Soft Information, Hard Decisions: AI Advising *

Jing Huang Shumiao Ouyang

June 2026

Abstract

Large language models (LLMs) perform well on well-defined tasks, but effective prompting is challenging when tasks depend on users' soft traits. We introduce preference uncertainty—capturing soft information—into a cheap talk framework (Crawford and Sobel, 1982) and model the communication of soft traits as the user's optimal stopping problem with Brownian information flow. The investor debiases the LLM's recommendation as it misunderstands her objective. We show that prompt quality substitutes for AI memory, and that better memory unambiguously helps only when the LLM has an aligned prior. 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 based on the Survey of Consumer Finances and conduct role-structured LLM advising experiments, benchmarked against standard portfolio questionnaires.

JEL Classification: D81, D83, G11, G41, O33, C45, C63, D91
Keywords: AI, LLMs, Hardening Soft Information, AI Alignment, Financial Advising

*Huang: Mays Business School, Texas A&M University; email: jing.huang@tamu.edu; Ouyang: Saïd Business School, University of Oxford; 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, Dan Luo, Shane Miller, Hongwei Mo, Cecilia Parlatore, Jian Sun, Yizhou Xiao (discussant), Yao Zeng, Leifu Zhang, and seminar and conference participants at Texas A&M University (Mays), University of Michigan (Ross), and the 2026 PHBS Finance Symposium. 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

The proliferation of artificial intelligence (AI) in advisory roles—spanning financial planning, healthcare, and education—has created a paradox in professional services. While AI advisors demonstrate objective advantages, including the elimination of conflicts of interest, scalable delivery, and performance comparable to human professionals, their market penetration remains surprisingly limited. In financial markets, the paradox is especially pronounced: although robo-advisors create portfolios with historical returns comparable to those of professional benchmarks (Fieberg, Hornuf, Streich, and Meiler, 2024) and assist seasoned investors in minimizing behavioral errors (Guo, Tong, Chen, Kumar, and Ouyang, 2022), they account for less than 2% of global assets under management as of 2025 and are expected to grow only modestly in the near future.[1] This limited adoption persists even as traditional human advisors face well-documented incentive problems—commission-driven recommendations, product pushing, and strategic exploitation of client biases (Bolton, Freixas, and Shapiro, 2007; Carlin and Manso, 2011; Chalmers and Reuter, 2020). The persistence of potentially biased human advisors over unbiased AI alternatives suggests that current AI systems cannot replicate crucial aspects of human advisory value.

Effective AI outputs rely on well-designed prompts, which are especially challenging to construct for advisory tasks that hinge on users' *soft traits*—the subjective, often unarticulated preferences and goals that clients themselves may not be able to communicate clearly or even fully understand (Liberti and Petersen, 2019). Soft information is traditionally collected via human interactions. Nowadays, while large language models (LLMs) enable conversational exchanges, the communication of soft traits remains inefficient. Digitization inherently incurs information loss (such as tone, facial expressions, and body language), and clients often cannot articulate their needs through effective prompts, creating a "garbage-in, garbage-out" problem. Moreover, the efficiency of back-and-forth clarification depends on the AI's memory, its ability to synthesize information across extended conversations (see, e.g., Sam Altman, "people want memory"). The result is generic recommendations that overlook the nuanced requirements underlying high-stakes decisions.

This paper develops a formal framework for prompt quality and AI memory when user objectives are soft traits, and validates model predictions through LLM-driven simulations that generate multi-turn, role-structured conversations. We introduce *preference uncertainty*—capturing soft information—into a canonical cheap talk framework à la Crawford and Sobel (1982). The investor's prompt-based communication to clarify her objective is modeled as an optimal stopping problem with Brownian information flow. While AI is free of the misaligned

[1] See analysis in https://resoinsights.com/insight/robo-advisors-in-wealth-management/

incentives common to human advisors, it may misunderstand the investor's preferences and objectives. Consequently, the investor debiases the AI's recommendation. Further, we analyze how prompt quality, AI memory, and training jointly affect the investor's value.

In the model, the investor has a quadratic loss utility function and faces two layers 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 unverifiable recommendation. We assume that the investor does not observe her true preference at prior. She knows the relevant information, such as income and tax status, years until retirement, and general risk attitudes in daily life, but does not know how to formulate the investment problem or how these characteristics map into an investment objective. The communication with AI elicits such information, helping both the investor and the AI learn about her underlying preference.

We model this prompt-based communication about soft information as the investor's optimal stopping problem with Brownian information flow. As she discusses her situation, the conversation generates public signals about her underlying preferences 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 s_t, 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.[2] Our specification yields a simple relationship between the LLM's and the investor's beliefs: the update in the LLM's belief, measured in log-likelihood ratio, is exactly a κ fraction of the investor's. Notably, asymmetric beliefs between the investor and the AI persist, because, unlike in the standard literature, the AI is not strategic and does not infer the investor's belief from its own. 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,

[2]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.

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 certain about the objective and the AI's belief is aligned, minimizing the debiasing friction. In addition, because the LLM operates as a black box with limited transparency into asset fundamentals, a more extreme posterior renders its recommendation more informative about the underlying state, thereby improving the investor's payoff.

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 the benchmark case of no memory and full memory plus aligned priors, we solve the equilibrium in closed form up to one unknown variable. For general cases, we characterize the value function in closed form.

When does memory help? First, prompt quality can substitute for memory. When prompts are sufficiently informative, the investor's belief converges rapidly toward certainty, whereupon even minimal memory lets the AI reach certainty too, achieving the first best. Conversely, under worse prompt quality, the investor stops communication before certainty, incurring a loss, even when AI 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 AI 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 AI's prior. Under full memory, aligned training is optimal, since any initial misalignment is pure friction. Under no memory, however, the investor generally 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 the misalignment is as likely to shrink as to widen, the informational gain when her belief moves toward the AI's more extreme side dominates, because a more extreme AI belief yields a more informative recommendation about asset fundamentals.

We derive a set of testable implications that connect the model's primitives to observable advising behavior and outcomes. In the paper, we focus on four simulation-testable hypotheses. 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 recommenda-

tion may remain generic, but dialogue helps investors reduce "preference uncertainty" and choose portfolios closer to their true objectives. H2 predicts that higher impatience 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 advisor. H4 tests the model's prediction about optimal AI training: investors benefit from opinionated advisors whose recommendations are more extreme than their own prior beliefs, rather than merely aligned advisors.

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, which is fed only the most recent Q&A and thus approximates fixed priors due to context-length limits, with a memory-augmented advisor that can use full chat history, and a full-information counterfactual that observes the complete profile. Conversations conclude according to an optimal stopping rule, modeled as a 0.10 exogenous probability of termination per round. The investor assesses the costs and benefits in each round and may choose to end the conversation early, with a maximum of eleven questions permitted.

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 14.1 percentage points even before any specific recommendations, and by 15.7 points afterward. Each additional round of exchange adds about 1.03 points to accuracy, while every extra word contributes roughly 0.017 points. Second, in line with H2, we find that ending conversations prematurely, outside the advisor's control, reduces accuracy by around 2.62 points. Even when accounting for total rounds, this drop remains substantial at about 0.96 points, indicating a real penalty for impatience. Third, as H3 predicts, access to memory significantly boosts recommendation quality: the system works best when it has full access to information, performs next best with full memory, and does worst with no memory at all. Fourth, to match the binary preference in the model, we focus on a subsample of 50 profiles from the tails of the allocation distribution with 30,250 observations and confirm H4: the empirical optimal advisor recommendation curve qualitatively matches the theoretical prediction, with investors benefiting from advisors whose recommendations are more extreme than their own prior beliefs.

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, and show that autonomous AI primarily benefits the most knowledgeable workers whereas AI used as a co-pilot benefits the least knowledgeable. Zhong (2025) analyzes the optimal allocation of authority between humans and AI in multilayered decision processes. Chen and Han (2026) analyze how AI adoption affects agency conflicts and show that finer-grained AI estimates can exacerbate agency conflicts. Weitzner (2024) attributes algorithm aversion to reputational concerns of human decision-makers. Most closely related is Ferreira and Li (2025), who also study AI in an advisory role and show that CEO AI adoption can substitute for board advice, weaken monitoring incentives, and lead firms to reduce board independence. Our paper complements these studies by focusing on AI's role in providing advice to non-expert users, with emphasis on the communication frictions and AI-specific features such as prompt quality, memory, and training.

Cheap talk and financial advising. We build on classic cheap talk model and introduce preference uncertainty to capture soft information. We highlight the key features of AI advisors by comparing them with human advisors, who are misaligned à la the classic cheap talk (Crawford and Sobel, 1982). This seminal framework is later applied to analyst settings where reputational and underwriting incentives drive distorted advice—see Bénabou and Laroque, 1992; Rüdiger and Vigier, 2019.[3] Our paper explores AI advisors as a new alternative and identifies their limitations: while unbiased, AI systems face challenges in interpreting soft, contextual information.[4]

Our emphasis on soft-information frictions is consistent with the view that human advisors retain a complementary role in tasks requiring nuance, context, and trust (Gennaioli, Shleifer, and Vishny, 2015; Linnainmaa, Melzer, Previtero, and Foerster, 2018).

Soft information and financial technology. Our application builds on the frictions of digitizing soft information. The literature on soft vs. hard information (e.g., Liberti and Petersen, 2019; Stein, 2002) emphasizes that hard information is verifiable and thus transferable within organizations, while soft information is often non-verifiable and modeled as cheap talk (e.g., Bertomeu and Marinovic, 2016; Corrao, 2023). Recent advances in Big Data and AI have made it possible to digitize even contextual, traditionally soft, subjective information into hard, objective data. He, Huang, and Parlatore (2024) examines this technology trend in

[3]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).

[4]A parallel empirical literature documents the effects of robo- and AI-based advisors on investor outcomes; see Mo and Ouyang (2025) for a review.

the context of credit market competition. Our paper microfounds the process of digitizing soft information as the investor communication with the LLM. We further study how LLM memory and training shape the communication of soft information.

Empirical contribution via LLM simulations. We complement the theory with prompt-based LLM simulations that generate multi-turn, role-structured conversations—a methodology that lets us test theoretical predictions about soft-information transmission and optimal stopping which 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 games (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

We build on the cheap talk framework of Crawford and Sobel (1982). Our key innovation is to introduce the *quality of prompts* as the central friction: it becomes critical when the task depends on the investor's soft traits, so that effective advising requires communicating such needs through prompts.

While the model applies to a broad range of advisory applications, to fix ideas we use the context of financial advising throughout the paper. Before making an investment choice, an investor seeks advice from an AI advisor that is better informed about asset fundamentals but uncertain about the investor's underlying objective, which depends on her soft traits and must be elicited through prompts.

