belief based on the full history of signals $\{s_t\}$ before termination; if they share the same prior, then $\hat{p}_t = p_t$.

Finally, the *advisor with full information* is a counterfactual that receives the investor's entire questionnaire profile at once. It directly observes ω and effectively faces no preference uncertainty. This case can also mimic a human advisor who perfectly elicits soft information instantaneously, with no bias. The full-information advisor's recommendation provides an upper bound on achievable accuracy. Note that the advisor's recommendation may still fall short of the frictionless Vanguard benchmark because of noise in the fundamentals. While the model assumes that advisors observe the fundamental realizations θ_1 and θ_0, in practice, the LLM only observes a noisy signal of these fundamentals.

5.2.3 Termination and Cost of Communication

We cap the interaction at a maximum of eleven rounds to mirror the finite length of the Vanguard questionnaire. However, an investor does not necessarily complete all eleven rounds, as the dialogue can end sooner based on one of two mechanisms: exogenous or endogenous termination. To model impatience or external factors that cut a conversation short, we introduce an exogenous termination probability of 0.10 per round. Before the simulation begins,

we conduct a Bernoulli draw for each round to determine if an "impatience shock" occurs. If a shock is scheduled for a given round, the conversation stops automatically after that round's question and answer are complete. If no exogenous shock occurs, the investor agent makes an active, endogenous decision to continue or stop. After answering the advisor's question, the investor is prompted with a decision frame that explicitly asks them to weigh the costs and benefits of more interaction: *"Your time is valuable, so each round of communication with the advisor carries a cost. You should choose to continue interacting with the advisor if and only if your expected informational gain from an additional round of communication exceeds your subjective cost of communication. Would you like to continue or terminate the conversation...?"* This prompt directly simulates the optimal stopping trade-off.

This dual-termination mechanism serves as a discrete analogue of the continuous-time optimal stopping problem in our theoretical model. Each question-answer round in the simulation corresponds to a small increment of time dt. The exogenous termination probability models the Poisson shock intensity λ, while the cost of continuing the conversation mirrors the flow cost c. The endogenous decision to terminate corresponds to the investor's optimal stopping thresholds $\underline{p}, \overline{p}$.

5.3 Empirical Results and Hypothesis Testing

We now present the empirical results from our LLM simulations to test the three main hypotheses derived from the theoretical model. Our analysis uses data from 500 simulated investor profiles, each with 5 iterations, yielding 2,500 total observations. The results provide strong support for the model's predictions about the role of investor learning, the impact of communication costs, and the benefits of memory augmentation in AI advising.

5.3.1 Testing H1: Investor Learning and Decision Quality

The first hypothesis posits that LLM advising help investor clarify their own preferences. Table 1 presents the regression results testing this hypothesis. Across measures, the data show notable gains in investment accuracy that arise from the advisory process itself.

Investors improved significantly even before receiving any recommendations. As shown in Columns (1) and (2), accuracy increased by 14.4 percentage points before any advice, reaching 17.6 percentage points after the advisor's input. The 3.2 point boost from recommendations, while meaningful, suggests that much of the learning comes from the interactive process itself rather than solely from the advice received.

Interaction intensity is vital for accuracy improvement. Results from Columns (3) to (4) show that each additional round between advisor and investor increased interim accuracy

Table 1: Investor Learning and Decision Quality

This table reports regression results examining factors influencing the accuracy of investors' investment choices, measured at interim and final stages, as well as accuracy improvements. Columns (1)-(2) display improvements in investment choice accuracy at pre- and post-recommendation stages. Column (3) specifically illustrates investors' interim investment choice accuracy, while Columns (4)-(8) present investors' final investment choice accuracy. Accuracy here is defined as 100 minus the deviation from the optimal portfolio allocation, expressed in percentage points. Independent variables include the number of interaction rounds (# Rounds, Total # Rounds), total words exchanged, and separately, the number of words contributed by the advisor and the investor. Standard errors, clustered by investor profile, are shown in parentheses. Profile fixed effects are incorporated in Columns (3)-(8) to control for investor-specific characteristics. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively.

	Accuracy Improvement			Investor's Accuracy				
	Pre-Rec	Post-Rec	Interim			Final		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# Rounds			0.727***					
			(0.079)					
Total # Rounds				0.771***				
				(0.120)				
Total # Words					0.012***			
					(0.002)			
Total # Advisor's Words						0.020***		0.029***
						(0.003)		(0.006)
Total # Investor's Words							0.026***	-0.018*
							(0.006)	(0.009)
Constant	14.4***	17.6***						
	(0.717)	(0.887)		—	—	—	—	—
Profile FE	N	N	Y	Y	Y	Y	Y	Y
Clustered by Profile	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,500	2,500	8,673	2,500	2,500	2,500	2,500	2,500
Adjusted R^2	-	-	0.674	0.903	0.903	0.903	0.902	0.903
Note:								$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

by 0.727 points and final accuracy by 0.771 points. This implies iterative exchanges help investors gradually refine their understanding of risk and preferences through structured dialogue.

The actual words exchanged drive learning as well. Column (5) shows that every additional word spoken by either party increases accuracy by 0.012 points. Columns (6) and (7) reveal interesting dynamics when examining advisor and investor contributions separately. Advisor words improve accuracy by 0.020 points per word, while investor words show an even stronger effect at 0.026 points per word when considered in isolation. This suggests that articulating and reflecting on one's own beliefs is particularly valuable for preference clarification.

However, the multivariate model in Column (8) reveals important nuance. When both advisor and investor word counts are included simultaneously, advisor words show an even stronger positive effect, while investor words surprisingly become negative though only

marginally significant. This reversal suggests potential multicollinearity or that once the advisor's structuring role is accounted for, excessive investor verbalization may indicate confusion rather than clarity.

The consistently strong effects of structured dialogue, combined with the substantial but not dominant role of final recommendations, suggest the primary benefit of LLM advising lies in facilitating investor self-discovery through guided interaction.

5.3.2 Testing H2: Investor Impatience and Early Termination

The second hypothesis examines whether investor impatience, manifested through early conversation termination, undermines the advisor's ability to provide accurate recommendations. Table 2 presents regression results analyzing the factors that influence the accuracy of the AI advisor's investment recommendations.

Table 2: Investor Impatience and Early Termination

This table reports regression results examining factors influencing the accuracy of the AI advisor's investment recommendations. All columns present the accuracy of the advisor's recommendations, measured as 100 minus the deviation from the optimal portfolio allocation, expressed in percentage points. Columns (1)-(4) examine individual factors in isolation, while Column (5) presents a specification including both advisor and investor word counts. Column (6) examines the effect of exogenous termination. Independent variables include the total number of interaction rounds (Total # Rounds), total words exchanged, separately the number of words contributed by the advisor and the investor, and an indicator for whether the conversation was terminated exogenously rather than reaching a natural conclusion. Standard errors, clustered by investor profile, are shown in parentheses. Profile fixed effects are incorporated in all columns to control for investor-specific characteristics. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively.

	Accuracy of Advisor's Recommendation					
	(1)	(2)	(3)	(4)	(5)	(6)
Total # Rounds	0.931***					
	(0.131)					
Total # Words		0.017***				
		(0.002)				
Total # Advisor's Words			0.025***		0.021***	
			(0.004)		(0.006)	
Total # Investor's Words				0.040***	0.008	
				(0.006)	(0.010)	
$\mathbb{I}${Exogenous Termination}						-0.988***
						(0.351)
Profile FE	Y	Y	Y	Y	Y	Y
Clustered by Profile	Y	Y	Y	Y	Y	Y
Observations	2,500	2,500	2,500	2,500	2,500	2,500
Adjusted R^2	0.372	0.371	0.371	0.367	0.371	0.350
Note:						$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

The results demonstrate a strong positive relationship between interaction intensity and

advisor recommendation accuracy. Column (1) shows that each additional round of interaction increases the advisor's recommendation accuracy by 0.931 percentage points, indicating that extended dialogue enables the advisor to better understand investor preferences and circumstances. This finding is reinforced by the word-count analyses in Columns (2)-(4).

The decomposition of word contributions reveals interesting dynamics in the advisory relationship. When examined separately, advisor words (Column 3) increase recommendation accuracy by 0.025 points per word, while investor words (Column 4) show an even larger effect of 0.040 points per word. However, when both are included simultaneously in Column (5), the advisor's words remain highly significant at 0.021 points per word, while the investor's words become statistically insignificant. This pattern, aligned with Table 1, suggests that the advisor's ability to process and synthesize investor information is the critical factor in generating accurate recommendations.

The distribution of conversation lengths by termination type is reported in Online Appendix OA.8: exogenous stopping dominates in the first few rounds and endogenous stopping thereafter, the eleven-round cap never binds, and just a few rounds already suffice for most investors.

The central finding regarding investor impatience emerges from Column (6), which shows that exogenous termination—conversations ended artificially due to external shocks rather than reaching natural completion—reduces advisor recommendation accuracy by 0.988 percentage points.

These results provide strong support for H2, indicating that investor impatience creates meaningful costs in advisory quality.

5.3.3 Testing H3: Memory Augmentation and Advisor Performance

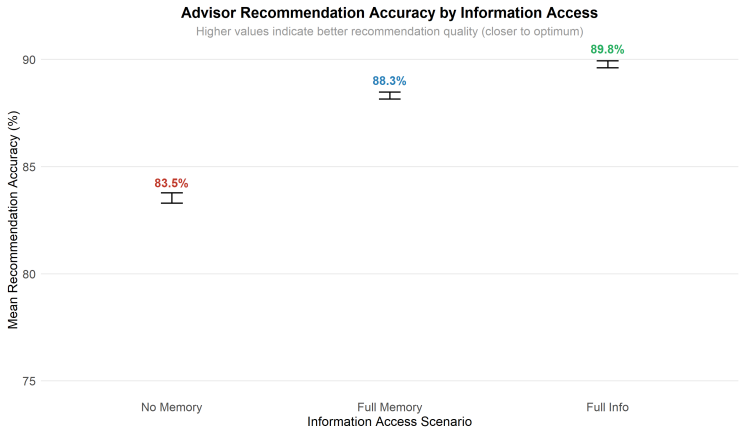
The third hypothesis examines whether providing LLM advisors with persistent memory improves their performance relative to memoryless systems. Figure 6 presents the results.

The figure shows a clear hierarchy in recommendation quality by information access: the full-information advisor achieves the highest accuracy, followed by the full-memory advisor, with the no-memory advisor performing worst. This strongly supports Hypothesis 3: both conversation memory and comprehensive profile information enhance AI advisor performance.

5.3.4 Testing H4: Optimal AI Training and Advisor Priors

The fourth hypothesis examines whether investors benefit from AI advisors whose trained priors are more opinionated than their own beliefs, rather than merely aligned. To test

Figure 6: Memory Augmentation and Advisor Performance



This figure shows mean advisor recommendation accuracy, defined as 100 minus the absolute deviation from the Vanguard-optimal allocation, under the three advisor information conditions: no memory (most recent question–answer pair only), full memory (complete conversation history), and full information (complete investor profile). Error bars represent standard errors.

this prediction, we implement a weighted pool analysis that directly maps to the theoretical structure of the model.

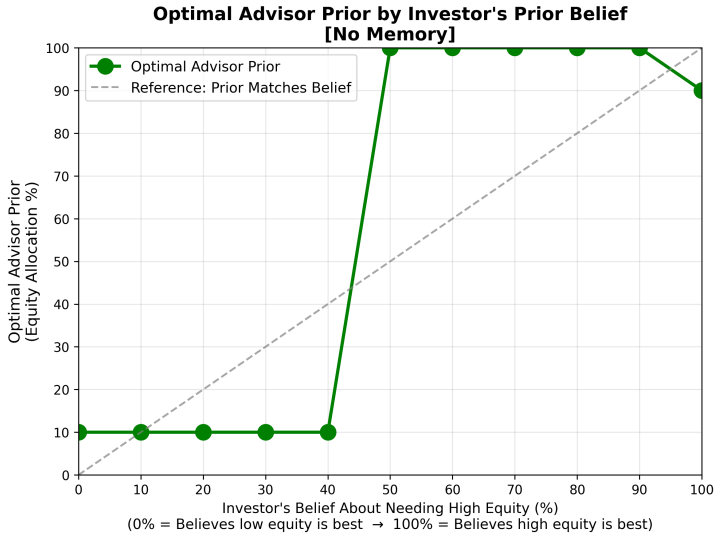
The key insight is to partition investor profiles into two pools based on their true preference type, then compute expected payoffs by weighting these pools according to the investor prior belief. We define two groups. The conservative pool consists of investors whose Vanguard optimal allocation is 0% equity, comprising 25 profiles and 15,125 observations. The aggressive pool consists of investors whose Vanguard optimal allocation is 80 to 100% equity, also comprising 25 profiles and 15,125 observations.

