Ex Machina: Financial Stability in the Age of Artificial Intelligence

Kartik Anand* Sophia Kazinnik†
Agnese Leonello‡ Ettore Panetti§

Abstract

Does artificial intelligence (AI) pose a threat to financial stability? This paper develops a simulation-based framework to study how AI agents behave in a mutual-fund redemption game with strategic complementarities and multiple equilibria. Different AI technologies, namely Q-learning (QL) algorithms and large language models (LLMs), generate distinct redemption profiles. QL-investors coordinate among themselves but exhibit a bias toward excessive early redemption that amplifies fund fragility. LLM-investors instead internalize the equilibrium structure of the problem and better align with theoretical predictions. However, their belief heterogeneity weakens coordination, thereby making their redemptions less predictable. Thus, our findings highlight that the design of AI systems is material for financial stability.

Keywords: Financial stability; strategic uncertainty; coordination games; reinforcement learning; Q-learning; large language models; AI agents.
JEL classifications: G01, G23, C63.

*Deutsche Bundesbank; kartik.anand@bundesbank.de
†Stanford University, HAI; kazinnik@stanford.edu
‡European Central Bank, DG-Research and CEPR; agnese.leonello@ecb.europa.eu
§University of Naples Federico II, CSEF, SUERF, UECE; ettore.panetti@unina.it

We thank Markus Brunnermeier, Emilio Calvano, Mehmet Ekmekci, Itay Goldstein, John J. Horton, and seminar audiences at the Deutsche Bundesbank for useful comments. The views expressed here are the authors' and do not reflect those of the European Central Bank, the Deutsche Bundesbank, or the Eurosystem.

I. Introduction

In ancient Greek tragedy, when a plot reached an impasse, a mechanical crane would lower actors playing gods onto the stage to impose a resolution, a device later known as *deus ex machina*. Two and a half millennia later, our growing reliance on artificial intelligence (AI) to navigate financial complexity and uncertainty invites a comparison. Will AI become the deus ex machina of modern finance, or will it instead threaten financial stability? This paper investigates this question.

AI is playing an increasingly central role in modern finance. For example, the release of a large language model (LLM) by DeepSeek has spurred an AI arms race amongst other hedge funds and mutual funds in China to automate the processing of market data and the generation of trading signals (Reuters, 2025). This trend extends to the retail sector, where industry reports suggest that AI-enabled applications that autonomously generate financial advice will be the leading source for investors by 2027 (MIT Sloan, 2024; Deloitte, 2024). The frontier is now shifting toward fully agentic AI, autonomously executing tasks, adapting to new information, and coordinating with other systems under minimal human oversight in areas like fraud detection and transaction management (Nvidia, 2025; Financial Times, 2025). This clear progression towards more sophisticated autonomous AI raises a central question: How will these systems behave in complex environments with strategic and economic uncertainty?

A growing literature offers some insights. In pricing games, Q-learning algorithms have been shown to sustain tacit collusion, even when standard theory predicts competitive outcomes (e.g., Calvano et al., 2020; Colliard et al., 2025; Dou et al., 2025). This shows that AI can learn sophisticated, nontrivial strategies without explicit communication or intent. Yet, most of these studies focus on settings with a unique equilibrium, where learning places limited demands on strategic reasoning. Events such as bank runs and financial market freezes are, in contrast, shaped by multiple equilibria driven by strategic complementarities: investors' actions depend on their beliefs about others' behavior (Diamond and Dybvig, 1983). In such environments, panic-driven crises can emerge with large economic costs (Reinhart and Rogoff, 2009). Furthermore, because beliefs respond to fundamentals, equilibrium selection varies with economic conditions (Goldstein and Pauzner, 2005).

This paper studies how different types of AI agents behave in such environments

and the implications for financial stability. To this end, we consider a stylized model of mutual fund redemptions based on Chen et al. (2010), in which AI agents decide whether or not to redeem their shares. Using a simulation-based experimental setup, we show that reinforcement learning agents tend to converge on coordinated strategies, but their learning dynamics bias them toward redeeming their shares, so they often coordinate on inefficient outcomes characterized by excessively high asset liquidations. By contrast, reasoning-oriented LLM agents use context: rather than relying solely on trial-and-error, they interpret the game structure and adapt accordingly. This keeps them close to theoretical benchmarks and prevents systematic over-redemption, but heterogeneity in their beliefs hinders their ability to coordinate on equilibrium outcomes. Our result thus highlights that the design of AI systems matters for financial stability: reinforcement learning promotes coordination but at the expense of financial stability, whereas LLM reasoning improves stability, but heterogeneity in beliefs makes the outcomes more difficult to predict.

Framework We build on a canonical mutual fund run model (Chen et al., 2010), a parsimonious variant of the models in Diamond and Dybvig (1983) and Goldstein and Pauzner (2005), that delivers multiple redemption equilibria.[1] The model yields theoretical predictions about equilibrium behavior, which we use as a benchmark for evaluating AI agents' decisions and their effects on financial stability.

In the model, investors choose whether to redeem their shares before the fund's investment project matures. Their payoffs depend on both economic fundamentals and the actions of other investors. Strategic complementarities imply that the incentive to redeem rises when more investors do so, while stronger fundamentals reduce the appeal to redeem. We examine equilibria under two dimensions of uncertainty: (i) payoff uncertainty, contrasting whether the fund's returns are risky or risk-free, and (ii) fundamental uncertainty, where investors may or may not observe the true underlying fundamentals.

With these results in hand, we design an experiment that replaces investors with AI agents to test whether they replicate the model's equilibrium predictions. We use two types of agents: Q-learning algorithms that learn through trial and error, and large language model (LLM) that make decisions using contextual understanding and

[1] While our focus is on mutual fund redemptions, the insights extend to other environments with similar strategic complementarities and payoff structures, such as bank runs and currency crises.

chain-of-thought reasoning. The experiment is built around four core predictions from the theoretical model. When there is no uncertainty, investors should redeem when fundamentals are weak, stay when fundamentals are strong, and exhibit multiple possible outcomes when fundamentals are in between, i.e., either everyone redeems or everyone stays. Second, the share of redemptions should not be affected by the payoff uncertainty since only expected returns matter. Third, with fundamental uncertainty, equilibrium behavior should take the form of threshold strategies, with investors redeeming whenever their private signal falls below a critical cutoff. Finally, financial fragility, defined as early inefficient redemptions in excess of the first-best level, should increase as markets become more illiquid.

Results We find consistent differences between reinforcement-learning and reasoning-based AI agents in environments with strategic complementarities.

First, in the absence of uncertainty, both types of investors reproduce the dominance regions predicted by theory: all redeem when fundamentals are weak, all stay when they are strong. However, in the intermediate range, behavior diverges. Q-learning (QL)-investors converge to a sharp threshold rule, effectively collapsing multiplicity into the risk-dominant equilibrium. In contrast, LLM-investors, who have heterogeneous beliefs about the actions of other investors, exhibit less coordination on equilibrium outcomes, resulting in partial redemptions.

Second, on introducing payoff uncertainty, LLM-investors continue to align with the theoretical benchmark: by reasoning in expected-value terms, they treat the environment as equivalent to one without payoff uncertainty. QL-investors, however, display a systematic bias toward redemption. Their learning dynamics penalize staying whenever a zero return is realized, which drives Q-values below the true expected payoff and pushes them toward inefficient redemptions.