2.1 Agents

This subsection introduces the advisor's recommendation and the investor's choice, taking their information as given. The communication process through which this information is endogenously determined is described in Section 2.2.

2.1.1 Investor

The investor is a non-expert AI user who faces a complex, personalized decision problem. Although she knows her own circumstances—such as income and tax status, years until retirement, and general risk attitudes in daily life—she does not know how to formulate the investment problem or how these characteristics map into the right objective. To model this, we introduce into the cheap talk framework of Crawford and Sobel (1982) a second layer of uncertainty—about the investor's objective—which we refer to as *preference uncertainty*, in addition to the standard uncertainty about the realization of states.

To capture this in the simplest possible way, we assume that the investor has a quadratic loss utility function and chooses an action a to match the realization of a fundamental state. However, it is uncertain which fundamental—either $\tilde{\theta}_1$ or $\tilde{\theta}_0$—is the relevant target. That is, the investor's objective function is either

$$-\mathbb{E}(a-\tilde{\theta}_1)^2,$$

or

$$-\mathbb{E}(a-\tilde{\theta}_0)^2.$$

We introduce $\omega \in \{1,0\}$ to refer to the investor's underlying preference: $\omega = 1$ for $\tilde{\theta}_1$ and $\omega = 0$ for $\tilde{\theta}_0$. We interpret preference uncertainty as soft information, and refer to the two terms interchangeably. To fix ideas, throughout the paper we use the example of equity ($\tilde{\theta}_1$) versus fixed income ($\tilde{\theta}_0$) investment. More broadly, the setting applies whenever inexperienced decision makers face complex, personalized choices, such as selecting medical insurance or self-diagnosing medical conditions.

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. With an LLM, this communication must be digitized into prompts, which inevitably entails information loss, and the LLM's limited memory further prevents synthesis across extended exchanges. We formalize this communication in Section 2.2.

As standard in the literature, the realizations of $\tilde{\theta}_1$ and $\tilde{\theta}_0$, denoted by θ_1 and θ_0, are unobservable to the investor but observable to the advisor. We refer to uncertainty about θ_1 and θ_0 as *fundamental uncertainty*, to distinguish it from *preference uncertainty*, which concerns ω. 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 $\Delta_\mu^2 \equiv (\mu_1 - \mu_0)^2$, the squared gap in fundamental means.

We introduce $p \in [0, 1]$ as the investor's belief about her preference:

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

This belief is endogenously determined by the investor's communication with the LLM, which we describe in Section 2.2. Given p, the investor's expected utility from choosing action a is

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

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: any developer objectives beyond minimizing prediction error are likely orthogonal to recommendation bias in specific applications. However, LLMs serve the investor's interest only under their own understanding of what she wants, and building this understanding requires *effective prompts*—which are challenging when investors' soft traits are involved. The communication of soft information ω with AI is subject to two frictions.

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

Second, LLMs operate within limited, largely stateless context windows, so each response is generated primarily from the current context rather than from a persistent memory of prior exchanges.[6] This limitation is architectural: the Transformer underlying most LLMs scales quadratically with input length, making long-context processing computationally costly. Although memory capacity has been improving rapidly—through expanded context windows and emerging Agentic memory systems—current tools still fall short of the nuanced memory humans use to track clients' subtle preferences. This shortfall is especially acute for non-expert users: memory systems are only as useful as what users articulate, and including long chat histories can itself degrade performance (Liu, Lin, Hewitt, Paranjape, Bevilacqua,

[5] This is without loss of generality because if we introduce a variance σ_ϵ^2 to capture fundamental uncertainty, one can show that only $\frac{(\mu_1 - \mu_0)^2}{\sigma_\epsilon^2}$ affects the investor's choice.

[6] Certain LLMs such as ChatGPT may include prior texts in the working memory, creating the impression of memory.

Petroni, and Liang, 2023). Our model accommodates such improvements in memory.

The LLM is unbiased: its recommendation maximizes the investor's payoff, but under its own understanding of what the investor wants. Let $\hat{p} \equiv \mathbb{P}^L(\omega = 1) \in [0, 1]$ denote the LLM's belief about the investor's latent preference ω; the LLM does not observe the investor's belief p, which remains soft information. Both beliefs p and $\hat{p}$ are endogenously determined through the communication process introduced in Section 2.2. The LLM observes the fundamental realizations θ_1 and θ_0. Its payoff U^L is therefore the investor's payoff evaluated at its own belief $\hat{p}$:

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

Maximizing this payoff, the LLM recommends

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

The investor understands that m reflects the LLM's conjecture about her preferences, and therefore debiases the recommendation and treats m as a signal of θ_1 and θ_0 combined according to Eq. (4). She then chooses her optimal action $a(m, \hat{p})$ to maximize her expected utility 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, if $\hat{p} = 1$, the recommendation fully tilts toward equity, $m = \theta_1$. The investor perfectly learns the equity fundamental θ_1 but obtains no information about the fixed-income fundamental θ_0; able to infer $\hat{p}$, the investor shifts her action toward fixed income $\tilde{\theta}_0$ instead.

Let $a^*(m, \hat{p})$ denote the optimal action attaining this maximum. Then, for any pair of beliefs $(p, \hat{p})$, the investor's expected utility from seeking a recommendation is

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

Because the investor seeks the recommendation when she stops communicating about ω, Eq. (6) serves as her stopping value at any belief pair $(p, \hat{p})$ 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. Our two-state structure abstracts from the LLM's underlying algorithms, and "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.

2.2 Communication with the LLM

Before seeking a recommendation m, the investor can engage in multiple rounds of communication with the LLM to discuss her needs ω. Communication is in continuous time, starting at $t = 0$ with an infinite horizon. Each round lasts an interval dt, during which information about the investor's preference ω is revealed through a Brownian diffusion process introduced below.

Communication may end exogenously with a Poisson shock of intensity λ, capturing the possibility that the investor becomes distracted. At each time t, conditional on the Poisson shock not having occurred, the investor chooses whether to stop communication. If she continues during the interval $[t, t + dt)$, she incurs a cost of $cdt > 0$. Whenever communication stops—either endogenously by the investor or exogenously by a Poisson event—the investor receives recommendation m as defined in (4).

If at time t the game has not yet ended exogenously and the investor chooses to stop communication, her payoff is

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

where p_t and $\hat{p}_t$ are the investor's and the LLM's belief about ω at time t, respectively, and $g(p, \hat{p})$ is the stopping value defined in (6).

There are multiple ways to introduce the cost of communication. We use a Poisson shock primarily to capture the practical motivation that AI consultation is less engaging and more likely to be disrupted, and because it yields a clean closed form well suited to studying LLM memory and training. The additional flow cost c is needed because Poisson stopping per se does not penalize continued communication; without c, the value function becomes convex and the investor never stops endogenously. Other formulations such as a plain discount rate preserve the local properties and yield qualitatively similar results, though the investor's payoff needs to be adjusted because she would prefer communicating at cost rather than stop and receive a negative stopping value (quadratic loss function).

Signals about soft information ω. During communication, information about the investor's true preference ω is revealed through a continuous-time signal process modeled as a Brownian diffusion with drift ω. Specifically, the signal process evolves according to

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

where $B = \{B_t, \mathcal{F}_t, 0 \leq t \leq \infty\}$ is standard Brownian motion on the canonical probability space. The signal ds_t in (7) is more informative when σ is small, so *prompt quality* is captured by signal precision $\frac{1}{\sigma}$. At each time t, the entire history of signals, $\{s_\tau, 0 \leq \tau \leq t\}$, is publicly observable. However, as will be discussed later, the LLM only partially learns from the signals depending on its memory capacity.

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):

$$z_t = z_0 + \ln \frac{f^1(s_t)}{f^0(s_t)} = z_0 + \frac{1}{\sigma^2}(s_t - \frac{t}{2}), \tag{9}$$

and thus

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

From the investor's perspective, $\omega = 1$ with probability p_t, so the signal is released according to $ds_t = p_t dt + \sigma dB_t$, where $p_t = \frac{e^{z_t}}{1+e^{z_t}}$. Therefore, the evolvement of investor's belief measured in log-likelihood ratio is

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

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 A.1),

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

There are a few things worth noting. First, the belief process $\{p_t\}$ is a martingale without a drift term, reflecting prior consistency—the expectation of posterior beliefs equals the prior. Second, the term $\frac{1}{\sigma}$ is the signal-to-noise ratio—with the numerator equal to the gap in drifts

between $\omega = 1$ and $\omega = 0$—and it captures *prompt quality*, governing the speed at which the investor's objective is learned through communication. Finally, the belief p_t is absorbing at $p_t = 0$ or 1: once the investor becomes certain about ω, there is nothing further to learn.

Turning to the LLM, what it can learn from communication depends on its memory. Importantly, the LLM does not learn the investor's belief p_t from the conversation for two reasons: as a non-strategic agent, it does not back out p_t from its own belief; and although p_t is a sufficient statistic for the investor's information, she cannot articulate or digitize it in a prompt—this is the nature of soft information (were she able to partially convey p_t, the setting would resemble the case of better LLM memory). The LLM therefore updates its belief $\hat{p}_t$ only from the signals ds_t it absorbs, with the amount governed by its memory.

LLM's belief $\hat{p}$: no-memory baseline. As discussed in Section 2.1.2, LLMs have limited memory, and the context window largely resets between prompts. Consistent with this feature, our baseline model assumes that the LLM advisor exhibits *one-shot memory*: it updates its belief solely based on signal ds_{t^-} from the most recent round of communication. In Appendix A.1, we use a standard Binomial approximation to show that, in the continuous-time setting, one-shot memory implies that the LLM's belief remains effectively fixed at its trained prior $\hat{p}_0$:

$$\hat{p}_t = \mathbb{E}[\omega = 1 | ds_{t^-}] = \hat{p}_0. \tag{12}$$

Although the LLM's belief remains constant over the course of communication, the trained prior $\hat{p}_0$ itself can vary across LLMs and plays a central role in our analysis.

The baseline captures the following scenario : investors who are uncertain about their needs and unable to articulate their soft traits interact with an LLM whose one-shot memory leaves no room to accumulate information across the conversation, so the LLM generates recommendations based on a "typical" customer profile—that is, its trained prior $\hat{p}_0$. Meanwhile, the investor herself gradually learns about her own preference through communication, as reflected in the evolution of p_t.