For each hypothetical investor with prior belief p_0 ranging from 0 to 100% in 10 percentage point increments, we compute the expected accuracy for each advisor prior level. The expected accuracy is a weighted average of accuracy in the aggressive pool, weighted by p_0, and accuracy in the conservative pool, weighted by $1 - p_0$. The optimal advisor prior for each investor belief is then the prior level that maximizes this expected accuracy. This methodology directly implements the theoretical expectation operator and allows us to trace out the empirical analog of Figure 5.

We use simulation data from 50 investor profiles drawn from the tails of the Vanguard allocation distribution. These include the 25 profiles with the lowest optimal allocations, all recommending 0% equity, and the 25 profiles with the highest optimal allocations (18 profiles at 80% and 7 profiles at 100%, yielding a mean of 85.6% equity). Each profile is simulated for 605 iterations, the full cross of 5 base iterations, 11 advisor priors, and 11 principal priors, yielding 30,250 total observations. The advisor's prior is exogenously assigned through

its system prompt over the grid $0, 10, \ldots, 100$, and the investor's preliminary allocation is likewise preset, so the treatment is the model's $\hat{p}_0$ itself rather than a proxy based on the advisor's endogenous recommendations. The outcome is the investor's final decision after hearing the recommendation of the no-memory advisor, the empirical analog of the $\kappa = 0$ baseline characterized in Proposition 6.[11]

Figure 7: Optimal Advisor Prior by Investor Prior Belief



This figure plots the optimal exogenously assigned advisor prior, measured as the equity allocation percentage stated in the advisor's system prompt, against the investor prior belief about needing high equity. The x-axis represents the investor belief, where 0% indicates certainty of needing low equity and 100% indicates certainty of needing high equity. The solid line shows the empirically optimal advisor prior for each belief level. The dashed diagonal line represents the reference case where the advisor's prior exactly matches the investor's belief. When the solid line lies below (above) the diagonal, the investor benefits from an advisor more conservative (more aggressive) than herself, indicating that opinionated advisors outperform aligned ones. The outcome is the investor's final decision after hearing the recommendation of the no-memory advisor, corresponding to the model's $\kappa = 0$ baseline. Data consists of 50 profiles with 605 iterations each, totaling 30,250 observations from LLM simulated advisor investor conversations.

Figure 7 presents the main finding. The empirical optimal advisor prior exhibits a striking pattern that qualitatively matches the theoretical prediction in Figure 5.

For investors with prior beliefs of 40% or lower, the optimal advisor prior is 10% equity: throughout the 20–40% range the curve lies below the 45 degree line, so these investors benefit from an advisor more conservative than themselves. For investors with prior beliefs of 50% or higher, the optimal advisor prior jumps to 100% equity, well above their own belief through the 90% level, so these investors benefit from an advisor more aggressive than

[11] Under the memory and full-information conditions, the optimal advisor prior remains at extreme values (10% or 90–100%) rather than at the aligned interior for every belief level, with the switch point between the conservative and aggressive corners shifting toward higher beliefs, consistent with the trained prior mattering less as the advisor's own information improves.

themselves. At the extremes, the optimal prior approximately aligns with the relevant pool's true optimum: 10% near the conservative pool's 0%, and 90% at a belief of 100%, close to the aggressive pool's mean optimum of 85.6%, consistent with the model's prediction that investors who would stop immediately prefer an aligned advisor. The one departure from the theoretical curve occurs at maximal confusion, $p_0 = 0.5$, where the model predicts an equally confused, aligned advisor but the empirical optimum selects the aggressive corner; this reflects a near-tie between the two corners generated by the slightly higher attainable accuracy in the aggressive pool.

The step function pattern in Figure 7, rather than the smooth curve in the theoretical Figure 5, reflects the discrete 10 percentage point grid on which advisor priors are assigned. Because the prior is set exogenously, every level has full data support by design, including priors that strongly contradict the investor's profile, a benchmark that is unavailable when the analysis must rely on the advisor's endogenous recommendations. Despite the discreteness, the qualitative pattern strongly supports Hypothesis 4. Investors benefit from opinionated AI advisors whose trained priors are more extreme than their own beliefs, consistent with the option value of learning emphasized in the theoretical model.

6 Conclusion

This paper studies human–machine communication: how effectively the user conveys her task to an AI, as governed by prompt quality and AI memory. This communication friction matters most for tasks involving soft information—preferences that are difficult to articulate and that the user may not fully know herself. We model prompt-based communication as the user's optimal stopping problem with Brownian information flow; because communication is inefficient, the LLM misunderstands her objective and the user debiases its recommendation.

The analysis yields three main results. First, prompt quality can substitute for AI memory: when prompts are sufficiently informative, the investor reaches certainty about her objective quickly, and even a minimally retentive AI attains the first best, which makes precise prompting a low-cost alternative to costly memory. Second, better memory benefits the investor unambiguously only when the AI's trained prior is aligned with hers; under misalignment, a more retentive AI may amplify disagreement, so limited memory can serve as a commitment device. Third, the optimal training of the AI is generally not alignment: under limited memory, the investor prefers an AI trained to be more "opinionated" than her own prior, because a more extreme AI belief delivers a recommendation that is more informative about fundamentals.

Methodologically, we propose two-sided LLM simulations as a tool for testing commu-

nication theories: both the advisor and the investor are simulated, generating multi-turn, role-structured conversations whose information sets we control. Using investor profiles drawn from the Survey of Consumer Finances and a rule-based allocation benchmark, the simulations support the model's predictions on investor learning, premature termination, memory access, and opinionated advisors. More broadly, the method occupies a middle ground between numerical simulation and field or laboratory data: unlike a simulation of the model itself, LLM agents' behavior is not dictated by the theory and can contradict it; unlike field data, the environment affords full control of information sets and complete observability. The design is portable to other theories where agents interact under specified information structures.

Finally, we view this paper as a step toward the economics of human–machine interaction. Our analysis focuses on AI advising alone, a departure from canonical advising by human experts who are strategically misaligned. An important direction for future research is supervised AI that combines the two. Our framework provides a natural starting point: a human advisor improves the elicitation of soft information, as if raising prompt quality or memory, but strategic misalignment results in information loss.

References

Adrian, Tobias, and Mark M. Westerfield, 2009, Disagreement and learning in a dynamic contracting model, *The Review of Financial Studies* 22, 3873–3906.

Anand, Kartik, Victor Duarte, and Victor Veitch, 2025, Ex machina: Collective action, coordination, and game theory with ai agents, *Available at SSRN*.

Banerjee, Snehal, and Martin Szydlowski, 2025, Trading against algorithms: Price dynamics and risk-sharing in a market with Q-learners, *Available at SSRN*.

Chatterji, Aaron, Thomas Cunningham, David J Deming, Zoe Hitzig, Christopher Ong, Carl Yan Shan, and Kevin Wadman, 2025, How people use chatgpt, Discussion paper National Bureau of Economic Research.

Chen, Joanne Juan, and Brandon Yueyang Han, 2026, When corporate AI adoption backfires, *Available at SSRN*.

Cong, Lin William, Steven R. Grenadier, and Yunzhi Hu, 2020, Dynamic interventions and informational linkages, *Journal of Financial Economics* 135, 1–15.

Crawford, Vincent P, and Joel Sobel, 1982, Strategic information transmission, *Econometrica* pp. 1431–1451.

Daley, Brendan, Thomas Geelen, and Brett Green, 2024, Due diligence, *The Journal of Finance* 79, 2115–2161.

Daley, Brendan, and Brett Green, 2012, Waiting for news in the market for lemons, *Econometrica* 80, 1433–1504.

Dou, Winston Wei, Itay Goldstein, Zigang Li, and Liyan Yang, 2025, The limits of AI trading: Market sophistication and algorithmic herding, *Available at SSRN*.

Ferreira, Daniel, and Jin Li, 2025, Artificial intelligence in the boardroom, *ECGI Working Paper*.

Grenadier, Steven R., and Andrey Malenko, 2010, A Bayesian approach to real options: The case of distinguishing between temporary and permanent shocks, *The Journal of Finance* 65, 1949–1986.

——— , and Nadya Malenko, 2016, Timing decisions in organizations: Communication and authority in a dynamic environment, *American Economic Review* 106, 2552–2581.

Gryglewicz, Sebastian, and Barney Hartman-Glaser, 2020, Investment timing and incentive costs, *The Review of Financial Studies* 33, 309–357.

He, Zhiguo, Jing Huang, and Cecilia Parlatore, 2024, Information span and credit market competition, Discussion paper National Bureau of Economic Research.

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

Ide, Enrique, and Eduard Talamàs, 2025, Artificial intelligence in the knowledge economy, *Journal of Political Economy* 133.

Levit, Doron, and Nadya Malenko, 2011, Nonbinding voting for shareholder proposals, *Journal of Finance* 66, 1579–1614.

Liberti, José María, and Mitchell A Petersen, 2019, Information: Hard and soft, *Review of Corporate Finance Studies* 8, 1–41.

Liu, Nelson F, et al., 2023, Lost in the middle: How language models use long contexts, *arXiv preprint arXiv:2307.03172*.

Lopez-Lira, Alejandro, 2025, Can large language models trade? testing efficient markets with ai agents, *Available at SSRN*.

Malenko, Andrey, and Nadya Malenko, 2019, Proxy advisory firms: The economics of selling information to voters, *The Journal of Finance* 74, 2441–2490.

Malenko, Andrey, and Anton Tsoy, 2019, Selling to advised buyers, *American Economic Review* 109, 1323–1348.

Mo, Hongwei, and Shumiao Ouyang, 2025, (Generative) AI in Financial Economics, *Available at SSRN 5287110*.

Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov, 2020, Persuading the principal to wait, *Journal of Political Economy* 128, 2542–2578.

Ouyang, Shumiao, Hayong Yun, and Xingjian Zheng, 2024, AI as decision-maker: Ethics and risk preferences of LLMs, *Available at SSRN 4851711*.

Stein, Jeremy C., 2002, Information production and capital allocation: Decentralized versus hierarchical firms, *The Journal of Finance* 57, 1891–1921.

Wang, Chupei, and Jiaqiu Vince Sun, 2025, Unable to forget: Proactive lnterference reveals working memory limits in llms beyond context length, *arXiv preprint arXiv:2506.08184*.

Weitzman, Martin L., 1974, Prices vs. quantities, *The Review of Economic Studies* 41, 477–491.

Zhong, Hongda, 2025, Optimal integration: Human, machine, and generative AI, *Management Science, forthcoming*.

A Technical Appendices

A.1 Proof of Proposition 1

Proof. We solve for the value function in the continuation region.