Third, with fundamental uncertainty, LLM-investors recognize the problem as a global game and coordinate around the theoretical panic threshold. QL-investors, by contrast, exhibit a over-redemption when signals are noisy. Heterogeneity in private signals generates disagreement that lowers the payoff from staying, thus reinforcing the bias toward redemption. However, unlike the scenario with payoff uncertainty, their behavior converges to persistent partial redemptions—an equilibrium outcome consistent with theory, but inefficient relative to the first best.

Finally, for LLM-investors, fragility rises monotonically with illiquidity, thus closely

tracking the theoretical benchmark. For QL-investors, the relationship is more complex: with precise signals and no payoff uncertainty, outcomes align with theory; with noisy signals or payoff uncertainty, fragility increases, and its sensitivity to illiquidity depends on the interaction of these two sources of uncertainty.

Taken together, these results highlight a central tradeoff: reinforcement learning promotes coordination but at the expense of financial stability, whereas LLM reasoning improves stability by anchoring to theoretical benchmarks, but reduces coordination through heterogeneous beliefs. A key implication is that the design of AI systems is not neutral – different approaches shape market dynamics in systematically different ways, and therefore the stability of the financial system critically depends on how AI agents are built.

Literature Our paper contributes to several strands of literature at the intersection of AI and finance. First, it advances the emerging work on AI and financial stability, a topic that has attracted growing attention from policymakers (Shabsigh and Boukherouaa, 2023; Aldasoro et al., 2024; Financial Stability Board, 2024; Leitner et al., 2024; Bank of England, 2025) but has only recently begun to receive systematic treatment in academia. Notable contributions include Danielsson et al. (2022) and Danielsson and Uthemann (2025), who examine channels through which AI may generate or amplify systemic risk and crises. We move this agenda forward by providing an experimental, model-grounded assessment of AI in a canonical coordination game with multiple equilibria. In a multiagent mutual fund framework, we compare distinct AI technologies across economic conditions that vary payoff risk and information. This design allows us to isolate the key economic drivers that shape Q-learning behavior and LLM reasoning, and to map how these drivers determine equilibrium selection and the level of fragility. Our study is related to Yang (2024), who shows that Q-learning can trigger currency attacks in a two-agent setting under alternative information structures. Relative to that work, we focus on a mutual fund withdrawal game with many interacting agents, and we conduct a head-to-head comparison of AI architectures to identify how technology choice and environment interact to produce fragility.

We also add to the growing literature that uses AI simulations in financial settings. Several papers use reinforcement learning to study asset pricing and trading: portfolio choice (Cong et al., 2023), liquidity provision (Colliard et al., 2025), and speculation

(Dou et al., 2025). We shift the focus from performance to mechanism by examining what drives AI agents' choices and what that implies for financial stability.

More broadly, our paper contributes to the economics of AI. Beginning with Calvano et al. (2020), much of this literature has examined pricing algorithms and documented emerging collusion in a variety of settings (Banchio and Mantegazza, 2022; Banchio and Skrzypacz, 2022; Cont and Xiong, 2024). Collusive outcomes also appear in many of the finance applications noted above, including our own. In our setting, however, collusion has a more nuanced interpretation; it is a mechanism through which AI agents resolve coordination failures. This interpretation links algorithmic "collusion" to equilibrium selection rather than illicit intent, and it helps explain why machine behavior can be predictable even in environments that admit multiple equilibria.

Along similar lines, we add to the growing literature that examines LLM agents as decision-makers in economic and strategic contexts. Recent work shows that LLMs can participate in strategic experiments (Horton, 2023), sustain cooperation in repeated games (Akata et al., 2023), and serve as stand-ins for human subjects with promising fidelity (Anthis et al., 2025). With appropriate prompts, LLMs can emulate investors (Fedyk et al., 2024), bank depositors (Kazinnik, 2023), and loan officers (Cook and Kazinnik, 2025); they can also serve as proxies for survey respondents (Hansen et al., 2024) and replicate human-like macroeconomic expectations (Bybee, 2023; Zarifhonarvar, 2024). Within finance, Lopez-Lira (2025) introduce an open-source framework that pits LLM trading agents against value investors, momentum traders, market makers, and contrarians, while Gao et al. (2024) show that heterogeneous LLM-investors produce price dynamics that mirror observed markets. Bhagwat et al. (2025) elicit beliefs from LLM-investor personas and find that disagreement about firm news both tracks abnormal trading volume and predicts a subsequent return premium. We build on these advances by placing LLMs in a classic withdrawal game and comparing their behavior with that of reinforcement learners, thereby clarifying how explicit reasoning, as opposed to implicit learning, affects equilibrium selection and fragility.

Finally, we contribute to the literature on financial fragility. Following the seminal work of Diamond and Dybvig (1983), this literature has extended to other financial institutions subject to strategic uncertainty, such as mutual funds (Chen et al., 2010), hedge funds (Liu and Mello, 2011), credit markets (Bebchuk and Goldstein, 2011),

life insurers (Foley-Fisher et al., 2020) and stablecoin issuers (Gorton et al., 2025), among others. Despite rich theory, credible empirical evidence remains scarce, with Goldstein et al. (2017) and Chen et al. (2024) as notable exceptions. We open a new research avenue by experimentally testing theories of financial fragility using AI agents, allowing fine-grained control over information, payoffs, and strategic interaction, and showing clear links among market fundamentals, algorithmic design, and systemic risk.

The rest of the paper proceeds as follows. In Section II, we present a stylized theoretical framework characterized by strategic complementarities and derive the prediction that we test in the experiments with Q-learning algorithms and LLMs. In Section III, we describe the experimental environment for both AI types. The results of the simulations are reported in Section IV. Section V concludes.

II. Theoretical Framework

In this section, we characterize equilibria in a stylized coordination game building on Chen et al. (2010), and derive testable predictions. Proofs appear in Appendix A. The economy extends over two dates, $t = 1$ and $t = 2$, and consists of a mutual fund and $N > 2$ risk-neutral investors. Before $t = 1$, each investor holds one share of unit value, so the fund's total size is N.

The book value of the fund's investments at $t = 1$ is $R_1 N$, which is common knowledge. At $t = 2$, the per-unit investment return is $R_2(\theta) = R\theta$. This specification captures an equity fund, in which investors hold a pro rata claim on the portfolio's period-2 payoff. The variable θ represents the fundamental of the economy and positively affects fund's investment return at $t = 2$.[2] It is uniform on $[0, 1]$ and drawn at the beginning of $t = 1$.[3] In what follows, we distinguish between two informational environments regarding the fundamental: (i) no fundamental uncertainty, where θ is common knowledge; and (ii) fundamental uncertainty, where each investor, i, receives a noisy private signal $s_i = \theta + \epsilon_i$, with ϵ_i i.i.d. across agents and uniformly distributed over the interval $[-\eta, \eta]$.

[2] In Chen et al. (2010) θ is a measure of the fund performance. The two definitions are clearly linked, as better fundamentals can be associated with improved fund performance. For the purpose of our exercise, the precise definition is immaterial.

[3] Assuming uniformity is without loss for our results: they hold for any strictly increasing mapping $R(\theta)$.