LLM's belief $\hat{p}$: with memory. We also consider an extension with built-in LLM memory features. In practice, the LLM can incorporate chat histories into the working memory (for example, ChatGPT within a single chat), or rely on long-term agentic memory. However, the former leads to excessive input length which significantly hurts performance,[7] while the latter relies on summarization of past interactions. Both therefore result in information loss.

In the model, we assume that the LLM only *partially* learns from the signals. In each

[7]The Transformer architecture underlying most LLMs correlates every input token (in the context window) with the whole universe of tokens, making long inputs computationally intensive and less efficient.

interaction, 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. We assume that these missing events are independent across time.

Consider the LLM's belief as measured in the log-likelihood ratio, $\hat{z}_t \equiv \ln \frac{\hat{p}_t}{1-\hat{p}_t}$. Importantly, if the LLM misses signal ds_t, this is equivalent to receiving an alternative signal $d\tilde{s}_t$ with infinite noise. In this case, the belief update—which takes the same form as Eq. (10)—is

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

If instead the LLM absorbs ds_t, it shares the same belief update as the investor, $d\hat{z}_t = dz_t$.

Therefore, when measured in log-likelihood ratio, the LLM's belief update is exactly a κ fraction of that of the investor,[8]

$$d\hat{z}_t = \kappa dz_t. \tag{13}$$

Equivalently,

$$\hat{z}_t = \hat{z}_0 + \int_0^t d\hat{z}_s = \hat{z}_0 + \kappa(z_t - z_0). \tag{14}$$

Eq. (14) implies that the LLM's belief $\hat{p}_t$ is a deterministic function of the investor's belief p_t (given the priors p_0 and $\hat{p}_0$).

Note that this extension nests the baseline model as the special case $\kappa = 0$: the LLM never learns from the conversation, and its belief remains fixed at prior $\hat{p}_0$. At the other extreme, when $\kappa = 1$, the LLM updates one-for-one with the investor, and their beliefs coincide if and only if their priors match ($\hat{p}_0 = p_0$).

Remark 2 (Asymmetric information)**.** In our model, the investor is rational whereas the LLM is not. Given the LLM's learning in (14), the investor conjectures the LLM's posterior belief $\hat{p}_t(p_t)$ based on her own belief p_t. The LLM, however, is a "machine" rather than a strategic agent: it updates its belief mechanically but does not infer the investor's belief p_t from (14). If both parties were rational, asymmetric information would not arise in this extension.

[8] 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 $k = 1, 2, ..., K$.

2.3 Equilibrium Definition

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

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

where her belief p_t evolves according to Eq. (11). 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 $t < \tau$, the game ends exogenously and the investor receives $g(p_t, \hat{p}_t)$; otherwise, with probability $e^{-\lambda \tau}$, no shock arrives by τ, and she stops voluntarily and receives $g(p_\tau, \hat{p}_\tau)$. The stopping value g is given by Eq. (6).

Definition 1. 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 (14), with the baseline case of $\kappa = 0$ implying $\hat{p}_t = \hat{p}_0$;

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

(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

We construct the equilibrium with the LLM advisor and solve it in closed form.

3.1 Equilibrium construction

We first take the agents' beliefs $(p, \hat{p})$ as given and characterize the investor's expected payoff from seeking a recommendation. The communication process then endogenizes the agents' beliefs, and the investor's stopping value corresponds to her expected payoff characterized above. We construct the equilibrium of interest and provide a closed-form solution.

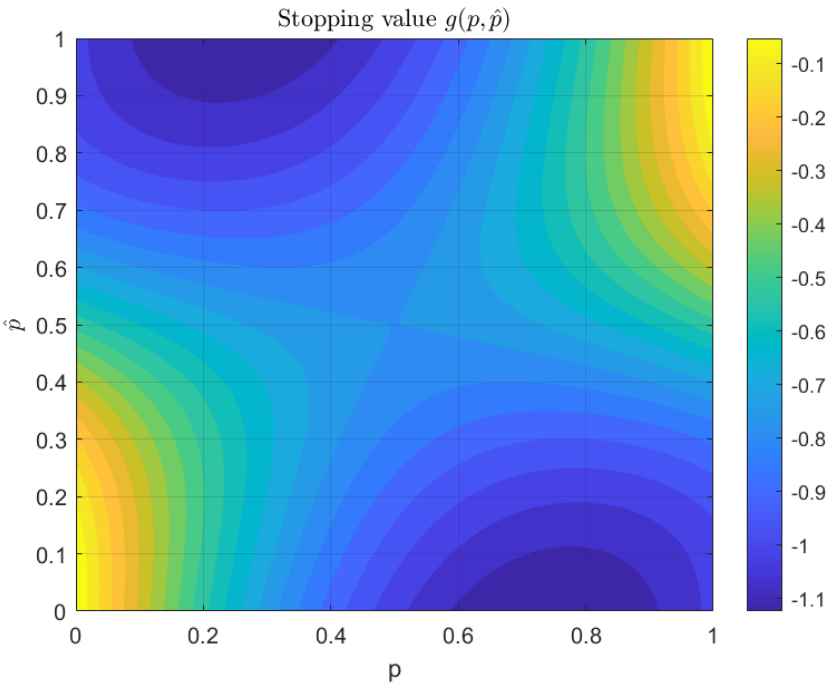


Figure 1: **Investor's stopping value** $g(p, \hat{p})$**.** Figure shows the investor's expected utility when seeking recommendation, $g(p, \hat{p})$, as a function of the investor's belief p in the horizontal axis and the LLM's belief $\hat{p}$ in the vertical axis. Parameters: $\Delta_\mu = \mu_1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

3.1.1 Stopping value

Recall that the investor's expected payoff when seeking a recommendation is defined in (6). The following lemma characterizes its value.

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. Unlike the human advisor in canonical models who is strategically biased (Crawford and Sobel, 1982), here the LLM is unbiased and maximizes the investor's expected payoff—albeit under its own belief $\hat{p}$, so the investor still needs to debias according to her own belief p. A more aligned LLM

belief thus reduces the need for debiasing. In addition, a more extreme LLM belief makes debiasing more efficient: as shown by the denominator $\hat{p}^2 + (1-\hat{p})^2$ of this term, the LLM's recommendation is more informative about the underlying fundamentals. (Remark 1 gives a short discussion of why the LLM does not send (θ_1, θ_0) as a message directly.)

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 Δ_μ between stock ($\tilde{\theta}_1$) and fixed income ($\tilde{\theta}_0$).

Figure 1 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 a more informative signal 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 both agents observe the same information flow, the LLM's belief $\hat{p}$ is a function of the investor's, allowing us to track only one state variable, p, and $g(p) \equiv g(p, \hat{p}(p))$ is the investor's stopping value.

The above properties of $g(\cdot)$ have two implications. First, the investor has an incentive to continue communication to become more certain about ω, trading off this benefit against the communication cost. Second, upon stopping, the investor most prefers a perfectly aligned LLM belief.

3.1.2 Value function and optimal policy

Recall that the investor communicates with the LLM about ω, and her belief p evolves endogenously through this communication according to (11); the LLM's belief is a function of p, $\hat{p} = \hat{p}(p; \kappa)$ (and its memory κ). The dynamic communication problem is stationary, and the investor's value function depends only on her belief p. Let $V(p)$ denote the investor's value in state p. 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 $g(p) \equiv g(p, \hat{p}(p))$ from Lemma 1.

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:

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

The first term, $-c$, is the flow communication cost. The second, $\lambda(g(p)-V(p))$, reflects the Poisson shock that forces the investor to stop at intensity λ, replacing her continuation value $V(p)$ with the stopping value $g(p)$. The last term is the option value of further belief updating, driven by the diffusion of p in (11).

We show in Appendix A.3 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 two constants C_1 and C_2 are yet to be determined.

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

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

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

3.2 Equilibrium under no LLM memory

We start by fully characterizing the equilibrium in the baseline case $\kappa = 0$: the LLM has one-shot memory and makes its recommendation based solely on the contemporaneous prompt ds_{t^-}, and under the continuous time setting, its belief therefore remains fixed at its trained prior throughout: $\hat{p}_t = \mathbb{E}[\omega = 1 \mid ds_{t^-}] = \hat{p}_0$.

In this case, we show that the investor's optimal stopping thresholds $\underline{p}$ and $\overline{p}$ are symmetric around $\frac{1}{2}$. Note that a tilted LLM belief $\hat{p}$ makes the investor's *value* asymmetric, with a higher payoff when her belief p is closer to $\hat{p}$. This asymmetry, however, does not carry over to the stopping rule. To see this, let $v(p) \equiv V(p) - g(p)$ denote the option value of waiting. Rearranging the HJB equation (17) gives

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

Under Poisson impatience, the investor's loss when the shock hits is exactly the option value of waiting $v(p)$, so her decision is governed by the marginal value of learning—the curvature $g''(p)$—not by the level of g. When $\kappa = 0$, g is quadratic, so g'' is constant; together with the symmetric belief volatility $\frac{p^2(1-p)^2}{2\sigma^2}$, this makes the option value $v(p)$ symmetric. Intuitively, absent LLM learning, the two states $\omega \in \{0, 1\}$ are interchangeable labels, so resolving uncertainty toward 0 or toward 1 is equally valuable; symmetric stopping thresholds arise.

In addition, the quadratic $g(\cdot)$ gives a closed-form particular solution. Taken together, this allows us to solve the equilibrium in closed form up to a single unknown variable as summarized in the following proposition.