Step 1 (homogeneous solution). We first derive the homogeneous solution, which depends only on the diffusion of p and is therefore common to all memory levels κ. We rewrite the differential equation in log-likelihood ratio $z = \ln \frac{p}{1-p}$, so that $p(z) = \frac{e^z}{1+e^z}$, $1 - p(z) = \frac{1}{1+e^z}$. Note that

$$p'(z) = \frac{1}{z'(p)} = \frac{1}{\frac{1}{p} + \frac{1}{1-p}} = p(1-p). \tag{29}$$

Define $\Psi(z) \equiv V(p(z))$. Then

$$\Psi_z = V'(p)p'(z) = V'(p)p(1-p), \tag{30}$$

or equivalently $V'(p) = \frac{\Psi_z}{p(1-p)}$. Then the second derivative is

$$V''(p) = \frac{d}{dp}\left[\frac{\Psi_z}{p(1-p)}\right] = z'(p)\frac{d}{dz}\left[\frac{\Psi_z}{p(1-p)}\right] = \frac{1}{p^2(1-p)^2}\left[\Psi_{zz} - (1-2p)\Psi_z\right].$$

Hence, the differential equation (17) becomes

$$-c + \lambda(g(p) - \Psi(z)) + \frac{1}{2\sigma^2}[\Psi_{zz} - (1-2p)\Psi_z] = 0. \tag{31}$$

Now we substitute $\Psi(z) = u(z)\hat{\Psi}(z)$ and choose $u(z)$ to remove the Ψ_z term in (31). Then

$$\Psi_z = u_z\hat{\Psi} + u\hat{\Psi}_z, \quad \Psi_{zz} = u_{zz}\hat{\Psi} + 2u_z\hat{\Psi}_z + u\hat{\Psi}_{zz}.$$

Using them in (31), we have:

$$-c + \lambda[g(p(z)) - u\hat{\Psi}] + \frac{1}{2\sigma^2}\left\{u\hat{\Psi}_{zz} + [2u_z - (1-2p(z))\,u]\,\hat{\Psi}_z + [u_{zz} - (1-2p(z))u_z]\,\hat{\Psi}\right\} = 0.$$

To eliminate the $\hat{\Psi}_z$ term, impose

$$2u_z - (1-2p)u = 0 \;\Leftrightarrow\; \frac{u_p}{u} = \frac{1-2p}{2p'(z)} = \frac{1}{2}\left(\frac{1}{p} - \frac{1}{1-p}\right). \tag{32}$$

Integrating with respect to p, we need to set

$$u(z) = \sqrt{p(z)(1-p(z))},$$

and then the differential equation becomes

$$\hat{\Psi}_{zz} + \left[\frac{u_{zz} - (1-2p)u_z}{u} - 2\lambda\sigma^2\right]\hat{\Psi} + \frac{2\sigma^2}{u}\left[\lambda g(p) - c\right] = 0. \tag{33}$$

From the definition of the u function in (32), we know that $\frac{u_z}{u} = \frac{1-2p}{2}$. Then

$$\frac{d}{dz}\left(\frac{u_z}{u}\right) = \frac{u_{zz}}{u} - \left(\frac{u_z}{u}\right)^2,$$

which implies the coefficient in front of $\hat{\Psi}$ in (33) is a constant:

$$\begin{aligned}\frac{u_{zz}}{u} - \frac{(1-2p)u_z}{u} - 2\lambda\sigma^2 &= p'(z)\frac{d}{dp}\left(\frac{u_z}{u}\right) + \left(\frac{u_z}{u}\right)^2 - \frac{(1-2p)u_z}{u} - 2\lambda\sigma^2 \\ &= p(1-p)\cdot(-1) + \left(\frac{1-2p}{2}\right)^2 - \frac{(1-2p)^2}{2} - 2\lambda\sigma^2 = -\frac{1}{4} - 2\lambda\sigma^2.\end{aligned}$$

Therefore, with the substitution of $z \equiv \ln\frac{p}{1-p}$ and $\hat{\Psi}(z) \equiv \frac{\Psi(z)}{\sqrt{p(z)(1-p(z))}} = \frac{V(p(z))}{\sqrt{p(z)(1-p(z))}}$, the original differential equation in (17) can be rewritten as

$$\hat{\Psi}_{zz} - (\frac{1}{4} + 2\lambda\sigma^2)\hat{\Psi} + \frac{2\sigma^2}{u(p)}[\lambda g(p) - c] = 0.$$

Let $\gamma = \sqrt{\frac{1}{4} + 2\sigma^2\lambda}$. The homogeneous solution of the above differential equation is

$$\hat{\Psi}_{homo}(z) = C_1 e^{z\gamma} + C_2 e^{-z\gamma},$$

where C_1 and C_2 are constants of integration, determined below by the boundary conditions. Hence, using $e^z = \frac{p}{1-p}$, the homogeneous solution of the original differential equation (17) is

$$V_{homo}(p) = \Psi_{homo}(p(z)) = u(p)\hat{\Psi}_{homo}(z) = C_1 p^{\gamma+\frac{1}{2}}(1-p)^{-\gamma+\frac{1}{2}} + C_2 p^{-\gamma+\frac{1}{2}}(1-p)^{\gamma+\frac{1}{2}}.$$

Let $U_1(p) \equiv p^{\gamma+\frac{1}{2}}(1-p)^{-\gamma+\frac{1}{2}}$ and $U_2(p) \equiv p^{-\gamma+\frac{1}{2}}(1-p)^{\gamma+\frac{1}{2}}$ denote the two homogeneous solutions. The homogeneous part above depends only on the diffusion of p and is therefore independent of κ; the entire dependence on AI memory enters through the inhomogeneous term $g(p, \hat{p}(p;\kappa))$.

Step 2 (particular solution). AI memory enters the HJB equation only through its inhomogeneous term. In the continuation region, the HJB is

$$-\lambda V(p) + \frac{p^2(1-p)^2}{2\sigma^2} V''(p) + \lambda\, g(p, \hat{p}(p;\kappa)) - c = 0. \tag{34}$$

By Step 1, every solution of (34) takes the form $V(p) = Q(p;\kappa) + C_1 U_1(p) + C_2 U_2(p)$ for some particular solution $Q(p;\kappa)$. We construct one by variation of parameters. Rewriting (34) in normal form gives

$$V''(p) - \frac{2\lambda\sigma^2}{p^2(1-p)^2} V(p) = -\frac{2\sigma^2[\lambda\, g(p, \hat{p}(p;\kappa)) - c]}{p^2(1-p)^2} \equiv -h(p;\kappa).$$

Applying variation of parameters with the homogeneous basis (U_1, U_2),

$$Q(p;\kappa) = \frac{1}{2\gamma}\left[U_1(p)\int_{1/2}^{p} U_2(s)\, h(s;\kappa)\, ds - U_2(p)\int_{1/2}^{p} U_1(s)\, h(s;\kappa)\, ds\right], \tag{35}$$

where the factor $\frac{1}{2\gamma}$ comes from $-\frac{1}{W}$, with $W \equiv U_1 U_2' - U_1' U_2 = -2\gamma$ denoting the Wronskian of (U_1, U_2), which is constant in p. Because $g(\cdot, \hat{p}(\cdot;\kappa))$ is continuous on $(0,1)$ and $h(\cdot;\kappa)$ is locally bounded there, the integrals in (35) are well defined on any compact subinterval of $(0,1)$, and one verifies directly that Q satisfies the inhomogeneous ODE. □

A.2 Proof of Proposition 2

Proof. The investor continues if $p \in (\underline{p}, \overline{p})$ and stops to seek a final recommendation if $p = \underline{p}$ or $\overline{p}$.

Step 1. The two corner cases share the homogeneous solution derived in Step 1 of the proof of Proposition 1: $U_1(p) = p^{\gamma+\frac{1}{2}}(1-p)^{-\gamma+\frac{1}{2}}$ and $U_2(p) = p^{-\gamma+\frac{1}{2}}(1-p)^{\gamma+\frac{1}{2}}$. Here the stopping value $g(p)$ is quadratic in p, and $q_2 \equiv g''(p)$ denote the curvature in (15). One particular solution is the following:

$$V_{part}(p) = g(p) - \frac{c}{\lambda} - \frac{q_2}{\gamma}\left[U_1(p)\int_{k_1}^{p} U_2(s)\, ds - U_2(p)\int_{k_2}^{p} U_1(s)\, ds\right], \tag{36}$$

where we have applied the constant Wronskian $W(p) \equiv U_1(p) U_2'(p) - U_1'(p) U_2(p) = -2\gamma$, and the lower limits k_1, k_2 can be chosen freely and are jointly determined with C_1, C_2 from the boundary conditions.

Therefore, adding the homogeneous solution V_{homo}, the value function is

$$V(p) = g(p) - \frac{c}{\lambda} - \frac{q_2}{\gamma}\left[U_1(p)\int_{k_1}^{p} U_2(s)\, ds - U_2(p)\int_{k_2}^{p} U_1(s)\, ds\right] + C_1 U_1(p) + C_2 U_2(p). \tag{37}$$

Step 2. We show that in the continuation region, $V(p) - g(p)$ is symmetric around $\frac{1}{2}$, and the optimal stopping thresholds satisfy $\underline{p} + \overline{p} = 1$.

To see this, define the adjusted value function

$$v(p) \equiv V(p) - g(p), \tag{38}$$

which captures the option value of waiting; it is strictly positive when waiting is optimal. Substituting $V(p) = v(p) + g(p)$ into the HJB equation (17) yields

$$\lambda v(p) = \frac{p^2(1-p)^2}{2\sigma^2} g''(p) - c + \frac{p^2(1-p)^2}{2\sigma^2} v''(p).$$

Importantly, $g(p)$ is a quadratic function of p, so $g''(p)$ is a constant.

First, note that both the flow benefit of waiting $\frac{p^2(1-p)^2}{2\sigma^2}g''(p)-c$ and the volatility $\frac{p^2(1-p)^2}{2\sigma^2}$ are larger when p is close to $\frac{1}{2}$ and smaller when p is close to 0 or 1. Hence, we conjecture that the investor continues if $p\in(\underline{p},\overline{p})$ and stops to seek a final recommendation if $p=\underline{p}$ or $\overline{p}$. The value-matching and smooth-pasting conditions in terms of v are

$$v(\underline{p})=0, v'(\underline{p})=0, v(\overline{p})=0, v'(\overline{p})=0.$$

Second, both the flow benefit of waiting $\frac{p^2(1-p)^2}{2\sigma^2}g''(p)-c$ and the volatility $\frac{p^2(1-p)^2}{2\sigma^2}$ are symmetric in p around $\frac{1}{2}$. Intuitively, the labels $\omega=0$ and $\omega=1$ are interchangeable. The symmetric ODE and boundary conditions then admit a symmetric solution, so $\underline{p}+\overline{p}=1$. The symmetry also implies that the waiting solution is valid whenever the learning cost lies below a critical threshold, $c<\overline{c}$, under which $v(\frac{1}{2})>0$.

Step 3. We specify $k_1=k_2=\frac{1}{2}$ and show that $C_1=C_2$. Let $I_1(p)\equiv\int_{\frac{1}{2}}^{p}U_1(s)ds$ and $I_2(p)\equiv\int_{\frac{1}{2}}^{p}U_2(s)ds$. Evaluating (37) at the thresholds and using $\overline{p}=1-\underline{p}$, the two value-matching conditions $V(\underline{p})=g(\underline{p})$ and $V(\overline{p})=g(\overline{p})$ are

$$0=V\left(\underline{p}\right)-g\left(\underline{p}\right)=-\frac{c}{\lambda}-\frac{q_2}{\gamma}\left[U_1\left(\underline{p}\right)I_2\left(\underline{p}\right)-U_2\left(\underline{p}\right)I_1\left(\underline{p}\right)\right]+C_1U_1\left(\underline{p}\right)+C_2U_2\left(\underline{p}\right), \tag{39}$$

$$\begin{aligned}0=V\left(1-\underline{p}\right)-g\left(1-\underline{p}\right)=&-\frac{c}{\lambda}-\frac{q_2}{\gamma}\left[U_1\left(1-\underline{p}\right)I_2\left(1-\underline{p}\right)-U_2\left(1-\underline{p}\right)I_1\left(1-\underline{p}\right)\right]\\&+C_1U_1\left(1-\underline{p}\right)+C_2U_2\left(1-\underline{p}\right).\end{aligned} \tag{40}$$

Note that for any $p\in[0,1]$,

$$U_1\left(1-p\right)=U_2\left(p\right), \qquad U_2\left(1-p\right)=U_1\left(p\right), \tag{41}$$

$$U_1'\left(1-p\right)=-U_2'\left(p\right), \qquad U_2'\left(1-p\right)=-U_1'\left(p\right). \tag{42}$$