At the beginning of $t = 1$, a number $A < N$ of investors are *active*, and choose whether to redeem their shares or stay until $t = 2$. Investors who redeem their shares receive the current value R_1. However, in order to service redemptions, the fund must liquidate assets, which is costly. Specifically, to raise R_1, the fund must liquidate $(1 + \lambda)R_1$ units of assets, where $\lambda > 0$ is a measure of illiquidity: the higher λ, the greater the share of assets that must be liquidated to service redemptions. Following Chen et al. (2010), we assume that $A \leq \frac{N}{1+\lambda}$, implying that the mutual fund has enough resources to meet the redemptions of all the A active investors at $t = 1$. Thus, all investors who redeem their shares receive R_1 for sure.[4]

Denoting the number of investors who redeem by W, an investor who decided to stay receives

$$\frac{N - W(1 + \lambda)}{N - W} R_1 R\theta \,. \tag{1}$$

The payoff from staying is increasing in θ, but is decreasing in both asset illiquidity λ and the number of redemptions, W.

A. Characterizing the Equilibria

We focus on characterizing equilibria in pure strategies. As a benchmark, we first consider the case without fundamental uncertainty when θ is common knowledge. Following standard lines of reasoning (Morris and Shin, 1998), we can partition the space of the fundamental into three intervals, which we report in Proposition 1 and Figure 1.

Proposition 1: *In the absence of fundamental uncertainty, equilibrium outcomes depend θ. For sufficiently extreme values of θ, there exists a unique Nash equilibrium in pure strategies, while for intermediate values multiple equilibria arise. Specifically,*

- *if $\theta < \underline{\theta} = \frac{1}{R}$, the unique equilibrium is that all investors redeem at $t = 1$;*
- *if $\theta > \overline{\theta} = \frac{1}{R}\frac{N-(A-1)}{N-(A-1)(1+\lambda)}$, the unique equilibrium is that all investors stay until $t = 2$;*
- *if $\theta \in [\underline{\theta}, \overline{\theta}]$, both "all redeem" and "all stay" are equilibria.*

An important prerequisite for the emergence of multiple equilibria is illiquidity. When the portfolio is perfectly liquid, i.e., $\lambda = 0$, asset liquidation does not impose

[4]The derivations for the case with illiquidity, i.e, when $A > \frac{N}{1+\lambda}$, are in Appendix C.

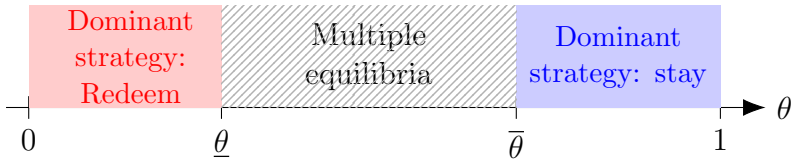


Figure 1: Equilibria without Fundamental Uncertainty

additional costs on the fund and investors who stay. The expected return from staying is independent of early redemptions. This implies that redeeming at $t = 1$ is the strictly dominant action for $\theta < \underline{\theta}$, while staying is strictly dominant for $\theta > \underline{\theta}$. From a welfare perspective, this is the efficient redemption strategy, since only negative NPV investments are liquidated. Hence, the threshold $\underline{\theta}$ also denotes the first-best redemption strategy.

In contrast, when $\lambda > 0$, redemptions no longer reflect solely the liquidation of unprofitable assets. In this case, the expected payoff difference between staying until $t = 2$ and redeeming at $t = 1$ decreases monotonically with the number of investors who redeem. As a result, even when fundamentals are sound (i.e., above $\underline{\theta}$), investors have an incentive to redeem if they expect others to do so. Put differently, redemption decisions are strategic complements: the incentive to redeem strengthens when others are expected to redeem. This strategic complementarity generates multiple equilibria in the intermediate region $[\underline{\theta}, \overline{\theta}]$.

Next, we consider the case with fundamental uncertainty, i.e., investors do not observe the realized θ but instead receive noisy signals. We follow the global game literature (Carlsson and Van Damme, 1993; Goldstein and Pauzner, 2005) and study how investors make their redemption decisions based on the signals. Proposition 2 characterizes the equilibrium, which is also illustrated in Figure 2.

Proposition 2: *With fundamental uncertainty, there exists a unique symmetric Bayes-Nash equilibrium in threshold strategies that is characterized by the threshold signal*

$$\theta^* \equiv \frac{A}{\sum_{W=0}^{A-1} \frac{N-W(1+\lambda)}{N-W} R}, \tag{2}$$

such that an investor redeems if and only if his signal is below $\theta^* \in (\underline{\theta}, \bar{\theta})$. For a given realization of the fundamental θ, the share of investors who redeem their shares at

$t = 1$ *is*

$$w^*(\theta, \theta^*) = \begin{cases} 0 & \text{if} \quad \theta > \theta^* + \eta \\ \frac{\theta^* - \theta + \eta}{2\eta} & \text{if} \quad \theta \in [\theta^* - \eta,\, \theta^* + \eta] \\ 1 & \text{if} \quad \theta < \theta^* - \eta \end{cases}. \tag{3}$$

The introduction of private information yields a unique equilibrium characterized by a threshold signal θ^*. This threshold determines the probability of redemption across investors. For fundamentals in the range $\underline{\theta} < \theta < \theta^*$, redemptions are driven by panic rather than weak fundamentals and are therefore inefficient. In this region, investors redeem not because redemption is a dominant strategy, but because the fundamental is weak enough to make each investor fear that others will redeem as well.

The threshold signal θ^*, and in turn the average number of investors redeeming their share, depends on the underlying economic parameters. In particular, θ^* increases with the illiquidity parameter λ. When λ is large, the fund must liquidate more assets to meet redemptions at $t = 1$, which reduces the payoff from staying. This strengthens investors' incentive to redeem early, thereby raising θ^*.

The results in Propositions 1 and 2 are derived under a specific assumption about the investment return at $t = 2$. As discussed above, the specification $R_2(\theta) = R\theta$ corresponds to the case of an equity fund. An alternative is a fund investing primarily in debt securities, i.e., a bond fund. In that case, the return at $t = 2$ takes the form

$$R_2(\theta) = \begin{cases} R & \text{with prob } \theta \\ 0 & \text{with prob } 1 - \theta \end{cases} \tag{4}$$

Relative to the equity-fund specification, this introduces payoff uncertainty: investors may receive zero even when fundamentals are favorable. By contrast, under the equity-fund specification there is no payoff uncertainty, since the payoff is deterministic given θ and is strictly positive whenever $\theta > 0$. Proposition 3 establishes that this distinction does not affect the theoretical results.

Proposition 3: *The results in Propositions 1 and 2 remain unchanged when the fund's return at* $t = 2$ *is given by Equation* (4).

This irrelevance follows from investors' risk neutrality: only the expected payoff matters for redemption decisions, and this expectation is identical across the two

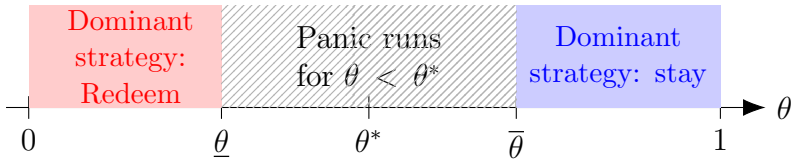


Figure 2: Equilibria with Fundamental Uncertainty

specifications and equal to $R/2$. What differs is the presence or absence of payoff uncertainty. In the equity-fund case (without payoff uncertainty), investors receive zero only when $\theta = 0$. In the bond-fund case (with payoff uncertainty), investors may receive zero even when fundamentals are strong.

B. Testable Predictions

Based on the characterization of the equilibria, we draw the following predictions. With common knowledge about θ, Proposition 1 highlights the tri-partite classification of the fundamental, with multiple equilibria emerging for intermediate values of θ. Multiple equilibria emerge because investors' beliefs over others are indeterminate.