Proposition 1. *When $q_2 \equiv \Delta_\mu^2 + 2 - \frac{1}{\hat{p}^2 + (1-\hat{p})^2} > 16c\sigma^2$, there exists a unique pair of stopping thresholds $\underline{p}$ and $\overline{p}$ that are symmetric around $\frac{1}{2}$, such that*

1. *When $p \notin (\underline{p}, \overline{p})$, the investor stops communication and receives $g(p, \hat{p})$ in (15);*
2. *When $p \in (\underline{p}, \overline{p})$, the investor communicates and her value is*

$$V(p) = g(p) - \frac{c}{\lambda} - \frac{q_2}{\gamma}\left[U_1(p) I_2(p) - U_2(p) I_1(p)\right] + C(\underline{p})\left[U_1(p) + U_2(p)\right], \tag{23}$$

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

$$C(\underline{p}) = \frac{\frac{c}{\lambda} + \frac{q_2}{\gamma}\left[U_1\left(\underline{p}\right) I_2\left(\underline{p}\right) - U_2\left(\underline{p}\right) I_1\left(\underline{p}\right)\right]}{U_1\left(\underline{p}\right) + U_2\left(\underline{p}\right)}. \tag{24}$$

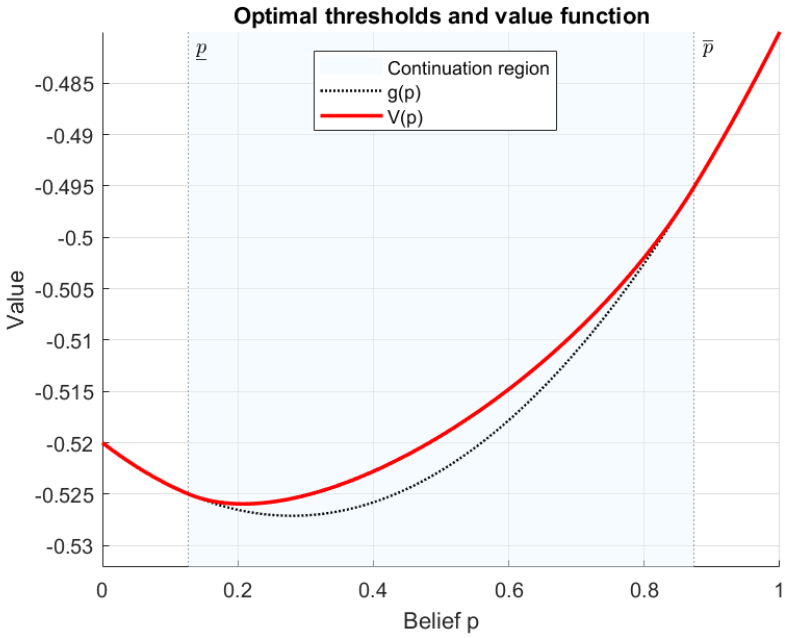


Figure 2: **Optimal policy and value function (baseline $\kappa = 0$).** The figure shows the continuation region $p \in (\underline{p}, \overline{p})$ (shaded area), the investor's value function $V(p)$ (red line) and stopping value $g(p) \equiv g(p, \hat{p}_0)$ (black dotted). Parameters: $\Delta_\mu = 0.3,\ \sigma = 0.3,\ \lambda = 0.3,\ c = 0.035,\ \hat{p}_0 = 0.51$.

The optimal stopping thresholds are uniquely determined by $\overline{p} = 1 - \underline{p}$ *and*

$$\frac{d}{dp}\left\{\frac{\frac{c}{\lambda} + \frac{q_2}{\gamma}\left[U_1(p) I_2(p) - U_2(p) I_1(p)\right]}{U_1(p) + U_2(p)}\right\}\Bigg|_{p=\underline{p}} = 0, \quad \text{where } \underline{p} \in (0, \tfrac{1}{2}). \tag{25}$$

The equilibrium requires that the option value of learning is large enough to justify the communication cost, formalized by the condition $g''(p) = q_2 > 16c\sigma^2$, where c is the flow cost of communication and $\frac{1}{\sigma^2}$ measures the quality of prompts.

Figure 2 illustrates the equilibrium. 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 $g(p)$, so she chooses to continue communication. Notably, even when preference uncertainty is fully resolved—$p = 1$ or $p = 0$—the investor still does not attain the first-best payoff: as the LLM never learns, frictions arises as the investor debiases due to belief misalignment. This is no longer the case once the LLM has even a bit of memory and first-best is attained at $p = 0, 1$, as we show in Section 4.1.

Importantly, the symmetry of the optimal policy depends on the quadratic form of $g(p)$

in the no-memory baseline; it breaks once the LLM learns ($\kappa > 0$), making learning more valuable towards the LLM's direction, as we show in Section 4.1. This quadratic form of $g(p)$ is general in economic applications. When an agent's payoff is linear in her information, as she also chooses an action based on that information, the resulting value function becomes quadratic in the information.

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 Equilibrium with AI memory

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 recent years. Yet memory is still costly and may lose information—a limitation especially consequential in our setting of soft traits and imprecise prompts.

As specified in Section 2.2.1, we capture AI memory by a parameter $\kappa \in [0, 1]$: for each signal s_t, the LLM absorbs it with probability κ—reflecting the memory capacity—and otherwise misses it. Asymmetric beliefs between the investor and the AI persist, because, unlike in the standard literature, the AI is not strategic and does not infer the investor's belief from (26). This is a new friction that sustains asymmetric beliefs in dynamic settings, distinct from canonical sources such as heterogeneous priors (e.g. Adrian and Westerfield, 2009), and that will become increasingly relevant as AI gains traction in economic settings.

In general, closed-form stopping thresholds are no longer available, except in the case of full memory $\kappa = 1$ and aligned priors, $\hat{p}_0 = p_0$, as summarized in Proposition 2 below.

Proposition 2. *Suppose the LLM has full memory and shares the investor's prior, $\kappa = 1$ and $\hat{p}_0 = p_0$. The equilibrium has the same structure as Proposition 1, evaluated at the new value of $q_2 = \Delta_\mu^2 + 2$: the value function is given by (23), and the stopping thresholds $\underline{p}, \overline{p} = 1 - \underline{p}$ are given by (25).*

The assumption of *aligned prior* $\hat{p}_0 = p_0$ is not merely an analytical convenience. We show in Section 4.3 that it is optimal for the investor. In this case, the LLM's belief tracks

the investor's exactly:

$$\hat{z}_t - \hat{z}_0 = z_t - z_0 \Rightarrow \hat{p}_t = p_t.$$

Then $\hat{p}(p; \kappa = 1) = p$ and the stopping value collapses to $g(p, \hat{p}) = -q_2\, p(1-p)$. This leads to a simple, quardratic stopping payoff in p, $g(p, \hat{p}) = -q_2\, p(1-p)$, which drives the simple structure of the equilibrium in Proposition 2.

Because the investor and the LLM always share the same beliefs, there is no debiasing of AI recommendations. The remaining frictions reduce to standard communication costs: in the model, whether the investor is willing to wait at cost c for further learning, given that prompts have informativeness $1/\sigma$.

This benchmark further clarifies the friction of soft information. Although the investor's belief p_t is a sufficient statistic of the conversation history, its soft information nature prevents her from conveying it directly to the AI as "hard" data—if she could, the resulting equilibrium be that of Proposition 2 regardless of AI memory. Therefore, memory matters precisely because soft information is communicated through back-and-forth interactions, which AI can now also partially mediate beyond canonical human interactions.

We now characterize the equilibrium in the general case.

Proposition 3. *Given AI memory $\kappa \in (0, 1]$, the LLM's belief $\hat{p}_t$ is a function of the investor's belief p_t:*

$$\hat{p}_t\left(p_t\right) = \frac{\frac{\hat{p}_0}{1-\hat{p}_0}\left[\frac{(p_t/p_0)}{(1-p_t)/(1-p_0)}\right]^{\kappa}}{1 + \frac{\hat{p}_0}{1-\hat{p}_0}\left[\frac{(p_t/p_0)}{(1-p_t)/(1-p_0)}\right]^{\kappa}}. \tag{26}$$

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 $V(p) = g(p, \hat{p}(p; \kappa))$, where g is defined in (15);*

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

$$V(p) = Q(p; \kappa) + C_1 U_1(p) + C_2 U_2(p), \tag{27}$$

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 $Q(p; \kappa)$ is a particular solution constructed in Appendix A.4. 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).

Figure 3 provides an illustration of the equilibrium with AI memory. A notable consequence of AI memory is the singularity near $p = 0, 1$. In the left panel, the investor attains the highest possible payoff, $U = 0$, when she is certain about her preference—that is, when

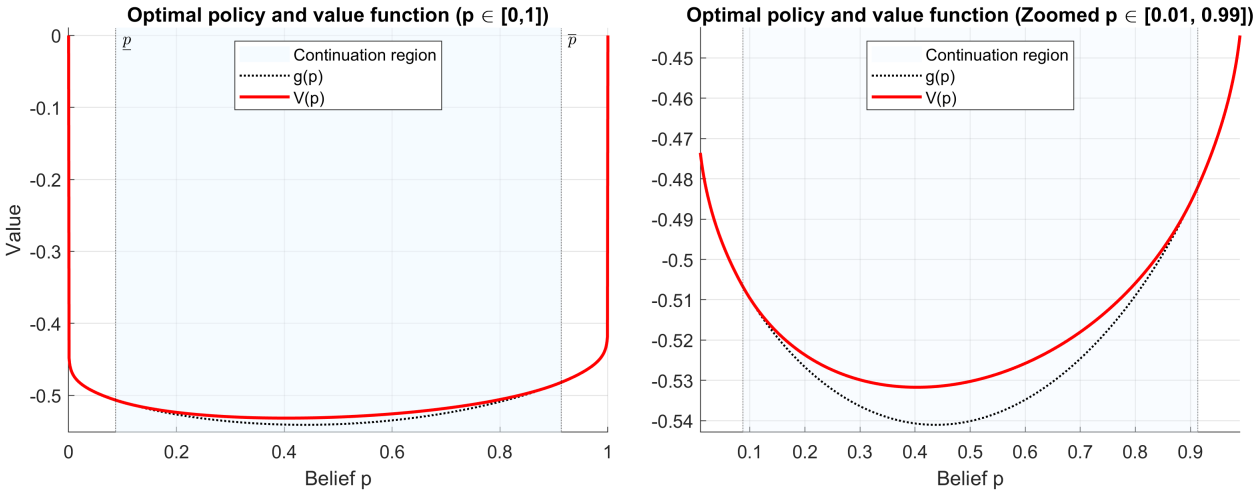


Figure 3: **Optimal policy and value function (extension $\kappa > 0$).** The left panel shows the full plot for $p \in [0, 1]$ and the right panel shows the zoomed plot $p \in (0.01, 0.99)$. The figure shows the continuation region $p \in (\underline{p}, \overline{p})$ (shaded area), the investor's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

$p = 0$ or 1. For such an expert user, adding a small amount of memory to the AI can achieve the first best. Investors without such extreme beliefs stop learning before reaching $p = 0$ or 1, but communication becomes more efficient and engaging when the LLM has memory. A comparison with the baseline shows that the investor stops at more extreme $\underline{p}, \overline{p}$, even at a higher communication cost.

In general, the stopping thresholds $\underline{p}$ and $\overline{p}$ are no longer symmetric around $\frac{1}{2}$ when $\hat{p}_0 \neq \frac{1}{2}$, because the LLM's option value of learning now matters as well. Since the LLM's learning is tilted toward its trained belief $\hat{p}_0$, the investor trades off correcting a misaligned LLM against obtaining a more informative recommendation from an LLM with a more extreme belief, which makes the continuation region asymmetric.

4.2 When does AI memory help?

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 corollary.

Corollary 1. *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.*

Thus, even arbitrarily small AI memory is sufficient to approximate the first best when prompt quality is high enough. The required prompt quality may be higher when memory is smaller.