Using them in (40),

$$\begin{aligned}0=&-\frac{c}{\lambda}-\frac{q_2}{\gamma}\left[U_2\left(\underline{p}\right)\int_{\frac{1}{2}}^{1-\underline{p}}U_1\left(1-s\right)ds-U_1\left(\underline{p}\right)\int_{\frac{1}{2}}^{1-\underline{p}}U_2\left(1-s\right)ds\right]+C_1U_2\left(\underline{p}\right)+C_2U_1\left(\underline{p}\right)\\\underbrace{=}_{t=1-s}&-\frac{c}{\lambda}-\frac{q_2}{\gamma}\left[-U_2\left(\underline{p}\right)\int_{\frac{1}{2}}^{\underline{p}}U_1\left(t\right)dt+U_1\left(\underline{p}\right)\int_{\frac{1}{2}}^{\underline{p}}U_2\left(t\right)dt\right]+C_1U_2\left(\underline{p}\right)+C_2U_1\left(\underline{p}\right)\\=&-C_1U_1\left(\underline{p}\right)-C_2U_2\left(\underline{p}\right)+C_1U_2\left(\underline{p}\right)+C_2U_1\left(\underline{p}\right)\\=&\left(C_1-C_2\right)\left[U_2\left(\underline{p}\right)-U_1\left(\underline{p}\right)\right].\end{aligned}$$

The second equality follows from changing integrand from s to $t=1-s$, and the third equality uses Eq. (39). Since $\underline{p}\neq\frac{1}{2}$ whenever the investor learns, we have $U_1(\underline{p})\neq U_2(\underline{p})$ and so

$$C_1=C_2=C.$$

Step 4. We solve for the remaining unknowns C and $\underline{p}$. Using $C_1=C_2=C$, the value-matching conditions (39) and (40) imply

$$C\left(\underline{p}\right)=\frac{\frac{c}{\lambda}+\frac{q_2}{\gamma}\left[U_1\left(\underline{p}\right)I_2\left(\underline{p}\right)-U_2\left(\underline{p}\right)I_1\left(\underline{p}\right)\right]}{U_1\left(\underline{p}\right)+U_2\left(\underline{p}\right)}. \tag{43}$$

The smooth-pasting condition, obtained by differentiating (37), then pins down $\underline{p}$:

$$0=V'\left(\underline{p}\right)-g'\left(\underline{p}\right)=-\frac{q_2}{\gamma}\left[U_1'\left(\underline{p}\right)I_2\left(\underline{p}\right)-U_2'\left(\underline{p}\right)I_1\left(\underline{p}\right)\right]+C\left(\underline{p}\right)\left[U_1'\left(\underline{p}\right)+U_2'\left(\underline{p}\right)\right]. \tag{44}$$

Substituting $C(\underline{p})$ from (43) into (44) yields

$$\left.\frac{d}{dp}\left\{\frac{\frac{c}{\lambda}+\frac{q_2}{\gamma}\left[U_1\left(p\right)I_2\left(p\right)-U_2\left(p\right)I_1\left(p\right)\right]}{U_1\left(p\right)+U_2\left(p\right)}\right\}\right|_{p=\underline{p}}=0, \quad \text{where } \underline{p}\in\left(0,\tfrac{1}{2}\right). \tag{45}$$

Step 5. We establish existence and uniqueness by showing that there is a unique $\underline{p}\in\left(0,\frac{1}{2}\right)$ that solves (45). Note that

$$\operatorname{sgn}\left\{\frac{d}{dp}\left\{\frac{\frac{c}{\lambda}+\frac{q_2}{\gamma}\left[U_1\left(p\right)I_2\left(p\right)-U_2\left(p\right)I_1\left(p\right)\right]}{U_1\left(p\right)+U_2\left(p\right)}\right\}\right\}$$

$$= \operatorname{sgn}\left\{ \underbrace{\frac{q_2}{\gamma}\left(U_1' I_2 - U_2' I_1\right)\left(U_1 + U_2\right) - \left[\frac{c}{\lambda} + \frac{q_2}{\gamma}\left(U_1 I_2 - U_2 I_1\right)\right]\left(U_1' + U_2'\right)}_{\equiv M(p)} \right\}, \tag{46}$$

and we determine the sign of $M(p)$ over $p \in \left[0, \frac{1}{2}\right]$.

First, we evaluate M at two endpoints 0 and $\frac{1}{2}$. Recall that $U_1(p) = p^{\gamma+\frac{1}{2}}(1-p)^{-\gamma+\frac{1}{2}}$, $U_2(p) = p^{\frac{1}{2}-\gamma}(1-p)^{\gamma+\frac{1}{2}}$ and $I_1(p) = \int_{\frac{1}{2}}^{p} s^{\gamma+\frac{1}{2}}(1-s)^{-\gamma+\frac{1}{2}}\,ds$, $I_2(p) = \int_{\frac{1}{2}}^{p} s^{\frac{1}{2}-\gamma}(1-s)^{\gamma+\frac{1}{2}}\,ds$. Then we have

$$U_1'(p) = \left(\gamma + \frac{1}{2}\right)\left(\frac{p}{1-p}\right)^{\gamma-\frac{1}{2}} - \left(-\gamma + \frac{1}{2}\right)\left(\frac{p}{1-p}\right)^{\gamma+\frac{1}{2}},$$
$$U_2'(p) = \left(\frac{1}{2} - \gamma\right)\left(\frac{p}{1-p}\right)^{-\gamma-\frac{1}{2}} - \left(\gamma + \frac{1}{2}\right)\left(\frac{p}{1-p}\right)^{\frac{1}{2}-\gamma}.$$

At $p = 0$. Since $\gamma = \sqrt{\frac{1}{4} + 2\sigma^2\lambda} > \frac{1}{2}$, the building blocks satisfy

$$U_1(0) = 0, \quad U_1'(0) = 0, \quad U_2(0) = +\infty, \quad U_2'(0) = -\infty,$$

while $I_1(0)$ and $I_2(0)$ are finite. Substituting into the definition of M and keeping only the terms involving U_2 and U_2',

$$\begin{aligned} M(0) &= \frac{q_2}{\gamma}\left(-U_2'(0) I_1(0)\right) U_2(0) - \left(\frac{c}{\lambda} - \frac{q_2}{\gamma} U_2(0) I_1(0)\right) U_2'(0) \\ &= -\frac{c}{\lambda} U_2'(0) = +\infty. \end{aligned}$$

At $p = \frac{1}{2}$. The integrals are taken from $\frac{1}{2}$, so $I_1\left(\frac{1}{2}\right) = I_2\left(\frac{1}{2}\right) = 0$. Moreover, U_1, U_2, U_1', U_2' are all finite at $p = \frac{1}{2}$, so every term of M that carries a factor of I_1 or I_2 is zero. The only remaining term is $-\frac{c}{\lambda}\left(U_1' + U_2'\right)$; using $U_1'\left(\frac{1}{2}\right) = 2\gamma$ and $U_2'\left(\frac{1}{2}\right) = -2\gamma$, it equals $-\frac{c}{\lambda}(2\gamma - 2\gamma) = 0$, so $M\left(\frac{1}{2}\right) = 0$.

Second, we examine the monotonicity of $M(p)$ over $p \in \left[0, \frac{1}{2}\right]$. We calculate

$$\begin{aligned} M'(p) =& \frac{q_2}{\gamma}\left(U_1'' I_2 - U_2'' I_1 + 2\gamma\right)\left(U_1 + U_2\right) + \frac{q_2}{\gamma}\left(U_1' I_2 - U_2' I_1\right)\left(U_1' + U_2'\right) \\ & - \frac{q_2}{\gamma}\left(U_1' I_2 - U_2' I_1\right)\left(U_1' + U_2'\right) - \left[\frac{c}{\lambda} + \frac{q_2}{\gamma}\left(U_1 I_2 - U_2 I_1\right)\right]\left(U_1'' + U_2''\right) \\ =& 2q_2\left(U_1 + U_2\right) + \frac{q_2}{\gamma}\left(I_1 + I_2\right)\left[U_1'' U_2 - U_2'' U_1\right] - \frac{c}{\lambda}\left(U_1'' + U_2''\right) \\ =& 2q_2\left(U_1 + U_2\right) + \frac{q_2}{\gamma}\left(I_1 + I_2\right)\frac{2\sigma^2\lambda}{p^2(1-p)^2}\left[U_1 U_2 - U_2 U_1\right] - \frac{c}{\lambda}\frac{2\sigma^2\lambda}{p^2(1-p)^2}\left(U_1 + U_2\right) \\ =& 2\left(U_1 + U_2\right)\left[q_2 - \frac{c\sigma^2}{p^2(1-p)^2}\right], \end{aligned}$$

where the third equality follows because $U_1(p)$ and $U_2(p)$ are roots of the homogeneous ODE, so that $\frac{p^2(1-p)^2}{2\sigma^2} U_i''(p) = \lambda U_i(p)$ for $i \in \{1, 2\}$. Since $U_1(p) \geq 0$ and $U_2(p) \geq 0$, the sign of $M'(p)$ is determined by $q_2 - \frac{c\sigma^2}{p^2(1-p)^2}$, which is increasing in $p \in \left[0, \frac{1}{2}\right]$ and negative at $p = 0^+$.

Two cases arise according to the sign of $q_2 - \frac{c\sigma^2}{p^2(1-p)^2}$ at $p = \frac{1}{2}$. First, if it is negative, i.e., $\frac{q_2}{c\sigma^2} \leq 16$, then $M'(p) < 0$ throughout $p \in \left[0, \frac{1}{2}\right]$, so $M(p)$ is decreasing. Combined with $M(0) = \infty$ and $M\left(\frac{1}{2}\right) = 0$, $M(p)$ is strictly positive on $\left[0, \frac{1}{2}\right)$, so no interior $\underline{p}$ solves $M(\underline{p}) = 0$. Second, if $\frac{q_2}{c\sigma^2} > 16$, $M(p)$ first decreases then increases on $\left[0, \frac{1}{2}\right]$. Combined with $M(0) = \infty$ and $M\left(\frac{1}{2}\right) = 0$, $M(p)$ crosses zero exactly once on $\left[0, \frac{1}{2}\right)$, pinning down a unique $\underline{p}$. This establishes existence and uniqueness of $\underline{p}$.

To summarize, when $\frac{q_2}{c\sigma^2} > 16$, in the continuation region the value function is

$$V(p) = g(p) - \frac{c}{\lambda} - \frac{q_2}{\gamma}\left[U_1(p) I_2(p) - U_2(p) I_1(p)\right] + \frac{\frac{c}{\lambda} + \frac{q_2}{\gamma}\left[U_1\left(\underline{p}\right) I_2\left(\underline{p}\right) - U_2\left(\underline{p}\right) I_1\left(\underline{p}\right)\right]}{U_1\left(\underline{p}\right) + U_2\left(\underline{p}\right)} \cdot \left[U_1(p) + U_2(p)\right], \tag{47}$$

where $\underline{p} \in \left(0, \frac{1}{2}\right)$ is uniquely determined by (45). □

A.3 Proof of Proposition 6

Proof. We find the optimal $\hat{p}$.

1. Stopping region. First clarify the dependence of the stopping region on the candidate LLM prior. For each $\hat{p}$, let

$$\mathcal{C}(\hat{p}) \equiv (\underline{p}(\hat{p}), \overline{p}(\hat{p}))$$

denote the continuation region induced by that $\hat{p}$. Define the set of priors for which some training choice makes continuation strictly valuable by

$$\mathcal{C}_0 \equiv \left\{ p_0 : \max_{\hat{p}} V(p_0; \hat{p}) > g(p_0, p_0) \right\}.$$

By symmetry and continuity, write this set as $\mathcal{C}_0 = (\underline{p}_0, \overline{p}_0)$ with $\underline{p}_0 + \overline{p}_0 = 1$. These are the threshold prior beliefs in the proposition. Since the stopping value in (15) is

$$g(p, \hat{p}) = -\left(\Delta_\mu^2 + 2\right) p(1-p) - \frac{(p-\hat{p})^2}{\hat{p}^2 + (1-\hat{p})^2},$$

it is clear that for any p, the value is maximized at $\hat{p} = p$.