Prediction 1: *Without fundamental uncertainty, all investors: (i) redeem their shares when $\theta < \underline{\theta}$ and (ii) stay when $\theta > \overline{\theta}$. In the intermediate range, both "all redeem" and "all stay" are equilibria.*

Since the characterization of the equilibria only depends on the expected return at $t = 2$, payoff uncertainty is irrelevant, as highlighted in Proposition 3. This result is the basis for our next prediction.

Prediction 2: *Investors' redemption flows are not impacted by the introduction of payoff uncertainty.*

With fundamental uncertainty, Proposition 2 shows that there is a unique signal threshold equilibrium characterized by θ^* and, for any given θ, the share of investors who redeem is given by $w^*(\theta, \theta^*)$. Our next prediction concerns the signal threshold and how signal precision impacts investors' redemptions.

Prediction 3: *With fundamental uncertainty, investors use threshold strategies, i.e., redeem whenever their private signal is below the critical threshold θ^*. The share of investors who redeem is given by $w^*(\theta, \theta^*)$, which is decreasing in signal precision.*

Redemptions are inefficient unless fundamentals θ are so low that redeeming is a strictly dominant action. This motivates our definition of *fragility* as the extent of inefficient, panic-driven redemptions. With fundamental uncertainty, fragility is measured by the expected number of redemptions that occur when $\theta > \underline{\theta}$. Formally, this corresponds to the share of early redemptions $w^*(\theta, \theta^*)$ defined in Equation (3). The cutoff $\underline{\theta}$ serves as the first-best benchmark: above $\underline{\theta}$, efficient behavior would entail no redemptions, so any observed redemptions reflect fragility. Since $\underline{\theta}$ is independent of λ, while θ^* (and thus $w^*(\theta, \theta^*)$) increases with the liquidation cost λ, we obtain the following prediction.

Prediction 4: *Fragility increases monotonically with fund illiquidity.*

III. Experimental Design

Our primary goal is to explore the impact of AI-based investors on financial stability. To this end, we construct a simulation-based experimental setup where we replace the investors from the theoretical model with algorithmic counterparts. We consider two distinct types of AI-based investors: Q-learning (QL) investors, who optimize behavior through trial-and-error and reward updating; and LLM-investors, who reason contextually using chain-of-thought inference.

To ensure comparability, both types of investors are evaluated under the same economic environment. There are $A = 30$ active investors out of a total of $N = 50$. The investment returns are $R_1 = 1$ and $R = 2$. The illiquidity parameter λ is drawn from an equally spaced seven-point grid spanning $[0.05, 0.35]$. The fundamental θ is drawn from an equally spaced twenty-five-point grid spanning $[0.4, 1]$, a range that includes both dominance bounds $\underline{\theta}$ and $\overline{\theta}$.

A. Q-learning Algorithm

We first consider QL-investors, implemented via the Q-learning algorithm. Q-learning is a standard reinforcement-learning method in which agents, without prior knowledge of the payoff structure, learn action values through trial-and-error interaction with the environment. By repeatedly updating these values across episodes, agents eventually converge toward a stable policy that approximates optimal behavior.

The algorithm has four components:

1. **States**: The state θ_i for each QL-investor i is derived from the information that they receive, i.e., either the true fundamental θ or the noisy signal s_i. As Q-learning requires a finite state space, this continuous information is mapped to a discrete classification, which constitutes the effective state of the algorithm. The set of all states is denoted $\Theta \subset \mathbb{N}$.

2. **Actions**: The action set is binary, $\mathcal{A} = \{a_R, a_S\}$, where a_R denotes "redeem" and a_S denotes "stay."

3. **Rewards**: The reward function $\pi(\theta, a)$ assigns a payoff based on the realization of the fundamental θ and the chosen action a. Crucially, QL-investors do not observe the underlying mapping $(\theta, a) \mapsto \pi$ but only the realized rewards.

4. **Episodes**: Learning unfolds over T episodes. In each episode, QL-investors observe their state, choose an action, and update their Q-values based on the realized payoff.

Each QL-investor $i = 1, \ldots, A$ starts episode $t = 1, \ldots, T$ with a Q-matrix $Q_{i,t} \in \mathbb{R}^{|\Theta| \times 2}$, with rows for states and columns for actions. Following the realization of the state $\theta_{i,t}$, the action is determined according to an ε-greedy policy: with probability $\varepsilon_t = \beta^t$, the QL-investor explores by randomizing between a_R and a_S, while with the complementary probability $1 - \varepsilon_t$, the QL-investor exploits by choosing the action that maximizes $Q_{i,t}(\theta_{i,t}, a)$.

If the chosen action is $a^*_{i,t} = a_R$, the reward is $\pi(\theta, a_R) = R_1$. If instead $a^*_{i,t} = a_S$, the reward depends on θ (the true fundamental) and the other QL-investors' actions and is given by Equation (1). The Q-matrix is then updated as

$$Q_{i,t+1}(\theta_{i,t}, a^*_{i,t}) = (1 - \alpha) Q_{i,t}(\theta_{i,t}, a^*_{i,t}) + \alpha \pi(\theta, a^*_{i,t}), \tag{5}$$

where $\alpha \in [0, 1]$ is the learning rate: higher α places more weight on new rewards, while lower α smooths learning over past experience. Figure 3 summarizes the iterative process.

In our simulations, we set $\beta = 0.99999$, $\alpha = 0.1$ and $|\Theta| = 75$. We run 25 independent rounds of training, each with $T = 500,000$ episodes. For these hyperparameters, each QL-investor experiments approximately 100,000 times in expectation, which corresponds roughly to 1/5 of total episodes.[5]

[5]Consistently with Colliard et al. (2025), we fix the hyperparameters and only perform compar-