Conversely, when prompts are imprecise (large σ), even substantial memory—say $\kappa = 1$—fails to resolve the uncertainty about ω in finite time, and the investor must communicate longer to extract useful recommendations. Therefore, prompt quality and memory are substitutes: better prompt quality relaxes the requirement on AI memory.

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 (i) is intuitive: with aligned priors, a more retentive LLM learns more about the investor's objective and produces better-aligned recommendations.

When the LLM is trained with a misaligned prior, however, greater memory need not benefit the investor, because the LLM's learning may overshoot. Suppose the investor leans toward fixed income ($\omega = 0$) at her prior, while the LLM is trained to be aggressive and equity-oriented ($\omega = 1$). If the conversation reveals the investor to be *more* aggressive than her prior, a more retentive LLM updates further in the wrong direction, amplifying rather than correcting the misalignment. In this case, a shorter memory acts as a commitment device: by limiting how much the LLM learns, it prevents recommendations from drifting too far from the investor's true objective.

4.3 Optimal AI training

Expanding AI memory is expensive, requiring costly advances in model architecture and infrastructure. A cheaper and more readily adjustable lever is *alignment*: how strong a view the model is trained to hold about the user it serves. Indeed, how opinionated an AI

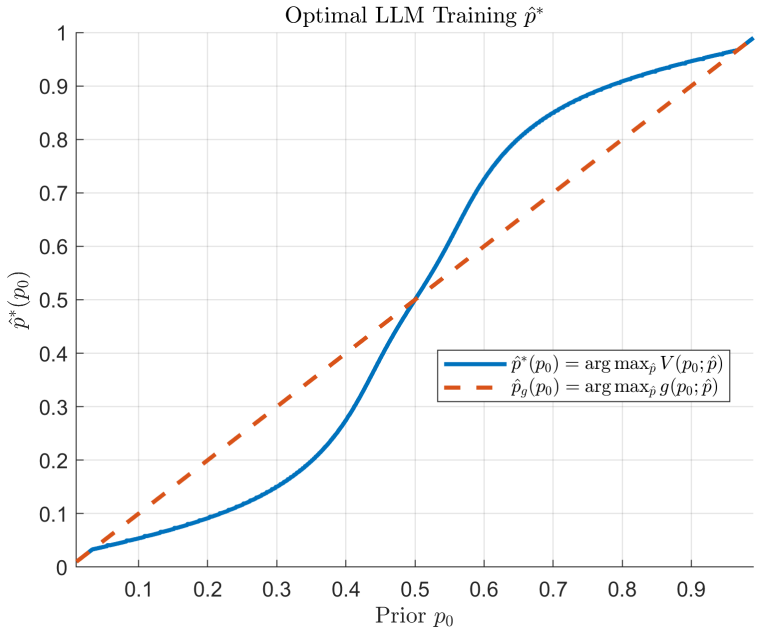


Figure 4: **Optimal LLM Training (baseline $\kappa = 0$).** The figure plots the optimal training of the LLM, $\hat{p}^*$, for users with different prior beliefs p_0. The dashed red line plots a benchmark case of the optimal LLM training if the user does not communicate and immediately seeks a recommendation. Parameters: $\Delta_\mu = 0.7,\ \sigma = 0.45,\ \lambda = 0.3,\ c = 0.08,\ \kappa = 0.$

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 belief 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 evolve endogenously as martingales during the conversation.

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

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. For simplicity, we focus on the baseline case of $\kappa = 0$, where the LLM's belief is fixed at the training level $\hat{p}_0$. The optimal LLM belief is plotted in Figure 4 and Proposition 6 presents the formal results.

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 user with prior belief $p_0 \in [0, 1]$, the optimal LLM training choice $\hat{p}^*(p_0) \equiv$*

$\arg\max_{\hat{p}} V(p_0;\hat{p})$ *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 receive $g(p_0;\hat{p})$ 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 4.

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 4, when the investor's prior is $p_0 = 0.2$, which leans toward fixed income ($\omega = 0$), her most preferred LLM prior is the more opinionated 0.1, rather than the aligned prior of 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's more extreme belief delivers a more informative recommendation about the realization of the fixed-income fundamental θ_0, and this gain outweighs the loss incurred when her belief instead drifts toward equity. On net, she prefers the opinionated LLM.

When the investor is most uncertain, with a prior of $p_0 = 0.5$, the problem is symmetric across the two directions, so neither tilt is preferred and she chooses an equally "confused" LLM, $\hat{p}^*(p_0 = 0.5) = 0.5$.

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.

The model posits that the investor is initially "confused" and does not fully understand her own needs summarized by ω. The dialogue with the LLM provides signals $\{ds_t\}$ that allow the investor to update her own beliefs about her type, moving her closer to certainty—that is, her belief about ω, p approaches 1 or 0. Because we control the LLM's information set to be the last prompt (short memory), its recommendation does not adapt to a deeper conversation. However, the now better-informed investor can make a more appropriate final decision based on her sharpened posterior belief.

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

In our model, an impatient investor may experience a Poisson "impatience shock" that prompts them to end the communication early. This premature truncation of the dialogue results in the investor gathering less information, which in turn leads to a portfolio decision that is less tailored to the investor's true 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.

The theoretical motivation suggests that human advisors naturally retain and recall earlier parts of conversations and the client's background, unlike a standard LLM which may lose mid-conversation context due to lack of memory. Enhancing the LLM with tools such as automatically generated summaries of past chat history should help reduce this information loss, enabling the LLM to better understand and address the investor's needs.

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.

Consider a retail investor who suspects she may be more risk-tolerant than average, but is not entirely sure. Should she consult a general-purpose robo advisor, or one specialized for aggressive investors? Counterintuitively, our model (Proposition 6) suggests she is better off with the aggressive specialist. Here is why. She will gain if she truly belongs to the aggressive type, as the specialist's recommendations will be well-suited; however, she will lose if she turns out to be conservative after all, as the mismatch with the advisor aggressive stance widens. The first effect dominates because an extreme advisor's recommendation is more informative about underlying asset fundamentals.

The same logic applies in reverse. An investor who leans conservative benefits from consulting an advisor trained on even more conservative clients. The key insight is that, for an investor subject to both preference and fundamental uncertainty, the option value of learning primarily reduces the fundamental uncertainty when interacting with an opinionated LLM, while preference uncertainty remains unchanged ex ante because belief updating follows a martingale. The exception is the investor who is genuinely uncertain and has no directional leaning. For her, an equally uncertain, neutral advisor is optimal, since there is no prior inclination to amplify.

5.2 Hypothesis Testing with LLM Simulations

To complement our analytical framework, we use prompt-based LLM simulations, a methodological innovation that enables the testing of complex economic theories in realistic, interactive settings. Unlike traditional laboratory experiments or analytical models, these simulations can generate dynamic, multi-turn conversations that closely approximate real-world

advisory relationships. This approach bridges the critical gap between theoretical rigor and empirical realism, allowing us to observe how theoretical mechanisms like information acquisition and belief updating operate in practice. 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.[9] 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 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.

Table 1 presents the correlation structure of the questionnaire responses across our SCF-

[9]The service is available at the website: https://investor.vanguard.com/tools-calculators/investor-questionnaire/questions. The questionnaires are listed in Appendix B.1.

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 1: 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

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 ($b = 0$). 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 fundamental uncertainty. 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 choosing the optimal stopping time τ once her belief p_t makes further interaction suboptimal. The theoretical problem is expressed as:

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

where p_t is the investor's belief about her type and $g(\cdot)$ is the payoff from acting.

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 2 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 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 advisor's words appear to be the critical factor in guiding productive self-reflection, helping investors articulate their preferences efficiently rather than

Table 2: 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:								$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

meandering through unfocused self-expression.

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. For AI advisory systems, this means the greatest value comes from designing sophisticated questioning frameworks that help users systematically explore and clarify their preferences, rather than focusing solely on recommendation algorithms.

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 3 presents regression results analyzing the factors that influence the accuracy of the AI advisor's investment recommendations.

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 the findings in Table 2, suggests that while investor input is valuable for providing information, the advisor's ability to process, synthesize, and respond to that information is the critical factor in generating accurate recommendations.

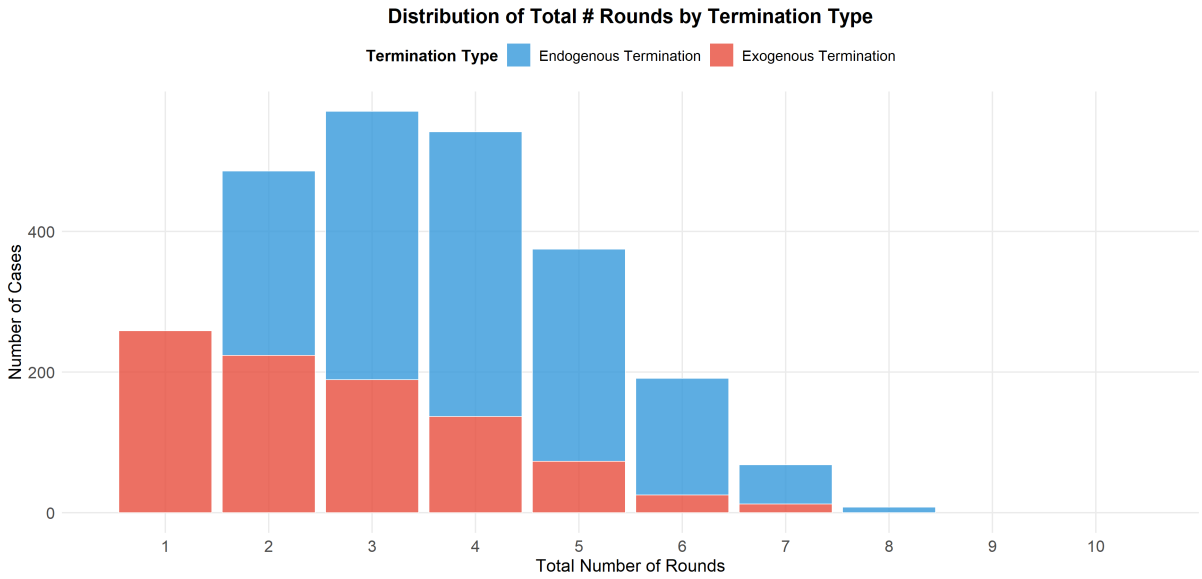
The distribution of the total number of rounds in Figure 5 distinguishes between two types of conversation endings, highlighting key dynamics of information acquisition. Endogenous terminations occur when investors stop early because their beliefs have sufficiently converged, representing optimal stopping 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

Table 3: 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:						$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Figure 5: 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.