Therefore, if $p_0 \notin \mathcal{C}_0$, no candidate $\hat{p}$ generates a continuation value above immediate stopping with the aligned LLM. Since $\max_{\hat{p}} g(p_0, \hat{p}) = g(p_0, p_0)$, the optimal AI training is $\hat{p} = p_0$.

2. Continuation region. Now fix $p_0 \in \mathcal{C}_0$ and consider a candidate $\hat{p}$ such that $p_0 \in \mathcal{C}(\hat{p})$. We rewrite (47) to highlight its dependence on $\hat{p}$, noting that $q_2(\hat{p})$ is a function of $\hat{p}$ and $\underline{p}$ is affected by $q_2(\hat{p})$:

$$\begin{aligned} V(p; \hat{p}) =& g(p; \hat{p}) - \frac{c}{\lambda} - \frac{q_2(\hat{p})}{\gamma} \left[U_1(p) I_2(p) - U_2(p) I_1(p) \right] \\ &+ \frac{\frac{c}{\lambda} + \frac{q_2(\hat{p})}{\gamma} \left[U_1\left(\underline{p}\right) I_2\left(\underline{p}(\hat{p})\right) - U_2\left(\underline{p}(\hat{p})\right) I_1\left(\underline{p}(\hat{p})\right) \right]}{U_1\left(\underline{p}(\hat{p})\right) + U_2\left(\underline{p}(\hat{p})\right)} \cdot \left[U_1(p) + U_2(p) \right] \end{aligned} \tag{48}$$

Introduce $H(p) \equiv U_1(p) I_2(p) - U_2(p) I_1(p)$ and $S(p) \equiv U_1(p) + U_2(p)$. Take derivative with respect to $\hat{p}$,

$$\begin{aligned} \frac{dV(p;\hat{p})}{d\hat{p}} =& \frac{dg(p;\hat{p})}{d\hat{p}} - \frac{q_2'(\hat{p})}{\gamma} H(p) + \frac{q_2'(\hat{p})}{\gamma} \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))} \cdot S(p) + \underbrace{\frac{d}{d\underline{p}} \left\{ \frac{\frac{c}{\lambda} + \frac{q_2(\hat{p})}{\gamma} H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))} \right\}}_{=0,\ \text{(Envelope Theorem)}} \cdot S(p) \cdot \frac{d\underline{p}(\hat{p})}{d\hat{p}} \\ =& \underbrace{\frac{dg(p;\hat{p})}{d\hat{p}}}_{\text{stopping-value effect}} - \underbrace{\frac{q_2'(\hat{p})}{\gamma} S(p) \left[\frac{H(p)}{S(p)} - \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))} \right]}_{\text{option-value effect}}, \end{aligned} \tag{49}$$

where we used the Envelope Theorem so that the indirect effect through $\underline{p}$ is zero (see (45).) The first term is the direct effect on the stopping payoff. The second term captures how changing $\hat{p}$ changes the value of waiting through the learning term $q_2(\hat{p})$; the square-bracketed term measures the option-value distance between the current belief and the voluntary stopping threshold.

Differentiating (15) directly with respect to $\hat{p}$ gives

$$\begin{aligned} \frac{dg(p;\hat{p})}{d\hat{p}} &= \frac{2(p-\hat{p})\left[(1-p)(1-\hat{p}) + p\hat{p}\right]}{\left[\hat{p}^2 + (1-\hat{p})^2\right]^2}, \\ q_2'(\hat{p}) &= \frac{4\hat{p} - 2}{\left[\hat{p}^2 + (1-\hat{p})^2\right]^2}, \end{aligned}$$

where $q_2(\hat{p}) = \Delta_\mu^2 + 2 - \frac{1}{\hat{p}^2 + (1-\hat{p})^2}$. Using these derivatives in (49),

$$\frac{dV(p;\hat{p})}{d\hat{p}} = \frac{2(p-\hat{p})\left[(1-p)(1-\hat{p}) + p\hat{p}\right]}{\left[\hat{p}^2 + (1-\hat{p})^2\right]^2} - \frac{4\hat{p} - 2}{\gamma\left[\hat{p}^2 + (1-\hat{p})^2\right]^2} S(p) \left[\frac{H(p)}{S(p)} - \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))} \right]. \tag{50}$$

Recall that $H(p) = U_1(p) I_2(p) - U_2(p) I_1(p)$, $S(p) = U_1(p) + U_2(p)$, and $I_i(p) = \int_{\frac{1}{2}}^{p} U_i(s)\, ds$, where

$U_1(p) = p^{\gamma+\frac{1}{2}}(1-p)^{\frac{1}{2}-\gamma}$ and $U_2(p) = p^{\frac{1}{2}-\gamma}(1-p)^{\gamma+\frac{1}{2}}$. The key object is $\frac{H(p)}{S(p)}$. On $p \in (0, \frac{1}{2})$,

$$\left(\frac{H(p)}{S(p)}\right)' = \frac{(U_1'U_2 - U_1U_2')(I_1+I_2)}{(U_1+U_2)^2} = \frac{2\gamma(I_1+I_2)}{(U_1+U_2)^2} < 0,$$

because $I_1(p) < 0$ and $I_2(p) < 0$. By symmetry, $\frac{H(p)}{S(p)} = \frac{H(1-p)}{S(1-p)}$. Thus, when $p \in (\frac{1}{2}, 1)$, the point $1-p$ lies in $(0, \frac{1}{2})$; as p increases, $1-p$ decreases, and the decreasing property on $(0, \frac{1}{2})$ implies that $\frac{H(p)}{S(p)}$ is increasing on $(\frac{1}{2}, 1)$. Therefore, within the continuation region, $\frac{H(p)}{S(p)}$ is maximized at the stopping thresholds, and for every interior belief,

$$\frac{H(p)}{S(p)} - \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))} < 0.$$

At $p_0 = \frac{1}{2}$, symmetry of the problem implies $\hat{p}^*(\frac{1}{2}) = \frac{1}{2}$. To see the derivative directly, note that $I_1(\frac{1}{2}) = I_2(\frac{1}{2}) = 0$, so $H(\frac{1}{2}) = 0$ and $S(\frac{1}{2}) = 1$. Hence (50) gives

$$\frac{dV(\frac{1}{2};\hat{p})}{d\hat{p}} = \frac{4(\frac{1}{2}-\hat{p})}{\gamma\left[\hat{p}^2 + (1-\hat{p})^2\right]^2}\left[\frac{\gamma}{4} - \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))}\right]. \tag{51}$$

The square-bracketed term $\frac{\gamma}{4} - \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))}$ is positive. Since $\frac{H}{S}$ decreases on $(0, \frac{1}{2})$ and $\underline{p}(\hat{p}) > 0$,

$$\frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))} < \frac{H(0)}{S(0)} = -I_1(0) = \int_0^{\frac{1}{2}} s^{\gamma+\frac{1}{2}}(1-s)^{\frac{1}{2}-\gamma} ds < \frac{1}{8} < \frac{\gamma}{4}.$$

Here $\frac{H(0)}{S(0)} = -I_1(0)$ follows by plugging $p = 0$ into $H = U_1I_2 - U_2I_1$ and $S = U_1 + U_2$: since $\gamma = \sqrt{\frac{1}{4} + 2\sigma^2\lambda} > \frac{1}{2}$, $U_2(p) = p^{\frac{1}{2}-\gamma}(1-p)^{\gamma+\frac{1}{2}}$ diverges as $p \downarrow 0$, while $U_1(p) = p^{\gamma+\frac{1}{2}}(1-p)^{\frac{1}{2}-\gamma}$ remains finite. The bound on the integral is also immediate from $\gamma > \frac{1}{2}$: for $s \in (0, \frac{1}{2})$,

$$s^{\gamma+\frac{1}{2}}(1-s)^{\frac{1}{2}-\gamma} = s\left(\frac{s}{1-s}\right)^{\gamma-\frac{1}{2}} < s,$$

and therefore the integral is smaller than $\int_0^{\frac{1}{2}} s\, ds = \frac{1}{8}$. Finally, $\frac{\gamma}{4} > \frac{1}{8}$. Therefore, by (51), the value increases on $\hat{p} \in (0, \frac{1}{2})$ and decreases on $(\frac{1}{2}, 1)$, so $\hat{p}^*(\frac{1}{2}) = \frac{1}{2}$.

Now take $p_0 > \frac{1}{2}$. First, no optimizer can lie below $\frac{1}{2}$. For any $\hat{p} < \frac{1}{2}$, compare it with the mirrored prior $1-\hat{p} > \frac{1}{2}$. Since $q_2(\hat{p}) = q_2(1-\hat{p})$, the stopping thresholds and all option-value terms are the same under $\hat{p}$ and $1-\hat{p}$. Hence the value comparison is determined only by the stopping payoff. Because $1-\hat{p}$ is closer to $p_0 > \frac{1}{2}$ than $\hat{p}$ is,

$$g(p_0, 1-\hat{p}) > g(p_0, \hat{p}),$$

By the closed-form value representation in (48), the continuation-value terms depend on $\hat{p}$ only through $q_2(\hat{p})$. Since $q_2(1-\hat{p}) = q_2(\hat{p})$, mirroring $\hat{p}$ leaves these terms unchanged; therefore the value comparison inherits the sign of the stopping-payoff comparison:

$$V(p_0; 1-\hat{p}) > V(p_0; \hat{p}).$$

Thus $\hat{p}^*(p_0) \geq \frac{1}{2}$.

Second, on $\hat{p} \in [\frac{1}{2}, p_0]$, the stopping-value effect in (49) is nonnegative because $g_{\hat{p}}(p_0, \hat{p}) \geq 0$, and it is strictly positive whenever $\hat{p} < p_0$. Also, $q_2'(\hat{p}) \geq 0$, while the square-bracketed term in (49) is negative, so the option-value effect is nonnegative. At $\hat{p} = p_0$, the stopping-value effect is zero, but $q_2'(p_0) > 0$ and the square-bracketed term is strictly negative because p_0 is in the continuation region; hence the option-value effect is strictly positive there. Therefore $\frac{dV(p_0;\hat{p})}{d\hat{p}} > 0$ for every $\hat{p} \in [\frac{1}{2}, p_0]$, so $V(p_0; \hat{p})$ is strictly increasing on $[\frac{1}{2}, p_0]$. Since the derivative is still positive at $\hat{p} = p_0$, increasing $\hat{p}$ slightly above p_0 raises value. Thus the maximizer must satisfy $\hat{p}^*(p_0) > p_0$. The case $p_0 < \frac{1}{2}$ is symmetric and gives $\hat{p}^*(p_0) < p_0$.

□

OA Online Appendix

OA.1 Additional Details for Section 2.2

In this part, we provide omitted details for the communication with the LLM.

Derivation of Eq. (11) . Since $p_t(z_t) = \frac{e^{z_t}}{1+e^{z_t}}$, Ito's Lemma implies

$$\begin{aligned} dp_t =& \left(\frac{e^{z_t}}{1+e^{z_t}}\right)'_{z_t} dz_t + \frac{1}{2}\left(\frac{e^{z_t}}{1+e^{z_t}}\right)''_{z_t} (dz_t)^2 \\ =& \frac{e^{z_t}}{(1+e^{z_t})^2} dz_t + \frac{1}{2}\frac{e^{z_t}}{(1+e^{z_t})^2}\left(1 - 2\frac{e^{z_t}}{1+e^{z_t}}\right)(dz_t)^2 \\ =& p_t(1-p_t)\left[\frac{1}{\sigma^2}(p_t - \frac{1}{2})dt + \frac{1}{\sigma} dB_t\right] + \frac{1}{2} p_t(1-p_t)(1-2p_t)\frac{dt}{\sigma^2} \\ =& \frac{p_t(1-p_t)}{\sigma} dB_t. \end{aligned}$$

The third equation uses the evolvement of z_t in (10), $(dz_t)^2 = \frac{(dB_t)^2}{\sigma^2}$ and $p_t(z_t) = \frac{e^{z_t}}{1+e^{z_t}}$.