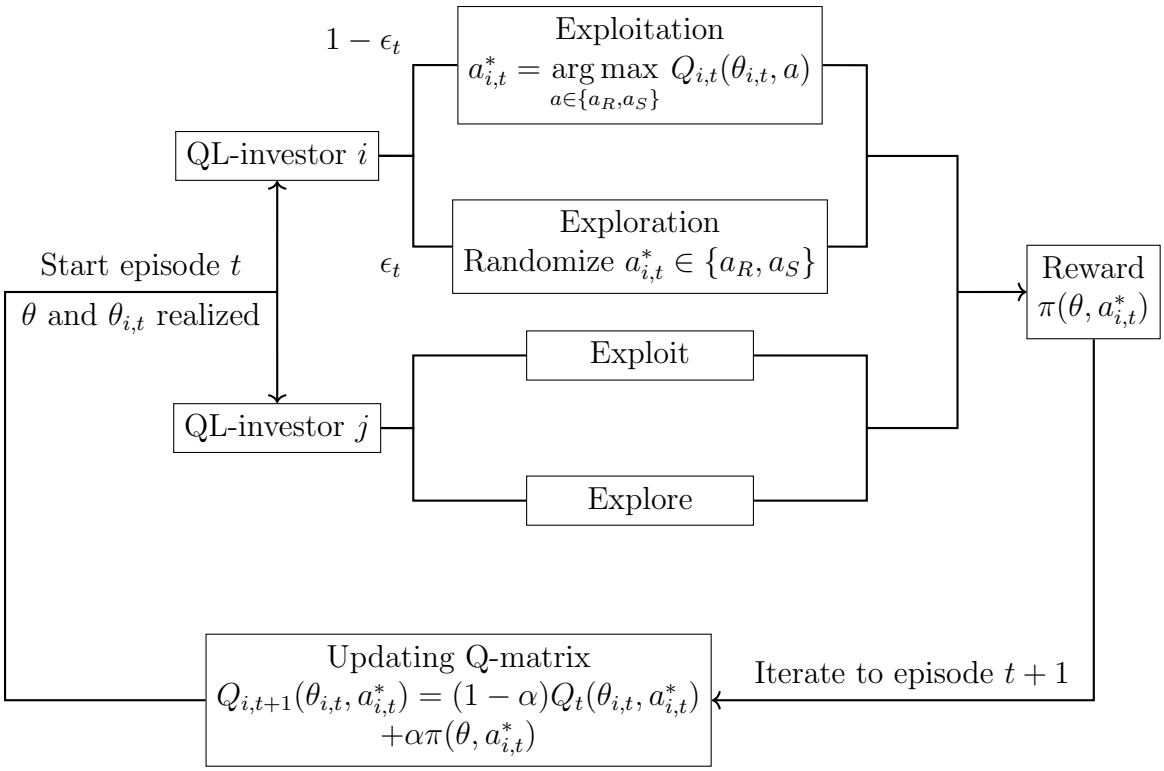


Figure 3: Iterative Process of Q-learning

B. Reasoning LLMs

Large language models (LLMs) are designed to predict the next word (or "token") in a sequence of text. Unlike the Q-learning algorithm, where parameters are updated through repeated trial-and-error, LLMs do not change their parameters when generating outputs. Instead, they rely on context-based inference: they read the information given in a prompt, identify relevant patterns, and use this to produce their response. Thus, rather than learning gradually through experience, LLMs adapt their behavior directly from the environment they are presented with.

Here, we consider LLM-investors, whose decision process differs fundamentally from QL-investors. Instead of learning through numerical rewards, LLM-investors rely on context-based inference to decide whether to redeem or stay.

We use DeepSeek's state-of-the-art reasoning model, `R1`. A recent class of large

ative statics exercises with respect to economic variables (e.g., λ). Modifying the hyperparameters in a way that is meaningful from an economic perspective would involve not only taking a stance on what their optimal value is, but also making assumptions about the preferences and objective functions of the agents coding the algorithms.

language models has been designed to *deliberate*, breaking down problems into intermediate steps and evaluating their reasoning before producing an answer. This is typically implemented through chain-of-thought style traces, verifier loops, and reflection. Like other large language models, `R1` is based on the transformer architecture and generates text by predicting the next token. Reasoning models build on this foundation by explicitly optimizing for step-by-step decomposition and self-verification using the chain of thought approach. Several such models are now available, including OpenAI's `o3` and Google's `Gemini 2.5 Pro`. A key advantage of DeepSeek's `R1` is that it provides direct access to the full chain-of-thought, enabling us to reconstruct the "mental model" behind each investor's decision. It is also an open-source model, supporting transparency and replicability.[6]

The experiment unfolds in two stages. First, we design the prompts. Each prompt distills the economic environment into a concise, textual description that specifies the investor's objective, the structure of payoffs, and the information available. The baseline version (Prompt 1) presents a world without uncertainty, neither in fundamentals nor in payoffs. To introduce payoff uncertainty, we modify line 3, as shown in Prompt 2. To introduce fundamental uncertainty and private signals, we append additional information after line 9, as detailed in Prompt 3.

Next, we run the model. Each prompt is instantiated with the relevant parameters and submitted to the DeepSeek `R1` model via an API call. Every investor is represented by a separate, independent call to the model. To encourage consistent behavior, we set the temperature hyperparameter to zero, reducing randomness in the outputs.[7] For each AI investor, we collect both the generated explanation and the final decision, which together determine the payoff.

[6] For a broader discussion on the use and evaluation of open-source LLMs, see Cook et al. (2023).

[7] The temperature setting controls the degree of randomness in the model's responses: a value near zero yields deterministic and focused outputs, while higher values introduce more variation.

```
You are one of A=\%d active investors in a mutual fund out of a total of N=\%d investors. Each investor holds one share. Your goal is to maximize your return.
If you redeem your share, then you earn R1.
If you do not redeem your share, i.e., you stay, then you earn fraction * R1 * R * \theta, where:
 - fraction = (N - W * (1 + λ)) / (N - W) is the fraction of assets remaining in the fund after serving redemptions;
 - W ≤ A-1 is the number of other active investors who redeem (the remaining N - A investors are passive and never redeem);
 - λ (lambda) = \%f is the illiquidity parameter;
 - R1 = \%f is the value of the share if you redeem;
 - R = \%f is the return earned by the fund from managing its portfolio;
 - θ (theta) = \%f is the fundamental, which measures the fund's performance.
Do you choose to redeem your share or stay? State your decision  with exactly one word: ''redeem'' or ''stay'' using the XML tag <decision>...</decision>.
```

Prompt 1: No Fundamental Uncertainty or Payoff Uncertainty

```
If you do not redeem your share (i.e., you stay), then with probability θ you earn fraction * R1 * R, and otherwise you get 0, where:
```

Prompt 2: Payoff Uncertainty Add-on

```
10 The fundamental value θ (theta) of the fund is randomly
     drawn from the interval [0,1] but you do not directly
     observe it. All values of θ are equally likely.
11 Instead you receive a private signal x_i defined as x_i =
     θ + ϵ_i, where ϵ_i is drawn uniformly from [−η, η] with η
     = \%f. Signals of different investors are drawn
     independently. Your private signal is x_i = \%f.
12 Do you choose to redeem your share or stay? State your
     decision  with exactly one word: ‘‘redeem’’ or ‘‘stay’’
      using the XML tag <decision>...</decision>.
```

Prompt 3: Fundamental Uncertainty Add-on

We subsequently iterate for the values of the illiquidity parameter and fundamental from the prespecified grid. We repeat this process for three independent rounds to generate statistical properties and construct confidence intervals around the outcomes.

To analyze the decision-making of LLM-investors, we use a second LLM as an analyst that converts the investors' explanation into directed acyclic graphs (DAGs).[8] The analyst LLM decomposes the explanations into three sub-graphs: context (inputs such as the economic environment, number of players, and payoff rules), reasoning (the intermediate logical or computational steps), and decision (the final action). We then compile the extracted variables, assumptions, and equations into a DAG that encodes functional dependence (nodes) and information flow (directed edges).

IV. Experimental Results

In this section, we report the results of our experiments with QL and LLM-investors to test our theoretical predictions. Before exploring the experiments in detail, it is instructive to consider the behavior of a single investor without fundamental or payoff uncertainty in order to isolate and highlight the role of strategic uncertainty. Figure 4 plots the withdrawal decision of a QL-investor and a LLM-investor as a function of the

[8] A DAG consists of nodes (variables, computed quantities, or decisions) and directed edges indicating the direction of influence or information flow. For a recent example of using DAGs to analyze LLM reasoning, see Bybee (2023).