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.

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. This nearly full percentage point penalty demonstrates that premature conversation endings significantly impair the advisor's performance.

These results provide strong support for H2, indicating that investor impatience creates meaningful costs in advisory quality. The consistent negative impact of early termination, combined with the positive effects of extended interaction, suggests that the full advisory process requires adequate time and engagement to function effectively. For AI advisory systems, this highlights the importance of designing mechanisms that encourage sustained engagement and discourage premature exit from the advisory process.

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 testing the benefits of memory augmentation.

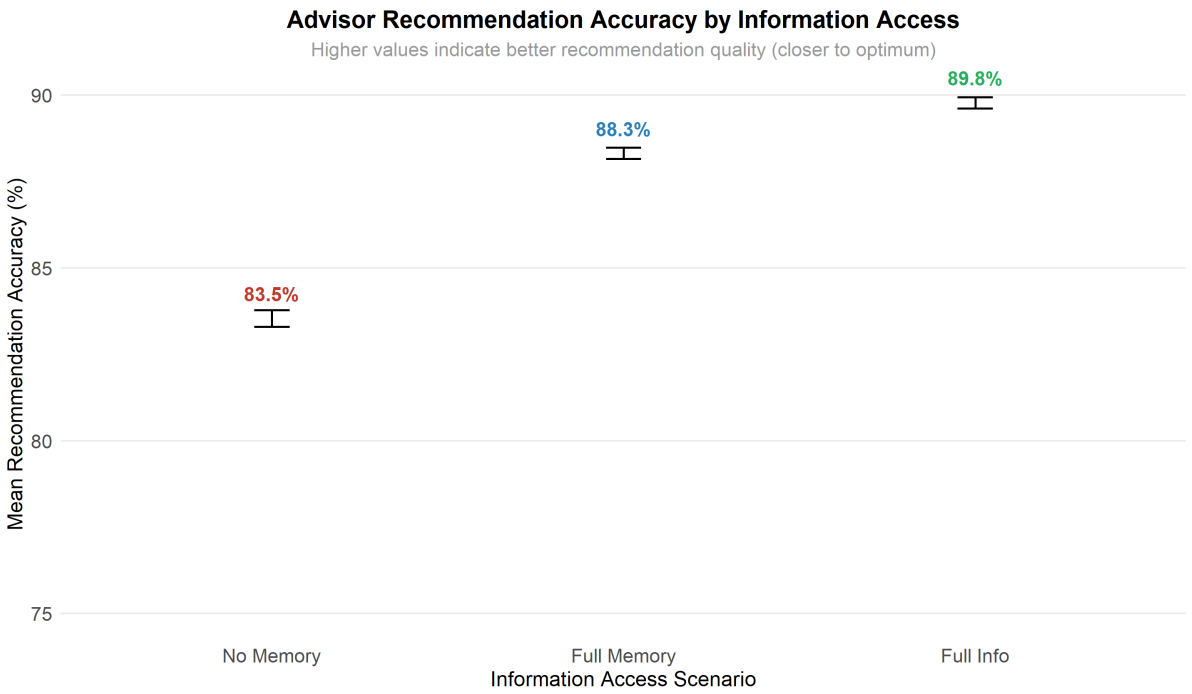
The visualization demonstrates a clear hierarchy in recommendation quality based on information access: scenarios where the advisor has access to all available information including the investor's complete profile achieve the highest accuracy, followed by scenarios where the advisor has access to the complete conversation history, while scenarios where the advisor only sees the current question perform the worst. This strongly supports Hypothesis 3, illustrating that both conversation memory and comprehensive profile information significantly enhance AI advisor performance. These results underscore that memory augmentation and full access to user information are both valuable for generating accurate investment recommendations.

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 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 prefer-

Figure 6: Memory Augmentation and Advisor Performance



This bar chart compares AI advisor recommendation accuracy across three information access scenarios: No Memory (advisor only sees the current question), Full Memory (advisor has access to the complete conversation history), and Full Info (advisor has access to all available information including the investor's complete profile). The y-axis shows mean recommendation accuracy as a percentage, with higher values indicating better recommendation quality and closer to the optimal equity allocation. Error bars represent standard errors.

ence 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 4.

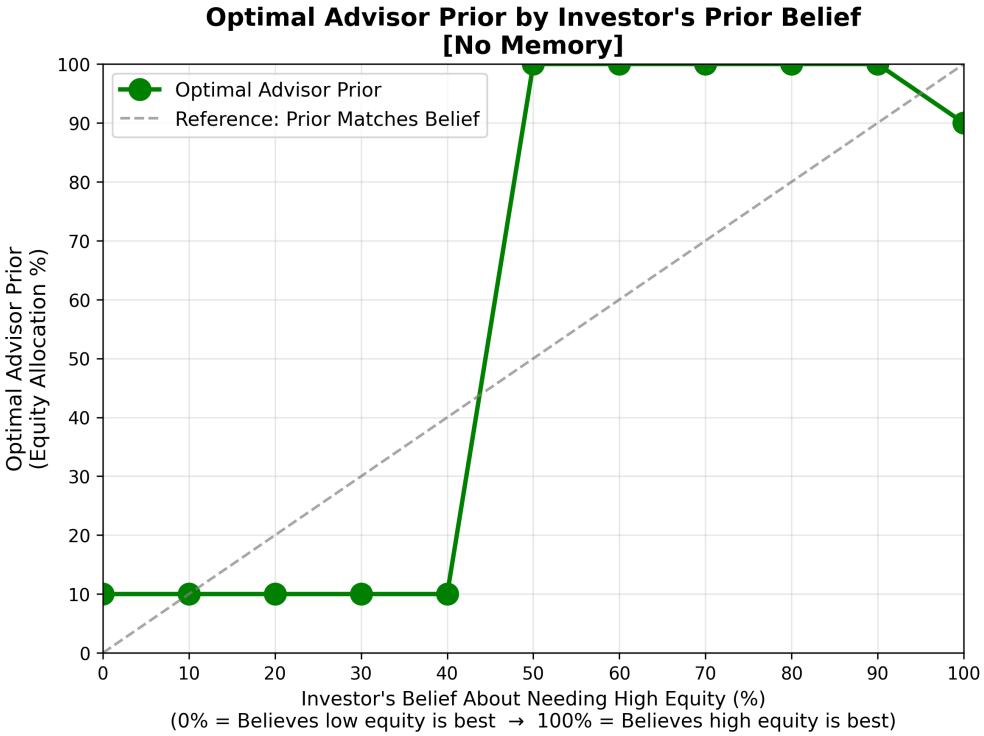
We use simulation data from 50 investor profiles drawn from the tails of the Vanguard allocation distribution. These include the 25 profiles with the lowest optimal allocations, all recommending 0% equity, and the 25 profiles with the highest optimal allocations (18 profiles at 80% and 7 profiles at 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.[10]

Figure 7 presents the main finding. The empirical optimal advisor prior exhibits a striking pattern that qualitatively matches the theoretical prediction in Figure 4.

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 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

[10]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.

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.

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 4, 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.

5.3.5 Summary of Empirical Findings

The empirical results provide strong support for all four hypotheses derived from the theoretical model. H1 demonstrates that the primary value of LLM advising comes from investors clarifying their own preferences through self-reflection rather than receiving increasingly personalized recommendations, with each additional round improving accuracy by 0.77 percentage points and each word exchanged contributing 0.012 points to accuracy. H2 reveals that investor impatience, manifested through early conversation termination, significantly undermines advisor performance, with exogenous termination reducing recommendation accuracy by 0.99 percentage points. H3 shows that memory augmentation substantially improves AI advisor performance, with scenarios providing full information access achieving the highest accuracy, followed by full memory access, while no-memory scenarios perform worst. H4 confirms the model prediction about optimal AI training. Using a focused subsample of 50 profiles from the tails of the allocation distribution with 605 iterations each, and exogenously assigning the advisor's prior through its system prompt, we find that investors benefit from opinionated advisors whose trained priors are more extreme than their own prior beliefs, with the empirical optimal prior curve qualitatively matching the theoretical prediction in Proposition 6.

The ability to generate large-scale, realistic simulations (2,500 conversations across 500 profiles, plus 30,250 conversations across 50 tail profiles for H4) while maintaining the complexity of individual interactions provides unprecedented opportunities to test theories that depend on realistic communication patterns and adaptive behavior. By allowing researchers to observe mechanisms like optimal stopping in a practical setting, this method-

ology opens new possibilities for economic research in domains where human behavior is context-dependent and communication plays a central role. It offers a promising path forward for both theoretical development and the design of practical AI systems.

6 Conclusion

This study models the interaction between human and AI financial advisors as a dynamic optimal stopping game under two layers of uncertainty. We find that human advisors excel at interpreting soft, subjective information and clarifying ambiguous investor goals, while AI advisors provide unbiased, scalable recommendations but struggle to process unspoken preferences. This "soft information gap" limits the efficiency of AI-driven advice in complex decision-making contexts. Our results show that advisor value depends on the clarity of investor preferences. When goals are uncertain or evolving, human advisors' interpretive strengths can outweigh their incentive biases. Conversely, when preferences are well-defined, unbiased AI advisors can match or outperform humans. Moreover, in contexts where human advice is heavily biased, AI guidance offers a clear advantage.

This work contributes a theoretical framework that integrates strategic communication, advisor incentives, and the role of soft information, extending classic models and complementing recent empirical findings on robo-advisors and generative AI. By formalizing the inefficiencies created by digitizing soft information, we highlight key trade-offs in advisory relationships and set the stage for deeper integration of AI in finance.

A key methodological innovation of this study is the use of LLM simulations to operationalize and test our theoretical framework. By simulating realistic advisory interactions, we demonstrate how LLMs can serve as scalable, controlled environments for theory validation and refinement. This approach not only bridges theoretical and empirical analysis but also opens new avenues for using AI-driven simulations in behavioral finance and decision sciences.

Practically, our findings support a hybrid advisory approach. Novice or uncertain investors may benefit from initial human interaction, while experienced clients with clear goals can rely on AI platforms for efficient, unbiased advice. Enhancing AI systems with better memory and context-handling could further reduce information loss. Policymakers, meanwhile, should ensure that AI-driven advice remains transparent, unbiased, and aligned with fiduciary standards.

Future research could explore hybrid human–AI models, adaptive AI systems with extended memory, and applications in other fields such as medical or legal advising. Our framework provides a foundation for advancing the theory and practice of financial advice

in the age of AI.

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* SSRN 5543139.