Derivation of $\hat{p}_t = \mathbb{E}(\omega = 1 | ds_{t^-})$ (no memory). We use the following Binomial approximation of the process of Brownian motion. The investor observes the whole binomial tree whereas the LLM updates its belief only based on the last signal. Specifically, The probability that the particle goes up is

$$\pi = 0.5 + \frac{\omega}{2\sigma}\sqrt{\Delta t},$$

and the particle goes down is

$$1 - \pi = 0.5 - \frac{\omega}{2\sigma}\sqrt{\Delta t},$$

and the size of jump is

$$u = \sigma\sqrt{\Delta t}, \quad d = -\sigma\sqrt{\Delta t}.$$

For a particle that starts from zero, we have

$$\pi u + (1-\pi)\, d = (2\pi - 1)\, u = \frac{\omega}{\sigma}\sqrt{\Delta t} \times \sigma\sqrt{\Delta t} = \omega \Delta t,$$

which means it has a drift of $\omega \Delta t$.

Suppose the particle goes up in the last round. The posterior belief of the LLM is

$$\hat{p}_t = \frac{\pi(\omega = 1)\hat{p}_0}{\pi(\omega = 1)\hat{p}_0 + \pi(\omega = 0)(1 - \hat{p}_0)} = \frac{\left(0.5 + \frac{1}{2\sigma}\sqrt{\Delta t}\right)\hat{p}_0}{\left(0.5 + \frac{1}{2\sigma}\sqrt{\Delta t}\right)\hat{p}_0 + 0.5(1-\hat{p}_0)}.$$

Under the continuous time setting where $\Delta t \to 0$, the above equation becomes $\hat{p}_t = \lim_{\Delta t \to 0} \frac{\left(0.5 + \frac{1}{2\sigma}\sqrt{\Delta t}\right)\hat{p}_0}{\left(0.5 + \frac{1}{2\sigma}\sqrt{\Delta t}\right)\hat{p}_0 + 0.5(1-\hat{p}_0)} = \hat{p}_0$.

OA.2 Smooth Pasting with $\kappa > 0$

The investor's problem is, in principle, two-dimensional in the pair of beliefs $(p, \hat{p})$. When both agents observe the same communication history, the LLM's belief is a function of the investor's, $\hat{p} = \hat{p}(p; \kappa)$ from (14), and the problem therefore reduces to the single state variable p with value function $V(p)$. This subsection verifies that the standard smooth-pasting conditions (21)–(22), applied to V, remain the correct optimality conditions at the stopping thresholds. The key is that, the LLM's posterior $\hat{p}(p; \kappa)$ also moves with p at the boundary into the stopping region.

Consider the upper threshold $\overline{p}$; the argument at $\underline{p}$ is symmetric. We replicate the investor's belief

evolvement (11) here,

$$dp_t = \frac{p_t(1-p_t)}{\sigma}\, dB_t.$$

At an optimal threshold, the investor must be indifferent between stopping at $\overline{p}$ and waiting one more instant before stopping. Stopping now yields $V(\overline{p}) = g\big(\overline{p}, \hat{p}(\overline{p})\big)$. Waiting an instant dt and then stopping yields the expected payoff

$$\tfrac{1}{2}\, V\big(\overline{p} + \tfrac{\overline{p}(1-\overline{p})}{\sigma}\, dB^-\big) + \tfrac{1}{2}\, V\big(\overline{p} + \tfrac{\overline{p}(1-\overline{p})}{\sigma}\, dB^+\big), \tag{52}$$

where $dB^- < 0$ moves the investor into the continuation region and $dB^+ > 0$ keeps her in the stopping region. Optimality requires that the marginal gain from waiting vanish, which is equivalent to equating the left and right derivative of V at $\overline{p}$:

$$\underbrace{\frac{V\big(\overline{p} + \frac{\overline{p}(1-\overline{p})}{\sigma}\, dB^-\big) - V(\overline{p})}{\frac{\overline{p}(1-\overline{p})}{\sigma}\, dB^-}}_{\text{continuation side}} = \underbrace{\frac{V\big(\overline{p} + \frac{\overline{p}(1-\overline{p})}{\sigma}\, dB^+\big) - V(\overline{p})}{\frac{\overline{p}(1-\overline{p})}{\sigma}\, dB^+}}_{\text{stopping side}}. \tag{53}$$

The continuation side converges to $V'(\overline{p}^-)$. On the stopping side, beyond the threshold the investor stops immediately, so V coincides with the stopping value evaluated at the *state-dependent* LLM belief:

$$V\big(\overline{p} + \tfrac{\overline{p}(1-\overline{p})}{\sigma}\, dB^+\big) = g\Big(\overline{p} + \tfrac{\overline{p}(1-\overline{p})}{\sigma}\, dB^+,\, \hat{p}\big(\overline{p} + \tfrac{\overline{p}(1-\overline{p})}{\sigma}\, dB^+\big)\Big),$$

where

$$\frac{d}{dp}\, g\big(p, \hat{p}(p)\big) = g_p\big(p, \hat{p}(p)\big) + g_{\hat{p}}\big(p, \hat{p}(p)\big)\, \hat{p}'(p),$$

which is exactly $g'(p)$ under $g(p) \equiv g\big(p, \hat{p}(p)\big)$. Taking $dt \to 0$ in (53) therefore yields

$$V'(\overline{p}) = g'(\overline{p}),$$

which is the smooth-pasting condition (22); the same steps at $\underline{p}$ deliver (21).

OA.3 Proof of Lemma 1

Proof. The investor understands that the recommendation m is based on the LLM's belief $\hat{p}$ and satisfies

$$m(\hat{p}) = \hat{p}\theta_1 + (1-\hat{p})\theta_0. \tag{54}$$

The investor chooses her optimal action $a(m)$ to solve

$$\max_a \mathbb{E}[U(a, p, \tilde{\theta}_1, \tilde{\theta}_0)] = -p\mathbb{E}[(a-\tilde{\theta}_1)^2|m] - (1-p)\mathbb{E}[(a-\tilde{\theta}_0)^2|m].$$

Her optimal action is

$$a(m) = p\,\mathbb{E}[\tilde{\theta}_1 \mid m] + (1-p)\,\mathbb{E}[\tilde{\theta}_0 \mid m]. \tag{55}$$

From the investor's perspective, $\tilde{\theta}_1$, $\tilde{\theta}_0$ and m in (54) are normal random variables. Hence, the conditional expectation of states given recommendation m are

$$\mu_{\theta_1|m} \equiv \mathbb{E}[\tilde{\theta}_1 \mid m] = \mu_1 + \frac{\hat{p}}{\hat{p}^2 + (1-\hat{p})^2}\big[m - \hat{p}\mu_1 - (1-\hat{p})\mu_0\big], \tag{56}$$

$$\mu_{\theta_0|m} \equiv \mathbb{E}[\tilde{\theta}_0 \mid m] = \mu_0 + \frac{1-\hat{p}}{\hat{p}^2 + (1-\hat{p})^2}\big[m - \hat{p}\mu_1 - (1-\hat{p})\mu_0\big]. \tag{57}$$

The investor's expected utility given any recommendation m is then

$$\begin{aligned}\mathbb{E}[U(a(m), p, \tilde{\theta}_1, \tilde{\theta}_0)|m] &= -a(m)^2 + 2a(m)\left[p\mu_{\theta_1|m} + (1-p)\mu_{\theta_0|m}\right] - p\mathbb{E}[\tilde{\theta}_1^2|m] - (1-p)\mathbb{E}[\tilde{\theta}_0^2|m] \\ &= -\underbrace{\left[p\,\mathrm{Var}(\tilde{\theta}_1 \mid m) + (1-p)\,\mathrm{Var}(\tilde{\theta}_0 \mid m)\right]}_{MSE} - p(1-p)(\mu_{\theta_1|m} - \mu_{\theta_0|m})^2.\end{aligned}$$

The conditional expectations $\mu_{\theta_1|m}, \mu_{\theta_0|m}$ are given in (56) and (57). The conditional variance $\mathrm{Var}(\tilde{\theta}_1 \mid m) = \mathrm{Var}(\tilde{\theta}_1) - \frac{\mathrm{Cov}(\tilde{\theta}_1, m)^2}{\mathrm{Var}(m)} = \frac{(1-\hat{p})^2}{\hat{p}^2+(1-\hat{p})^2}$ and $\mathrm{Var}(\tilde{\theta}_0 \mid m) = \frac{\hat{p}^2}{\hat{p}^2+(1-\hat{p})^2}$. Now we take expectation over $m(\hat{p})$ in (54)

to calculate the investor's expected utility when she seeks recommendation,

$$\begin{aligned}
g(p,\hat{p}) =& \mathbb{E}[\mathbb{E}[U(a(m),p,\tilde{\theta}_1,\tilde{\theta}_0)|m]] \\
=& -\frac{p(1-\hat{p})^2+(1-p)\hat{p}^2}{\hat{p}^2+(1-\hat{p})^2} - p(1-p)\mathbb{E}\left\{\mu_1-\mu_0+\frac{2\hat{p}-1}{\hat{p}^2+(1-\hat{p})^2}\left[m-\hat{p}\mu_1-(1-\hat{p})\mu_0\right]\right\} \\
=& -\frac{p(1-\hat{p})^2+(1-p)\hat{p}^2}{\hat{p}^2+(1-\hat{p})^2} - p(1-p)\left[\Delta_\mu^2+\frac{(2\hat{p}-1)^2}{\hat{p}^2+(1-\hat{p})^2}\right] \\
=& -\frac{p-2p\hat{p}+\hat{p}^2}{\hat{p}^2+(1-\hat{p})^2} - p(1-p)\left[\Delta_\mu^2+2-\frac{1}{\hat{p}^2+(1-\hat{p})^2}\right] \\
=& -\frac{(p-\hat{p})^2}{\hat{p}^2+(1-\hat{p})^2} - p(1-p)\left(\Delta_\mu^2+2\right)
\end{aligned}$$

Note that $g(p,\hat{p})$ is a quadratic function in p. □

OA.4 Proof of Proposition 3

Proof. Fix $\kappa > 0$ and fix the other primitives. The proof has three steps. First, any positive memory creates the endpoint singularity. From Proposition 1,

$$\frac{\hat{p}(p)}{1-\hat{p}(p)} = \frac{\hat{p}_0}{1-\hat{p}_0}\left[\frac{p/p_0}{(1-p)/(1-p_0)}\right]^\kappa.$$

Thus $\hat{p}(p) \to 0$ as $p \downarrow 0$ and $\hat{p}(p) \to 1$ as $p \uparrow 1$. Hence, by Lemma 1,

$$g(0,\hat{p}(0)) = g(1,\hat{p}(1)) = 0, \qquad g(p,\hat{p}(p)) < 0 \quad \text{for all } p \in (0,1).$$

Second, the continuation region expands as prompt quality rises. Introduce the value of waiting

$$v(p;\sigma,\kappa) \equiv V(p;\sigma,\kappa) - g(p,\hat{p}(p));$$

then $v > 0$ ($v = 0$) for the continuation (stopping) region. When $v(p;\sigma,\kappa) > 0$, the HJB for v, multiplied by σ^2, is

$$-c\sigma^2 - \lambda\sigma^2 v(p;\sigma,\kappa) + \frac{p^2(1-p)^2}{2}\left[v''(p;\sigma,\kappa) + \frac{d^2}{dp^2}g(p,\hat{p}(p))\right] = 0.$$

Take $0 < \sigma_1 < \sigma_2$. If the lower-quality prompt gave a higher option value somewhere, choose a point p^* where $v(p;\sigma_2,\kappa) - v(p;\sigma_1,\kappa)$ is maximized and positive. Then

$$v(p^*;\sigma_2,\kappa) > v(p^*;\sigma_1,\kappa) \geq 0, \qquad v''(p^*;\sigma_2,\kappa) - v''(p^*;\sigma_1,\kappa) \leq 0.$$

The HJB binds for σ_2 at p^*, while it is nonpositive under σ_1 as $v(p^*;\sigma_1,\kappa) \geq 0$. Subtracting the HJBs gives

$$\frac{(p^*)^2(1-p^*)^2}{2}\left[v''(p^*;\sigma_2,\kappa) - v''(p^*;\sigma_1,\kappa)\right] \geq \lambda\left[\sigma_2^2 v(p^*;\sigma_2,\kappa) - \sigma_1^2 v(p^*;\sigma_1,\kappa)\right] + (\sigma_2^2-\sigma_1^2)c > 0,$$

a contradiction. Hence $v(p;\sigma_1,\kappa) \geq v(p;\sigma_2,\kappa)$ for all p. The same comparison, with both HJBs binding, gives strict inequality whenever $v(p;\sigma_2,\kappa) > 0$.