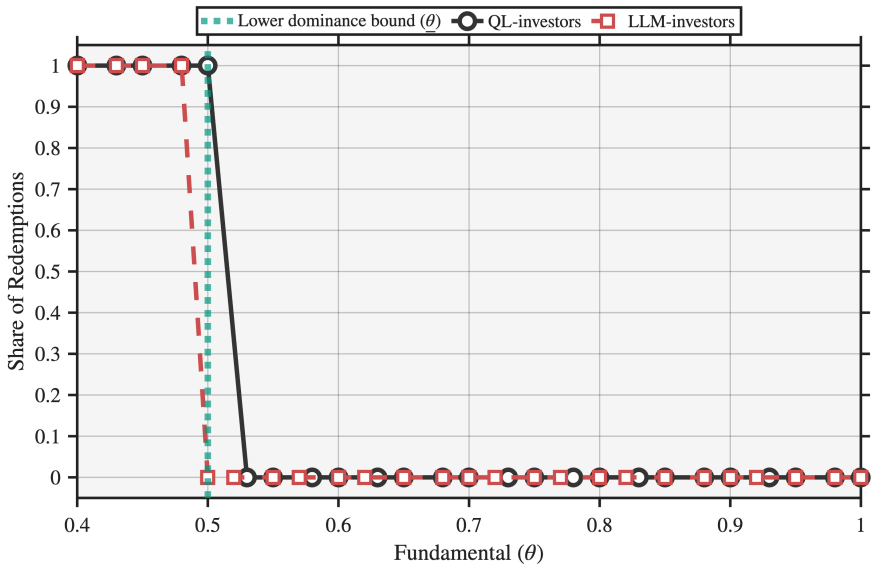


Figure 4: Decision of a single investor as a function of the fundamental, without fundamental or payoff uncertainty.

fundamental θ for $\lambda = 0.25$.[9] Both types of investors behave largely in accordance with theory: in the absence of strategic uncertainty, the first-best equilibrium is obtained, with the decision to switch from redeem to stay occurring around the first-best threshold, $\underline{\theta} = \frac{1}{R} = 0.5$. Changes in the illiquidity parameter play no role, since $\underline{\theta} = \frac{1}{R}$ is independent of λ.

Figure 5 shows the reasoning structure inferred from one LLM-investor's explanation for $\theta = 0.55$. The DAG begins with contextual inputs: the payoff from redeeming is fixed at 1, and the investor recognizes that it is the sole active participant, so no other investors redeem. From this, the reasoning proceeds to compute the payoff from staying, which equals 1.1. The model then compares the two options, noting that $1.1 > 1.0$, and concludes that staying yields the higher return. This inference flows into the decision sub-graph, where the investor chooses to stay.

[9] In all subsequent figures we set $\lambda = 0.25$ unless otherwise specified.

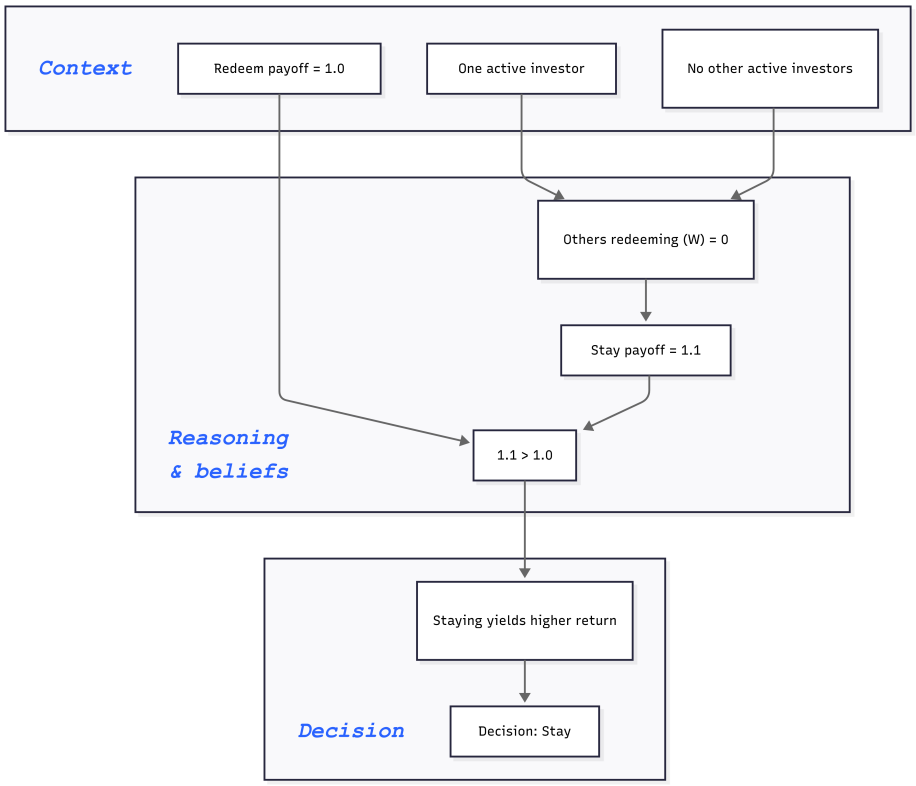


Figure 5: Reasoning structure for the LLM-investor. The DAG was constructed based on the explanation in the experiment with $\theta = 0.55$.

A. Coordination and Multiple Equilibria

Next, we turn to our analysis of experiments with multiple investors ($A = 30$) simultaneously choosing to redeem or stay in the absence of fundamental uncertainty. Figure 6 plots the share of investors who redeem as a function of the fundamental in the absence of payoff-uncertainty and for $\lambda = 0.25$. Within the dominance regions, investors perfectly coordinate on the Nash equilibria predicted by the theory. So, when the fundamental is weak, i.e., $\theta < \underline{\theta}$, they all redeem, while when fundamentals are strong, i.e., $\theta > \overline{\theta}$, they all stay. In the intermediate region, however, differences emerge. While QL-investors continue to coordinate on equilibrium outcomes, LLM-investors find it more difficult to coordinate leading to intermediate levels for redemptions.

The aggregate outcome obtained with QL-investors switches at some critical value of the fundamental, which we denote by $\tilde{\theta}$. So, for $\theta < \tilde{\theta}$, all agents redeem, while for $\theta \geq \tilde{\theta}$, they all stay. Across multiple runs of the QL algorithms, we obtain the same outcome. The implications of this are two-fold. First, QL-investors converge to equilibrium outcomes. And second, they consistently converge to the same equilibrium outcomes, implying that multiple equilibrium outcomes, for the same value of θ, do not materialize.

Rationalizing the behavior of QL-investors. The behavior of QL-investors and the cut-off, $\tilde{\theta}$ can be rationalized using K-level reasoning, which is a model of bounded rationality. In this framework, agents reason in layers: a level-0 agent chooses randomly between redeeming and staying. A level-1 agent assumes others are level-0 and best-responds to that assumption. A level-2 agent assumes others are level-1, and so on.