Bénabou, Roland, and Guy Laroque, 1992, Using privileged information to manipulate markets: Insiders, gurus, and credibility, *The Quarterly Journal of Economics* 107, 921–958.

Bertomeu, Jeremy, and Iván Marinovic, 2016, A theory of hard and soft information, *The Accounting Review* 91, 1–20.

Bolton, Patrick, Xavier Freixas, and Joel Shapiro, 2007, Conflicts of Interest, Information Provision, and Competition in the Financial Services Industry, *Journal of Financial Economics* 85, 297–330.

Carlin, Bruce Ian, and Gustavo Manso, 2011, Obfuscation, Learning, and the Evolution of Investor Sophistication, *The Review of Financial Studies* 24, 754–785.

Chalmers, John, and Jonathan Reuter, 2020, Is Conflicted Investment Advice Better Than No Advice?, *Journal of Financial Economics* 138, 366–387.

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*.

Corrao, Roberto, 2023, Mediation markets: The case of soft information, Discussion paper Working Paper.

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

Ferreira, Daniel, and Jin Li, 2025, Artificial intelligence in the boardroom, *ECGI Working Paper / Available at SSRN* SSRN id 5409542.

Fieberg, Christian, Lars Hornuf, David Streich, and Maximilian Meiler, 2024, Using Large Language Models for Financial Advice, .

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, 2015, Money doctors, *The Journal of Finance* 70, 91–114.

Guo, Yue, Siliang Tong, Tao Chen, Subodha Kumar, and Shumiao Ouyang, 2022, The impact of generative artificial intelligence on individual manual investment decisions: Empirical evidence from mutual funds, *Nanyang Business School Research Paper*.

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.

Linnainmaa, Juhani, Brian Melzer, Alessandro Previtero, and Stephen Foerster, 2018, Financial advisors and risk-taking, *Unpublished manuscript*.

Liu, Nelson F, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang, 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.

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

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

Rüdiger, Jesper, and Adrien Vigier, 2019, Learning about analysts, *Journal of Economic Theory* 180, 304–335.

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*.

Weitzner, Gregory, 2024, Reputational algorithm aversion, *Available at SSRN*.

Zhong, Hongda, 2025, Integrating humans and technologies in decision-making, *Available at SSRN*.

A Technical Appendices

A.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

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

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

Derivation of $\hat{p}_t = \mathbb{E}(\omega = 1 | ds_{t^-})$ (baseline). 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 $\hat{p}_t = \lim_{\Delta t \to 0} \frac{\left(0.5+\frac{1}{2\sigma}\sqrt{\Delta t}\right)\hat{p}_0}{\left(0.5+\frac{1}{2\sigma}\sqrt{\Delta t}\right)\hat{p}_0+0.5(1-\hat{p}_0)} = \hat{p}_0$.

A.2 Proof of Lemma 1

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

$$m(\hat{p}) = \hat{p}\theta_1 + (1-\hat{p})\theta_0. \tag{29}$$

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

$$a(m) = p\,\mathbb{E}[\tilde{\theta}_1 \mid m] + (1-p)\,\mathbb{E}[\tilde{\theta}_0 \mid m]. \tag{30}$$

From the investor's perspective, $\tilde{\theta}_1$, $\tilde{\theta}_0$ and m in (29) are normal random variables. Hence, the conditional expectation of states given recommendation m are

$$\mu_{\theta_1|m} \equiv \mathbb{E}[\tilde{\theta}_1 \mid m] = \mu_1 + \frac{\hat{p}}{\hat{p}^2+(1-\hat{p})^2}\Big[m - \hat{p}\mu_1 - (1-\hat{p})\mu_0\Big], \tag{31}$$

$$\mu_{\theta_0|m} \equiv \mathbb{E}[\tilde{\theta}_0 \mid m] = \mu_0 + \frac{1-\hat{p}}{\hat{p}^2+(1-\hat{p})^2}\Big[m - \hat{p}\mu_1 - (1-\hat{p})\mu_0\Big]. \tag{32}$$

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

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

The conditional expectations $\mu_{\theta_1|m}, \mu_{\theta_0|m}$ are given in (31) and (32). 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 (29) to calculate the investor's expected utility when she seeks

recommendation,

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

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

A.3 Proof of Proposition 1

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. We solve for the value function in the continuation region in closed form. 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

$$p'(z) = \frac{1}{z'(p)} = \frac{1}{\frac{1}{p}+\frac{1}{1-p}} = p(1-p). \tag{33}$$

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

$$\Psi_z = V'(p)p'(z) = V'(p)p(1-p), \tag{34}$$

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

$$-c+\lambda(g(p)-\Psi(z))+\frac{1}{2\sigma^2}[\Psi_{zz}-(1-2p)\Psi_z] = 0. \tag{35}$$

Now we substitute $\Psi(z) = u(z)\hat{\Psi}(z)$ and choose $u(z)$ to remove the Ψ_z term in (35). 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 (35), we have:

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

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

$$2u_z-(1-2p)u=0\;\Leftrightarrow\;\frac{u_p}{u}=\frac{1-2p}{2p'(z)}=\frac{1}{2}\left(\frac{1}{p}-\frac{1}{1-p}\right). \tag{36}$$

Integrating with respect to p, we need to set

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

and then the differential equation becomes

$$\hat{\Psi}_{zz}+\left[\frac{u_{zz}-(1-2p)u_z}{u}-2\lambda\sigma^2\right]\hat{\Psi}+\frac{2\sigma^2}{u}\left[\lambda g(p)-c\right]=0. \tag{37}$$

From the definition of the u function in (36), 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 (37) 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

$$\hat{\Psi}_{homo}(z)=C_1e^{z\gamma}+C_2e^{-z\gamma},$$

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 $U_2(p) = p^{-\gamma+\frac{1}{2}}(1-p)^{\gamma+\frac{1}{2}}$ denote the roots, and let q_0, q_1, q_2 denote the coefficients for stopping value $g(p) = q_0 + q_1 p + q_2 p^2$ in (15). Conjecture the following particular solution

$$V_{part}(p) = g(p) - \frac{c}{\lambda} + 2q_2 \left[U_1(p) \int_{k_1}^{p} \frac{U_2(s)}{W(s)} ds - U_2(p) \int_{k_2}^{p} \frac{U_1(s)}{W(s)} ds \right]. \tag{38}$$

where $W(p) \equiv U_1(p) U_2'(p) - U_1'(p) U_2(p)$ is the Wronskian, 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. Since $U_1(p)U_2(p) = p(1-p)$ and $\frac{U_2'(p)}{U_2(p)} - \frac{U_1'(p)}{U_1(p)} = -\frac{2\gamma}{p(1-p)}$, the Wronskian is the constant

$$W(p) = U_1(p)U_2(p)\left[\frac{U_2'(p)}{U_2(p)} - \frac{U_1'(p)}{U_1(p)}\right] = -2\gamma.$$

Substituting $W(s) = -2\gamma$, the particular solution (38) becomes

$$V_{part}(p) = g(p) - \frac{c}{\lambda} - \frac{q_2}{\gamma} \left[U_1(p) \int_{k_1}^{p} U_2(s)\, ds - U_2(p) \int_{k_2}^{p} U_1(s)\, ds \right]. \tag{39}$$

We verify that (39) satisfies the ODE in (17); that is,

$$\frac{2\lambda\sigma^2}{p^2(1-p)^2}\left[V_{part}(p) - g(p) + \frac{c}{\lambda}\right] = V_{part}''(p). \tag{40}$$

The left-hand side equals

$$\frac{2\lambda\sigma^2}{p^2(1-p)^2} \cdot \left(-\frac{q_2}{\gamma}\right) \left[U_1(p) \int_{k_1}^{p} U_2(s)\, ds - U_2(p) \int_{k_2}^{p} U_1(s)\, ds \right].$$

The right-hand side is, using $g''(p) = 2q_2$,

$$\begin{aligned}
V_{part}''(p) &= g''(p) - \frac{q_2}{\gamma}\left[U_1'(p) \int_{k_1}^{p} U_2(s)\, ds - U_2'(p) \int_{k_2}^{p} U_1(s)\, ds\right]' \\
&= 2q_2 - \frac{q_2}{\gamma}\left[U_1''(p) \int_{k_1}^{p} U_2(s)\, ds - U_2''(p) \int_{k_2}^{p} U_1(s)\, ds + U_1'(p) U_2(p) - U_2'(p) U_1(p)\right] \\
&= -\frac{q_2}{\gamma}\left[U_1''(p) \int_{k_1}^{p} U_2(s)\, ds - U_2''(p) \int_{k_2}^{p} U_1(s)\, ds\right],
\end{aligned}$$

where the last equation uses $U_1'(p)U_2(p) - U_2'(p)U_1(p) = 2\gamma$. Subtracting the right-hand side

from the left-hand side,

$$\frac{2\lambda\sigma^2}{p^2(1-p)^2}\left[V_{part}(p)-g(p)+\frac{c}{\lambda}\right]-V''_{part}(p)$$
$$=-\frac{q_2}{\gamma}\int_{k_1}^{p}U_2(s)\,ds\left[\frac{2\lambda\sigma^2}{p^2(1-p)^2}U_1(p)-U_1''(p)\right]+\frac{q_2}{\gamma}\int_{k_2}^{p}U_1(s)\,ds\left[\frac{2\lambda\sigma^2}{p^2(1-p)^2}U_2(p)-U_2''(p)\right]=0,$$

which holds because the square bracked terms are zero: $U_1(p)$ and $U_2(p)$ are roots of the homogeneous part of the ODE in (17).