At an interior regular stopping threshold for σ_2, value matching and smooth pasting give $v = 0$ and the HJB binds. Given the monotonicity discussed above, lowering σ makes the local continuation payoff strictly positive at the old stopping thresholds. Therefore, the continuation region strictly expands as σ decreases.

Third, we use this monotonicity to choose σ for a given $\varepsilon > 0$. Since $g(p,\hat{p}(p)) \to 0$ at both endpoints, choose $\delta \in (0, \frac{1}{2})$ such that

$$g(p,\hat{p}(p)) > -\varepsilon \quad \text{whenever } p \leq \delta \text{ or } p \geq 1-\delta.$$

It remains to show that the stopping thresholds eventually lie in these endpoint neighborhoods. Apply the time change $s = t/\sigma^2$ and write

$$\tilde{p}_s \equiv p_{\sigma^2 s}.$$

By Brownian scaling, $\tilde{B}_s \equiv \frac{1}{\sigma}B_{\sigma^2 s}$ is a Brownian motion in the s-clock, so

$$d\tilde{p}_s = \tilde{p}_s(1-\tilde{p}_s)d\tilde{B}_s.$$

For fixed S, following the feasible policy of communicating until calendar time $\sigma^2 S$ gives payoff

$$\mathbb{E}\left[\int_0^S e^{-\lambda\sigma^2 s}\sigma^2\left[-c+\lambda g(\tilde{p}_s,\hat{p}(\tilde{p}_s))\right]ds + e^{-\lambda\sigma^2 S}g(\tilde{p}_S,\hat{p}(\tilde{p}_S))\right].$$

Letting $\sigma \downarrow 0$ leaves only $\mathbb{E}[g(\tilde{p}_S,\hat{p}(\tilde{p}_S))]$. As $S\to\infty$, the bounded martingale $\tilde{p}_S$ converges to an endpoint, so the endpoint singularity gives $\mathbb{E}[g(\tilde{p}_S,\hat{p}(\tilde{p}_S))]\to 0$.

Now take $K=[\delta,1-\delta]$. Since $g(p,\hat{p}(p))<0$ on K and K is compact, stopping is uniformly bounded away from zero on K: there exists $\eta>0$ such that

$$g(p,\hat{p}(p))<-\eta \quad \text{for every } p\in K.$$

The convergence above is uniform on K: choose S large and then σ small so that the feasible payoff from every $p\in K$ is greater than $-\frac{\eta}{2}$. Since V is at least this feasible payoff,

$$V(p;\sigma,\kappa)>-\frac{\eta}{2}>g(p,\hat{p}(p)) \quad \text{for every } p\in K.$$

Thus stopping cannot occur in K. Therefore

$$\underline{p}(\sigma,\kappa)<\delta, \qquad \overline{p}(\sigma,\kappa)>1-\delta$$

for all sufficiently small σ. By value matching at the stopping thresholds,

$$V(\underline{p}(\sigma,\kappa);\sigma,\kappa)=g(\underline{p}(\sigma,\kappa),\hat{p}(\underline{p}(\sigma,\kappa))),$$

and similarly at $\overline{p}(\sigma,\kappa)$. The choice of δ therefore implies

$$V(\underline{p}(\sigma,\kappa);\sigma,\kappa)>-\varepsilon, \qquad V(\overline{p}(\sigma,\kappa);\sigma,\kappa)>-\varepsilon$$

whenever σ is sufficiently small. □

OA.5 Proof of Proposition 4

The following lemma gives a sufficient condition for the value function to be monotone in a generic parameter a; we apply it in the proofs below.

Lemma 2 (Comparative statics)**.** *Fix a parameter value a and suppose the continuation region is an interior interval $(\underline{p}(a),\overline{p}(a))$. Define*

$$r(p;a)\equiv -c+\lambda g(p;a).$$

If $r_a(p;a)>0$ throughout the continuation region and $g_a\geq 0$ at both stopping thresholds, then $V_a(p;a)>0$ throughout the continuation region.

Proof. Inside the continuation region, the HJB is

$$-\lambda V(p;a)+\frac{p^2(1-p)^2}{2\sigma^2}V''(p;a)+r(p;a)=0.$$

Let

$$W(p)\equiv\frac{\partial V(p;a)}{\partial a}.$$

Differentiating the HJB with respect to a gives

$$-\lambda W(p)+\frac{p^2(1-p)^2}{2\sigma^2}W''(p)=-r_a(p;a). \tag{58}$$

It remains to account for the moving stopping thresholds. At the lower threshold, value matching gives

$$V(\underline{p}(a);a)=g(\underline{p}(a);a).$$

Differentiating,

$$V_p(\underline{p}(a);a)\underline{p}_a(a)+V_a(\underline{p}(a);a)=g_p(\underline{p}(a);a)\underline{p}_a(a)+g_a(\underline{p}(a);a).$$

By smooth pasting, $V_p(\underline{p}(a);a)=g_p(\underline{p}(a);a)$, so the terms involving $\underline{p}_a(a)$ cancel and

$$W(\underline{p}(a))=g_a(\underline{p}(a);a);$$

this reflect the Envelope Theorem that there is only the direct effect of parameter change. The same argument gives

$$W(\overline{p}(a)) = g_a(\overline{p}(a); a).$$

Now suppose $r_a > 0$ and $g_a \geq 0$ at the two stopping thresholds. If W were negative somewhere, it would attain a negative interior minimum p^-. At that point $W(p^-) < 0$ and $W''(p^-) \geq 0$, so

$$-\lambda W(p^-) + \frac{(p^-)^2(1-p^-)^2}{2\sigma^2} W''(p^-) > 0,$$

contradicting (58), which implies

$$-\lambda W(p^-) + \frac{(p^-)^2(1-p^-)^2}{2\sigma^2} W''(p^-) = -r_a(p^-; a) < 0.$$

Hence $W \geq 0$. If W had an interior zero p^*, then p^* would be an interior minimum. Thus $W(p^*) = 0$ and $W''(p^*) \geq 0$, so the left-hand side of (58) is nonnegative at p^*. But the right-hand side is $-r_a(p^*; a) < 0$, a contradiction. Therefore $W > 0$ throughout the continuation region. □

Proof of Proposition 4

Proof. We focus on the parameter range where the optimal stopping thresholds are interior and the investor's prior p_0 lies in the continuation region, $0 < \underline{p}(\kappa) < p_0 < \overline{p}(\kappa) < 1$.

Direct differentiation of g in (15) together with $\hat{p}(p;\kappa)$ in (14) yields

$$\begin{aligned} \operatorname{sgn}\left(\frac{\partial}{\partial \kappa} g(p, \hat{p}(p;\kappa))\right) &= \operatorname{sgn}\left((p-\hat{p})\big[(1-p)(1-\hat{p}) + p\hat{p}\big] \frac{\partial \hat{p}(p;\kappa)}{\partial \kappa}\right) \\ &= \operatorname{sgn}\big((p - \hat{p}(p;\kappa))(z - z_0)\big), \end{aligned} \tag{59}$$

where $z = \ln \frac{p}{1-p}$ and $z_0 = \ln \frac{p_0}{1-p_0}$. The second equality follows because $(1-p)(1-\hat{p}) + p\hat{p} > 0$, $\hat{z} = \hat{z}_0 + \kappa(z - z_0)$, and $\hat{p} = \frac{e^{\hat{z}}}{1+e^{\hat{z}}}$ is strictly increasing in $\hat{z}$.

1. Aligned prior. Suppose $\hat{p}_0 = p_0$. Then (13) gives $\hat{z}(p;\kappa) = \kappa(z - z_0) + z_0$, so $\operatorname{sgn}(p - \hat{p}(p;\kappa)) = \operatorname{sgn}(z - \hat{z}) = \operatorname{sgn}(z - z_0)$ for $p \neq p_0$. Substituting into (59),

$$\operatorname{sgn}\left(\frac{\partial}{\partial \kappa} g(p, \hat{p}(p;\kappa))\right) = \operatorname{sgn}\big((z - z_0)^2\big) \geq 0,$$

with equality only at $p = p_0$. Hence $g(p, \hat{p}(p;\kappa))$ is weakly increasing in κ for every $p \in [0,1]$, and strictly increasing for $p \in (0,1) \setminus \{p_0\}$.

Apply Lemma 2 with $a = \kappa$. Since $r_\kappa = \lambda\, \partial_\kappa g > 0$ throughout the continuation region and $g_\kappa \geq 0$ at the two thresholds, Lemma 2 gives $V_\kappa(p;\kappa) > 0$ in the continuation region.

2. Misaligned prior. We construct a counterexample. For any κ, let

$$\tau(\kappa) = \inf\{t \geq 0 : p_t \notin (\underline{p}(\kappa), \overline{p}(\kappa))\}.$$

Since stopping occurs at one of the two thresholds,

$$\begin{aligned} V(p_0;\kappa) = {} & \mathbb{E}_{p_0}\Big[e^{-\lambda\tau(\kappa)} \mathbf{1}\{p_{\tau(\kappa)} = \underline{p}(\kappa)\}\Big]\, g\big(\underline{p}(\kappa), \hat{p}(\underline{p}(\kappa);\kappa)\big) \\ & + \mathbb{E}_{p_0}\Big[e^{-\lambda\tau(\kappa)} \mathbf{1}\{p_{\tau(\kappa)} = \overline{p}(\kappa)\}\Big]\, g(\overline{p}(\kappa), \hat{p}(\overline{p}(\kappa);\kappa)) \\ & - c\, \mathbb{E}_{p_0}\left[\int_0^{\tau(\kappa)} e^{-\lambda t}\, dt\right] + \lambda\, \mathbb{E}_{p_0}\left[\int_0^{\tau(\kappa)} e^{-\lambda t}\, g(p_t, \hat{p}(p_t;\kappa))\, dt\right]. \end{aligned}$$

By the envelope theorem, the indirect effects through $\underline{p}(\kappa)$ and $\overline{p}(\kappa)$ vanish, so

$$\begin{aligned} \frac{\partial V(p_0;\kappa)}{\partial \kappa} = {} & \lambda\, \mathbb{E}_{p_0}\left[\int_0^{\tau(\kappa)} e^{-\lambda t}\, \frac{\partial}{\partial \kappa} g(p_t, \hat{p}(p_t;\kappa))\, dt\right] \\ & + \mathbb{E}_{p_0}\Big[e^{-\lambda\tau(\kappa)} \mathbf{1}\{p_{\tau(\kappa)} = \underline{p}(\kappa)\}\Big]\, \frac{\partial}{\partial \kappa} g\big(\underline{p}(\kappa), \hat{p}(\underline{p}(\kappa);\kappa)\big) \\ & + \mathbb{E}_{p_0}\Big[e^{-\lambda\tau(\kappa)} \mathbf{1}\{p_{\tau(\kappa)} = \overline{p}(\kappa)\}\Big]\, \frac{\partial}{\partial \kappa} g(\overline{p}(\kappa), \hat{p}(\overline{p}(\kappa);\kappa)). \end{aligned} \tag{60}$$

Consider the no-memory case $\kappa = 0$, where $\hat{p}(p;0) = \hat{p}_0$, and take

$$\hat{p}_0 < \underline{p}(0) < p_0 < \overline{p}(0), \qquad p_0 \downarrow \underline{p}(0).$$

The LLM is trained more conservatively than the investor, $\hat{p}_0 < p_0$. At the lower threshold, $p - \hat{p} = \underline{p}(0) - \hat{p}_0 > 0$ and $z - z_0 = \underline{z}(0) - z_0 < 0$, so by (59),

$$\frac{\partial}{\partial \kappa} g\big(\underline{p}(\kappa), \hat{p}(\underline{p}(\kappa);\kappa)\big)\Big|_{\kappa=0} < 0.$$

Intuitively, the investor's belief has drifted down toward the LLM under $\kappa = 0$; a marginal increase in memory makes the LLM's belief drift down as well, enlarging the discrepancy.