The aggregate behavior observed among QL investors aligns with that of level-1 reasoners. That is, they behave as if they expect others to act randomly and adjust their strategy accordingly. As a result, the critical threshold $\tilde{\theta}$ corresponds to the fundamental value at which an agent is indifferent between redeeming and staying, given the belief that others are equally likely to choose either action, i.e.,

$$\tilde{\theta} = \frac{R_1}{\left(\frac{1}{2}\right)^{A-1} \sum_{W=0}^{A-1} \binom{A-1}{W} \left(\frac{N-W(1+\lambda)}{N-W}\right) R_1 R}. \tag{6}$$

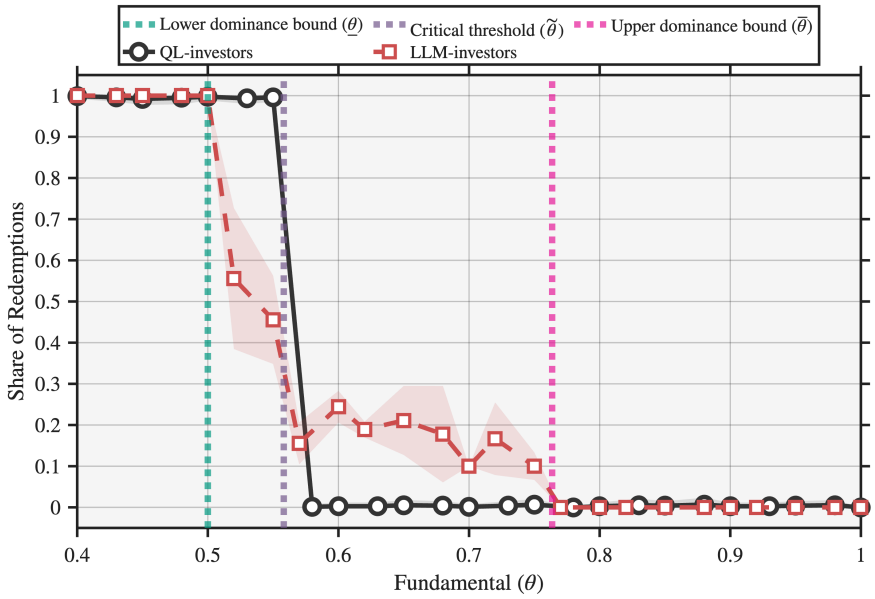


Figure 6: Share of redemptions as a function of the fundamental θ under full information (no fundamental or payoff uncertainty).

The threshold strategy, redeem whenever $\theta < \tilde{\theta}$, delivers the risk-dominant equilibrium. Christianos et al. (2023) and Albrecht et al. (2024) highlight that independent Q-learning algorithms converge to the risk-dominant equilibrium when they are uncertain about the actions of others.[10] Thus, QL-investors operate as if they all use the same threshold strategy that delivers the unique risk-dominant equilibrium.

Rationalizing the behavior of LLM-investors. In contrast, LLM-investors struggle to coordinate on equilibrium outcomes in the intermediate region. There are two key steps. First, investors recognize that the return from staying depends on the number of other investors who also stay. And for staying to be optimal, sufficiently many other investors must also stay. Having determined this critical number of other investors who must choose not to redeem for staying to be optimal, the next step involves reasoning over the actions of the other investors.

To understand the reason behind this finding, Figure 7 plots the DAG for the representative LLM-investor with $\theta = 0.55$ and $\lambda = 0.25$. Since the prompts to LLM-

[10]This result had also previously been established in the context of a two-player stag-hunt game (Bearden, 2001).

investors do not specify how they should form beliefs over the behavior of others, we find that different investors use different models to form their beliefs. In one model, the investor has an optimistic belief that all investors will coordinate on the Pareto-optimal action, which is to stay. However, in the second model, the investor holds a pessimistic belief that all others will choose to redeem. The other belief model identified in the reasoning also leads to the conclusion that redeeming is optimal. Our analysis thus implies that the failure to coordinate actions stems from LLM-investors using different approaches to form their beliefs over the behavior of others. We summarize our findings below.

Findings 1: *In the absence of both fundamental uncertainty and payoff uncertainty, both QL and LLM-investors coordinate on the theoretically predicted equilibria in the dominance regions: all investors redeem for $\theta < \underline{\theta}$ and all agents stay for $\theta > \overline{\theta}$. In the intermediate region, QL-investors behave as if they are level-1 reasoners, and so they all redeem whenever $\theta < \tilde{\theta}$. Consequently, there are no multiple equilibria. The LLM-investors fail to coordinate their actions in the intermediate region since different investors use different models to form beliefs about the behavior of others.*

B. Irrelevance of Payoff Uncertainty

We test Prediction 2 that payoff uncertainty is irrelevant for the aggregate outcome. Figure 8 plots the results for the share of redemptions as a function of the fundamental for both specifications of $R_2(\theta)$ and for both QL and LLM-investors. We note two key results. First, the outcome with LLM-investors is only slightly influenced by the payoff uncertainty. In fact, the results with and without payoff uncertainty are, for the most part, statistically indistinguishable. Second, QL-investors experience a much stronger bias towards redeeming, even for values of θ where staying is predicted to be the strictly dominant action.

Rationalizing the behavior of LLM-investors. An explanation for this result can be found by looking at the mental model that LLM-investors use to handle payoff uncertainty. As Figure 9 illustrates, the investors use the concept of expected value to treat the uncertainty, which renders the problem identical to that without payoff uncertainty. Thus, LLM-investors' ability to reason and contextualize delivers an aggregate outcome in line with our prediction.