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

$$V(p)=g(p)-\frac{c}{\lambda}-\frac{q_2}{\gamma}\left[U_1(p)\int_{k_1}^{p}U_2(s)\,ds-U_2(p)\int_{k_2}^{p}U_1(s)\,ds\right]+C_1U_1(p)+C_2U_2(p). \quad (41)$$

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

$$v(p)\equiv V(p)-g(p), \quad (42)$$

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, $c<\bar{c}$, 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 (41) 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

$$0 = V\left(\underline{p}\right) - g\left(\underline{p}\right) = -\frac{c}{\lambda} - \frac{q_2}{\gamma}\left[U_1\left(\underline{p}\right) I_2\left(\underline{p}\right) - U_2\left(\underline{p}\right) I_1\left(\underline{p}\right)\right] + C_1 U_1\left(\underline{p}\right) + C_2 U_2\left(\underline{p}\right), \tag{43}$$

$$\begin{aligned} 0 = V\left(1-\underline{p}\right) - g\left(1-\underline{p}\right) = &-\frac{c}{\lambda} - \frac{q_2}{\gamma}\left[U_1\left(1-\underline{p}\right) I_2\left(1-\underline{p}\right) - U_2\left(1-\underline{p}\right) I_1\left(1-\underline{p}\right)\right] \\ &+ C_1 U_1\left(1-\underline{p}\right) + C_2 U_2\left(1-\underline{p}\right). \end{aligned} \tag{44}$$

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

$$U_1(1-p) = U_2(p), \qquad U_2(1-p) = U_1(p), \tag{45}$$

$$U_1'(1-p) = -U_2'(p), \qquad U_2'(1-p) = -U_1'(p). \tag{46}$$

Using them in (44),

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

The second equality follows from changing integrand from s to $t = 1 - s$, and the third equality uses Eq. (43). 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 (43) and (44) imply

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

The smooth-pasting condition, obtained by differentiating (41), then pins down $\underline{p}$:

$$0 = V'\left(\underline{p}\right) - g'\left(\underline{p}\right) = -\frac{q_2}{\gamma}\left[U_1'\left(\underline{p}\right) I_2\left(\underline{p}\right) - U_2'\left(\underline{p}\right) I_1\left(\underline{p}\right)\right] + C\left(\underline{p}\right)\left[U_1'\left(\underline{p}\right) + U_2'\left(\underline{p}\right)\right]. \tag{48}$$

Substituting $C(\underline{p})$ from (47) into (48) yields

$$\frac{d}{dp}\left\{\frac{\frac{c}{\lambda}+\frac{q_2}{\gamma}\left[U_1(p)I_2(p)-U_2(p)I_1(p)\right]}{U_1(p)+U_2(p)}\right\}\Bigg|_{p=\underline{p}}=0, \quad \text{where } \underline{p}\in\left(0,\tfrac{1}{2}\right). \tag{49}$$

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 (49). Note that

$$\begin{aligned}&\operatorname{sgn}\left\{\frac{d}{dp}\left\{\frac{\frac{c}{\lambda}+\frac{q_2}{\gamma}\left[U_1(p)I_2(p)-U_2(p)I_1(p)\right]}{U_1(p)+U_2(p)}\right\}\right\}\\ =&\operatorname{sgn}\left\{\underbrace{\frac{q_2}{\gamma}\left(U_1'I_2-U_2'I_1\right)\left(U_1+U_2\right)-\left[\frac{c}{\lambda}+\frac{q_2}{\gamma}\left(U_1I_2-U_2I_1\right)\right]\left(U_1'+U_2'\right)}_{\equiv M(p)}\right\},\end{aligned} \tag{50}$$

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

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

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',

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

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

$$V(p) = g(p) - \frac{c}{\lambda} - \frac{q_2}{\gamma}\left[U_1(p) I_2(p) - U_2(p) I_1(p)\right] + \frac{\frac{c}{\lambda} + \frac{q_2}{\gamma}\left[U_1\left(\underline{p}\right) I_2\left(\underline{p}\right) - U_2\left(\underline{p}\right) I_1\left(\underline{p}\right)\right]}{U_1\left(\underline{p}\right) + U_2\left(\underline{p}\right)} \cdot \left[U_1(p) + U_2(p)\right], \tag{51}$$

where $\underline{p} \in \left(0, \frac{1}{2}\right)$ is uniquely determined by (49). □

A.4 Proof of Proposition 3

Proof. We first derive $\hat{p}(p; \kappa)$ as given in (26). By (13), the LLM's log-likelihood ratio evolves as $d\hat{z}_t = \kappa\, dz_t$. Integrating from 0 to t gives

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

Writing the log-likelihood ratios in terms of beliefs, $z_t = \ln \frac{p_t}{1-p_t}$ and $\hat{z}_t = \ln \frac{\hat{p}_t}{1-\hat{p}_t}$, this becomes

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

Exponentiating both sides and solving $\hat{p}_t = \frac{x}{1+x}$ for x equal to the right-hand side yields (26), giving the function $\hat{p}(p;\kappa)$ used throughout the proposition.

Second, AI memory enters the HJB equation only through its inhomogeneous term. In the continuation region, the HJB is

$$-\lambda V(p) + \frac{p^2(1-p)^2}{2\sigma^2} V''(p) + \lambda\, g(p, \hat{p}(p;\kappa)) - c = 0. \tag{52}$$

Since its homogeneous part is independent of κ and coincides with that of the baseline HJB, Step 1 of the proof of Proposition 1 implies that every solution of (52) takes the form $V(p) = Q(p;\kappa) + C_1 U_1(p) + C_2 U_2(p)$ for some particular solution $Q(p;\kappa)$.

Last, we construct a particular solution by variation of parameters. Rewriting (52) 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),

$$Q(p;\kappa) = \frac{1}{2\gamma} \left[U_1(p) \int_{1/2}^{p} U_2(s)\, h(s;\kappa)\, ds - U_2(p) \int_{1/2}^{p} U_1(s)\, h(s;\kappa)\, ds \right], \tag{53}$$

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 (53) are well defined on any compact subinterval of $(0,1)$, and one verifies directly that Q satisfies the inhomogeneous ODE.

□

A.5 Proof of Corollary 1

Proof. Fix $\kappa > 0$ and fix the other primitives. The proof has three steps. First, any positive memory creates the endpoint singularity. From Proposition 3,

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

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. □

A.6 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

$$-\lambda W(p) + \frac{p^2(1-p)^2}{2\sigma^2} W''(p) = -r_a(p; a). \tag{54}$$

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 (54), 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 (54) 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, $0 < \underline{p}(\kappa) < p_0 < \bar{p}(\kappa) < 1$.

Direct differentiation of g in (15) together with $\hat{p}(p;\kappa)$ in (26) yields

$$\begin{aligned} \operatorname{sgn}\left(\frac{\partial}{\partial \kappa} g(p, \hat{p}(p;\kappa))\right) &= \operatorname{sgn}\left((p-\hat{p})\Big[(1-p)(1-\hat{p}) + p\hat{p}\Big]\frac{\partial \hat{p}(p;\kappa)}{\partial \kappa}\right) \\ &= \operatorname{sgn}\Big((p - \hat{p}(p;\kappa))(z - z_0)\Big), \end{aligned} \tag{55}$$

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 (14) 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 (55),

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

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

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

Since stopping occurs at one of the two thresholds,

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

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

$$\begin{aligned} \frac{\partial V(p_0;\kappa)}{\partial\kappa} = {} & \lambda\,\mathbb{E}_{p_0}\left[\int_0^{\tau(\kappa)} e^{-\lambda t}\,\frac{\partial}{\partial\kappa} g(p_t, \hat{p}(p_t;\kappa))\,dt\right] \\ & + \mathbb{E}_{p_0}\left[e^{-\lambda\tau(\kappa)}\mathbf{1}\{p_{\tau(\kappa)} = \underline{p}(\kappa)\}\right] \frac{\partial}{\partial\kappa} g\Big(\underline{p}(\kappa), \hat{p}(\underline{p}(\kappa);\kappa)\Big) \\ & + \mathbb{E}_{p_0}\left[e^{-\lambda\tau(\kappa)}\mathbf{1}\{p_{\tau(\kappa)} = \bar{p}(\kappa)\}\right] \frac{\partial}{\partial\kappa} g(\bar{p}(\kappa), \hat{p}(\bar{p}(\kappa);\kappa)). \end{aligned} \tag{56}$$

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

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

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 (55),

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

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:

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

Substituting into (56),

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

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

A.7 Proof of Proposition 5

Proof. Using $\kappa = 1$ in (26) 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. □

A.8 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 (51) 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 $q_2(\hat{p})$:

$$\begin{aligned} V(p; \hat{p}) =& g(p; \hat{p}) - \frac{c}{\lambda} - \frac{q_2(\hat{p})}{\gamma} \left[U_1(p) I_2(p) - U_2(p) I_1(p)\right] \\ &+ \frac{\frac{c}{\lambda} + \frac{q_2(\hat{p})}{\gamma} \left[U_1\left(\underline{p}\right) I_2\left(\underline{p}(\hat{p})\right) - U_2\left(\underline{p}(\hat{p})\right) I_1\left(\underline{p}(\hat{p})\right)\right]}{U_1\left(\underline{p}(\hat{p})\right) + U_2\left(\underline{p}(\hat{p})\right)} \cdot \left[U_1(p) + U_2(p)\right] \end{aligned} \tag{57}$$

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 $\hat{p}$,

$$\begin{aligned} \frac{dV(p; \hat{p})}{d\hat{p}} =& \frac{dg(p; \hat{p})}{d\hat{p}} - \frac{q_2'(\hat{p})}{\gamma} H(p) + \frac{q_2'(\hat{p})}{\gamma} \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))} \cdot S(p) + \underbrace{\frac{d}{d\underline{p}} \left\{ \frac{\frac{c}{\lambda} + \frac{q_2(\hat{p})}{\gamma} H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))} \right\}}_{=0, \text{ (Envelope Theorem)}} \cdot S(p) \cdot \frac{d\underline{p}(\hat{p})}{d\hat{p}} \\ =& \underbrace{\frac{dg(p; \hat{p})}{d\hat{p}}}_{\text{stopping-value effect}} - \underbrace{\frac{q_2'(\hat{p})}{\gamma} S(p) \left[\frac{H(p)}{S(p)} - \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))}\right]}_{\text{option-value effect}}, \end{aligned} \tag{58}$$

where we used the Envelope Theorem so that the indirect effect through $\underline{p}$ is zero (see (49).) 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

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

where $q_2(\hat{p}) = \Delta_\mu^2 + 2 - \frac{1}{\hat{p}^2+(1-\hat{p})^2}$. Using these derivatives in (58),

$$\frac{dV(p;\hat{p})}{d\hat{p}} = \frac{2(p-\hat{p})\Big[(1-p)(1-\hat{p})+p\hat{p}\Big]}{\left[\hat{p}^2+(1-\hat{p})^2\right]^2} - \frac{4\hat{p}-2}{\gamma\left[\hat{p}^2+(1-\hat{p})^2\right]^2} S(p)\left[\frac{H(p)}{S(p)} - \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))}\right]. \quad (59)$$

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 (59) gives

$$\frac{dV(\frac{1}{2};\hat{p})}{d\hat{p}} = \frac{4(\frac{1}{2}-\hat{p})}{\gamma\left[\hat{p}^2+(1-\hat{p})^2\right]^2}\left[\frac{\gamma}{4} - \frac{H(\underline{p}(\hat{p}))}{S(\underline{p}(\hat{p}))}\right]. \quad (60)$$

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 (60), 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 (57), 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 (58) 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 (58) 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 $\hat{p}^*(p_0) < p_0$.

□

B Appendices for Empirical Exercise

B.1 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
   (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.