As $p_0 \downarrow \underline{p}(0)$, hitting the lower threshold becomes the dominant event:

$$\tau(0) \to 0, \quad \mathbb{E}_{p_0}\left[e^{-\lambda\tau(0)}\mathbf{1}\{p_{\tau(0)} = \underline{p}(0)\}\right] \to 1, \quad \mathbb{E}_{p_0}\left[e^{-\lambda\tau(0)}\mathbf{1}\{p_{\tau(0)} = \overline{p}(0)\}\right] \to 0.$$

Substituting into (60),

$$\frac{\partial V(p_0;\kappa)}{\partial \kappa} \to \frac{\partial}{\partial \kappa} g\big(\underline{p}(\kappa), \hat{p}(\underline{p}(\kappa);\kappa)\big)\Big|_{\kappa=0} < 0,$$

so $V(p_0;\kappa_1) > V(p_0;\kappa_2)$ for $\kappa_1 < \kappa_2$ sufficiently close to 0. □

OA.6 Proof of Proposition 5

Proof. Using $\kappa = 1$ in (14) and defining $A \equiv \frac{p_0/(1-p_0)}{\hat{p}_0/(1-\hat{p}_0)}$, we can write the AI's posterior belief as

$$\hat{p}\,(p) = \frac{p}{A\,(1-p) + p}.$$

Substituting into the stopping value (15) gives

$$\begin{aligned} g\,(p;A) &= -\left(\Delta_\mu^2 + 2\right) p(1-p) - \frac{\left(p - \frac{p}{A(1-p)+p}\right)^2}{\left(\frac{p}{A(1-p)+p}\right)^2 + \left[1 - \frac{p}{A(1-p)+p}\right]^2} \\ &= -\left(\Delta_\mu^2 + 2\right) p(1-p) - \frac{(1-A)^2\,p^2\,(1-p)^2}{p^2 + A^2\,(1-p)^2}. \end{aligned}$$

The second term in $g(p;A)$ is weakly negative and vanishes at $A = 1$, so $g(p;A) \le g(p;1)$ for every $p \in [0,1]$, with strict inequality for $p \in (0,1)$ whenever $A \neq 1$. Moreover,

$$g_A(p;A) = -\frac{2(A-1)p^2(1-p)^2\left[p^2 + A(1-p)^2\right]}{\left[p^2 + A^2(1-p)^2\right]^2},$$

so $g_A > 0$ when $A < 1$ and $g_A < 0$ when $A > 1$ for every interior p.

If $A < 1$, apply Lemma 2 with $a = A$. Since $r_A = \lambda g_A > 0$ throughout the continuation region and $g_A \ge 0$ at the stopping thresholds, the value is increasing in A. Hence $V(p;A) < V(p;1)$. If $A > 1$, apply the same lemma with $a = -A$. Then $r_a = -\lambda g_A > 0$ and $g_a = -g_A \ge 0$, so the value is increasing in $-A$, equivalently decreasing in A. Again, $V(p;A) < V(p;1)$. Since $A = 1$ corresponds to the aligned prior $\hat{p}_0 = p_0$, the aligned prior maximizes the investor's value. □

OA.7 Investor Profile Correlations

Table 3 presents the correlation structure of the questionnaire responses across our SCF-based investor profiles. The matrix reveals several important patterns that validate both our sampling approach and the questionnaire's internal consistency. First, the strong positive correlations among risk tolerance questions (Q6-Q8, Q10-Q11) with coefficients ranging from 0.457 to 0.865 demonstrate that households provide coherent responses across different risk scenarios. The particularly high correlation between Q10 and Q11 (0.865), which both measure willingness to accept short-term losses, confirms that investors have stable risk preferences that manifest consistently across similar questions. Second, the high correlation between time horizon questions (Q1-Q2: 0.712) indicates natural alignment between investment timeline and planning horizon in

our SCF households. Third, the negative correlations between independent thinking (Q9) and risk tolerance measures suggest that investors who rely more on their own judgment tend to be more conservative, possibly reflecting awareness of their own knowledge limitations. Finally, the moderate correlations between financial stability indicators (Q4-Q5) and other dimensions confirm that income stability and emergency preparedness represent distinct but related aspects of financial security. These correlation patterns, emerging from actual household data rather than random generation, provide confidence that our profiles capture realistic interdependencies among investor characteristics.

Table 3: Correlation Matrix of Investor Questionnaire Responses

This table presents the correlation matrix for responses to the 11-question Vanguard Investor Questionnaire based on 500 SCF household profiles. Each question captures different aspects of investor preferences and constraints: Q1-Q2 measure investment time horizon and planning period; Q3 captures past investment experience; Q4-Q5 assess income stability and emergency fund adequacy; Q6-Q8 evaluate risk tolerance through hypothetical portfolio performance scenarios; Q9 indicates independent thinking in financial decision making; Q10-Q11 gauge willingness to accept short-term losses for potential long-term gains. All correlations are calculated using Pearson correlation coefficients on the ordinal response scales.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Q1	1.000										
Q2	0.712	1.000									
Q3	0.425	0.505	1.000								
Q4	0.493	0.659	0.148	1.000							
Q5	0.153	0.145	0.310	-0.017	1.000						
Q6	0.443	0.441	0.405	0.267	0.042	1.000					
Q7	0.386	0.401	0.440	0.215	-0.001	0.700	1.000				
Q8	0.171	0.186	0.242	0.115	-0.068	0.457	0.645	1.000			
Q9	-0.062	0.001	0.125	-0.124	0.209	-0.229	-0.294	-0.184	1.000		
Q10	0.424	0.452	0.501	0.240	0.009	0.862	0.774	0.527	-0.248	1.000	
Q11	0.478	0.518	0.570	0.313	0.020	0.779	0.853	0.601	-0.277	0.865	1.000

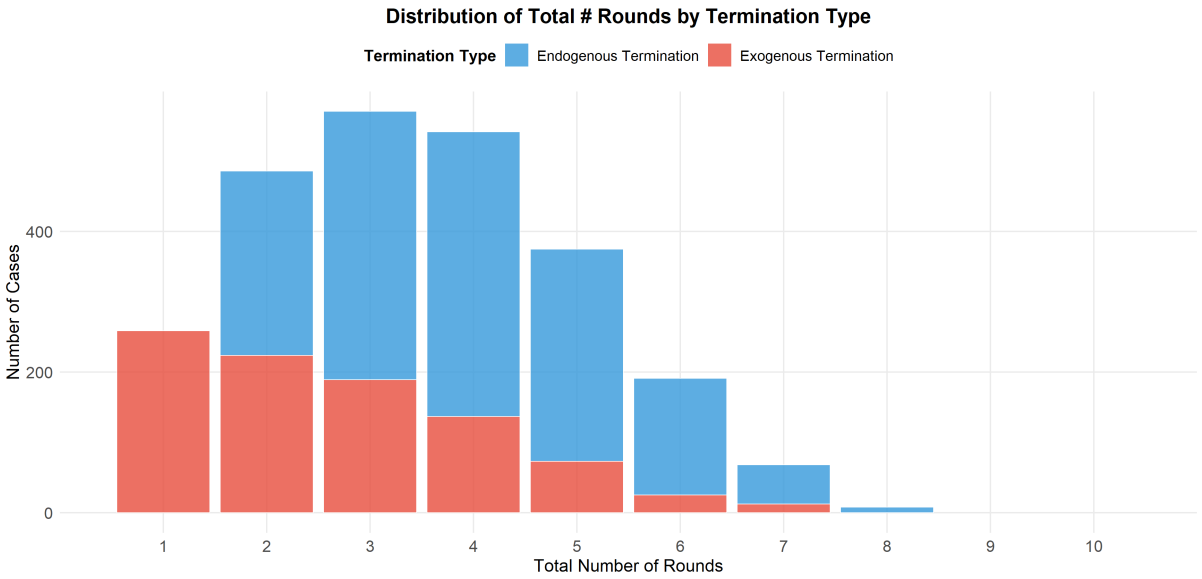
OA.8 Conversation Length Distribution

Figure 8 distinguishes between the two types of conversation endings, highlighting key dynamics of information acquisition. Endogenous terminations occur when investors stop early because their beliefs have sufficiently converged, representing optimal stopping in response to the tradeoff between information gained and the costs incurred. In contrast, exogenous terminations happen either due to a random stopping probability or upon reaching the eleven-question cap, independent of belief convergence, though the maximum length of 11 rounds is never actually reached. The distribution reveals that exogenous termination is more frequent in the first few rounds, while endogenous termination becomes the majority after the third round. This indicates that the investor actively weighs the tradeoff between gaining more information from conversation and the cost of communication. Moreover, just a few rounds of conversation are already quite helpful for the investor to become satisfied.

OA.9 Vanguard Questionnaire

1. Once you start withdrawing money from your investments, you plan to spend it over a period of...
   (a) 2 years or less
   (b) 3-5 years

Figure 8: Endogenous and Exogenous Terminations



This stacked bar chart shows the distribution of conversation rounds by termination type across the dataset. The x-axis represents the total number of question rounds, while the y-axis shows the absolute count of cases. Each bar is divided into two categories: Endogenous Termination (blue) and Exogenous Termination (red). Endogenous termination occurs when the conversation naturally concludes based on the conversation flow, while exogenous termination happens due to external shocks.

(c) 6-10 years

(d) 11-15 years

(e) More than 15 years

2. When making a long-term investment, you plan to keep the money invested for...

   (a) 1-2 years

   (b) 3-4 years

   (c) 5-6 years

   (d) 7-8 years

   (e) More than 8 years

3. When it comes to investing in stock or bond mutual funds or ETFs (or individual stocks or bonds) you would describe yourself as...

   (a) Very inexperienced

   (b) Somewhat inexperienced

   (c) Somewhat experienced

   (d) Experienced

   (e) Very experienced

4. You plan to begin taking money from your investments in...

(a) 1 year or less

(b) 1-2 years

(c) 3-5 years

(d) 6-10 years

(e) 11-15 years

(f) More than 15 years

5. Your current and future income sources (for example, salary, social security, pensions) are...

   (a) Very unstable

   (b) Unstable

   (c) Somewhat stable

   (d) Stable

   (e) Very stable

6. From September 2008 through October 2008, bonds lost 4%. If you owned a bond investment that lost 4% in two months, you would...

   (a) Sell all the remaining investment

   (b) Sell a portion of the remaining investment

   (c) Hold onto the investment and sell nothing

   (d) Buy more of the remaining investment

7. The below table shows the greatest 1-year loss and the highest 1-year gain on 3 different hypothetical investments of $10,000.

   - Investment A (gain $593; loss -$164)
   - Investment B (gain $1,921; loss -$1,020)
   - Investment C (gain $4,229; loss -$3,639)

   Given the potential gain or loss in any 1 year, you would invest your money in...

   (a) minimal volatility

   (b) moderate volatility

   (c) most volatility

8. During market declines, you tend to sell portions of your riskier assets and invest the money in safer assets. **(R)**

   (a) Strongly disagree

   (b) Disagree

   (c) Somewhat agree

   (d) Agree

   (e) Strongly agree

9. You would invest in a mutual fund or ETF (exchange-traded fund) based solely on a brief conversation with a friend, co-worker, or relative. **(R)**

   (a) Strongly disagree

(b) Disagree

(c) Somewhat agree

(d) Agree

(e) Strongly agree

10. From September 2008 through November 2008, stocks lost over 31%. If you owned a stock investment that lost about 31% in three months, you would...

    (a) Sell all the remaining investment

    (b) Sell a portion of the remaining investment

    (c) Hold onto the investment and sell nothing

    (d) Buy more of the remaining investment

11. Generally, you prefer an investment with little or no ups and downs in value, and you are willing to accept the lower returns these investments may make. **(R)**

    (a) Strongly disagree

    (b) Disagree

    (c) Somewhat agree

    (d) Agree

    (e) Strongly agree

Note: Questions marked with (R) are reverse-scored items where higher numerical responses indicate lower risk tolerance.