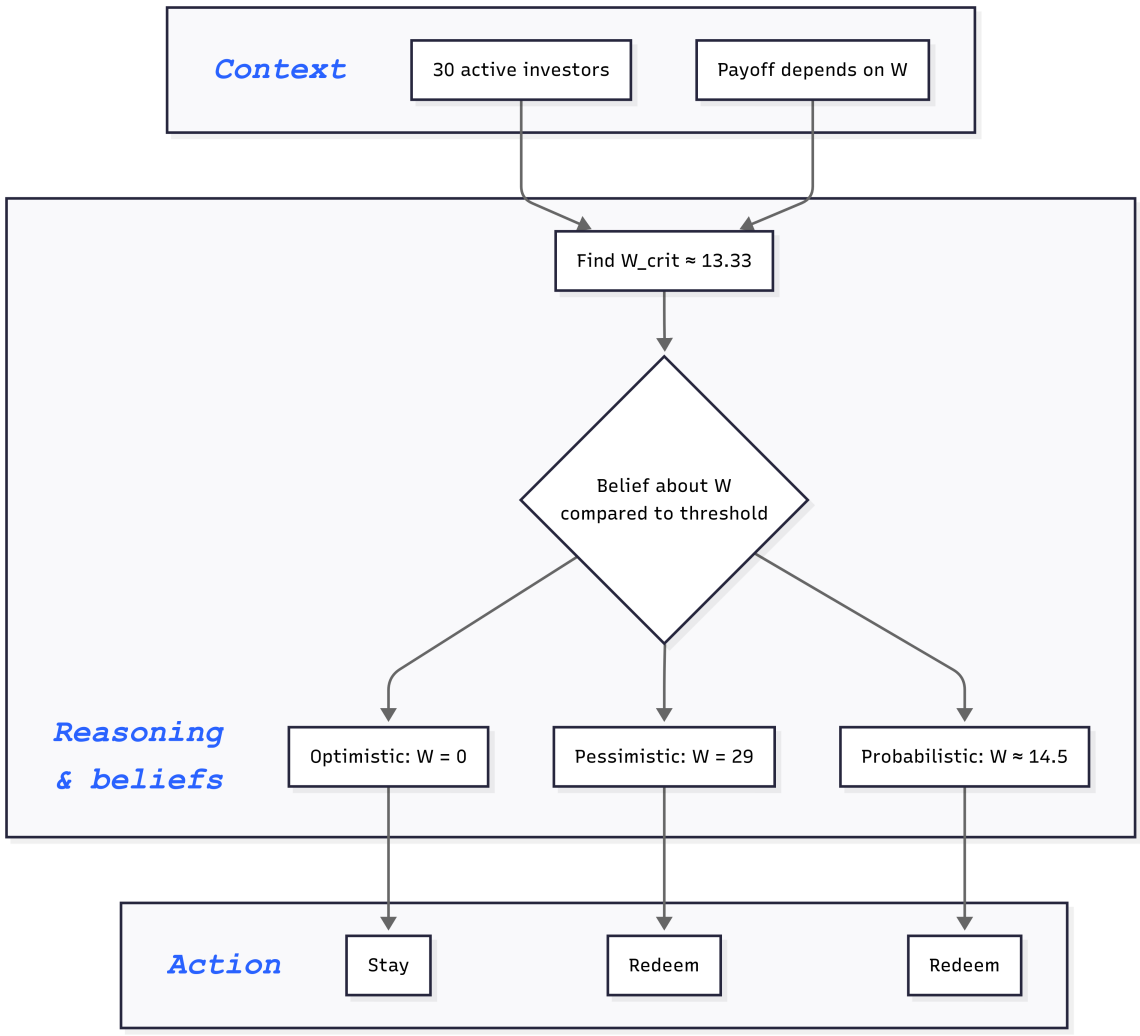


Figure 7: Reasoning structure for a representative LLM-investor. The DAG was derived from the explanations provided in the experiment with $\theta = 0.55$.

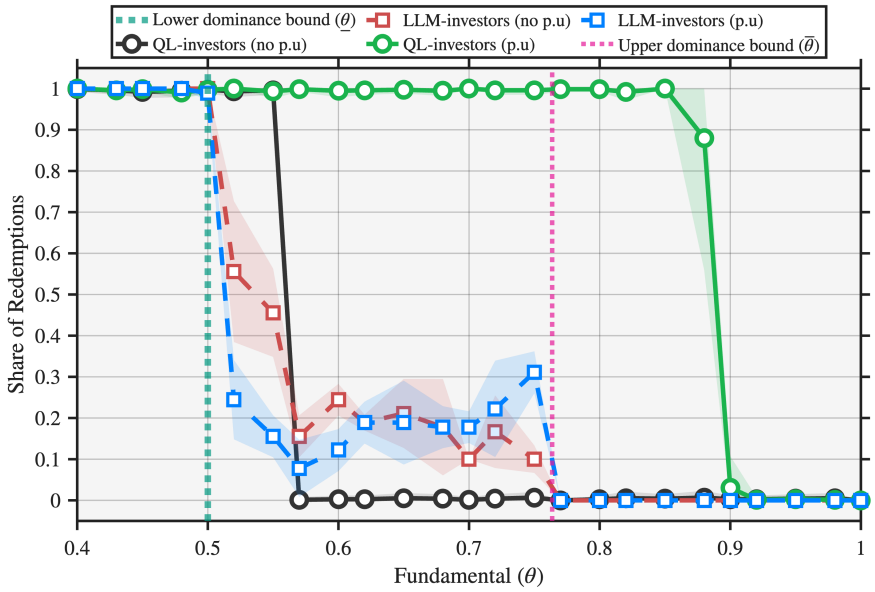


Figure 8: Payoff uncertainty and redemptions vs. θ (fundamentals known).

Rationalizing the behavior of QL-investors. An explanation for QL-investors' bias towards redeeming can be found by exploring the properties of reinforcement learning. Consider investor i in episode t, currently in state $\theta_{i,t}$, that chooses to stay (action A_S). With probability $1-\theta$, the fund's return is zero and the investor receives no reward, $\pi(\theta, a_S) = 0$. Since the Q-learning update rule is $Q_{i,t+1}(\theta_{i,t}, a) = (1-\alpha)Q_{i,t}(\theta_{i,t}, a) + \alpha \cdot \pi(\theta, a)$, a zero reward implies that

$$Q_{i,t+1}(\theta_{i,t}, A_S) = (1-\alpha)Q_{i,t}(\theta_{i,t}, a_S). \tag{7}$$

Thus, each zero-return episode revises the Q-value downward. An accumulation of such realizations, particularly during the exploration phase, causes the Q-values of staying to remain systematically below the true expected value $R\theta$, making QL-investors pessimistic about the benefits of staying.

In contrast, the payoff for redeeming is fixed at R_1 and not subject to uncertainty. Consequently, as the Q-values for redeeming rapidly converge to R_1. Once $Q_{i,t}(s, a_R) > Q_{i,t}(s, a_S)$, the investor increasingly chooses to redeem. This reduces opportunities to learn the true value of staying and further entrenches the bias towards redeeming.

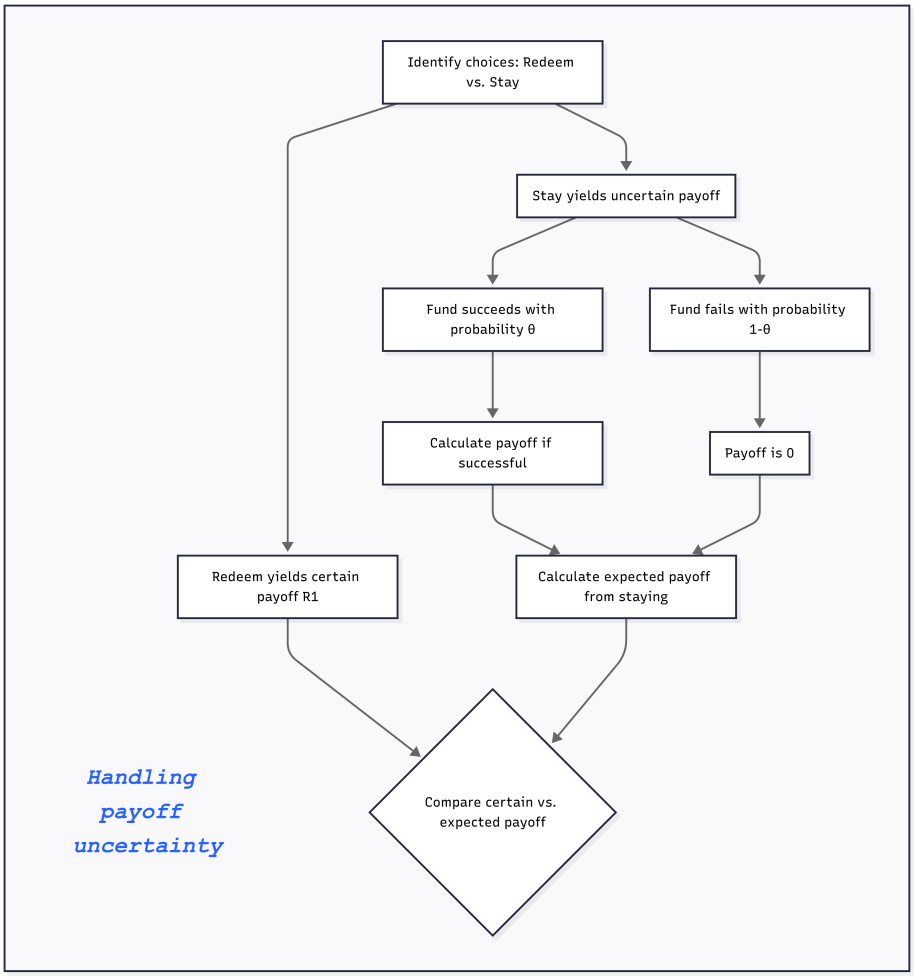


Figure 9: Inferred Mental model for how a representative LLM-investor handles payoff Uncertainty. The DAG was derived from the explanations provided in the experiment with $\theta = 0.55$.

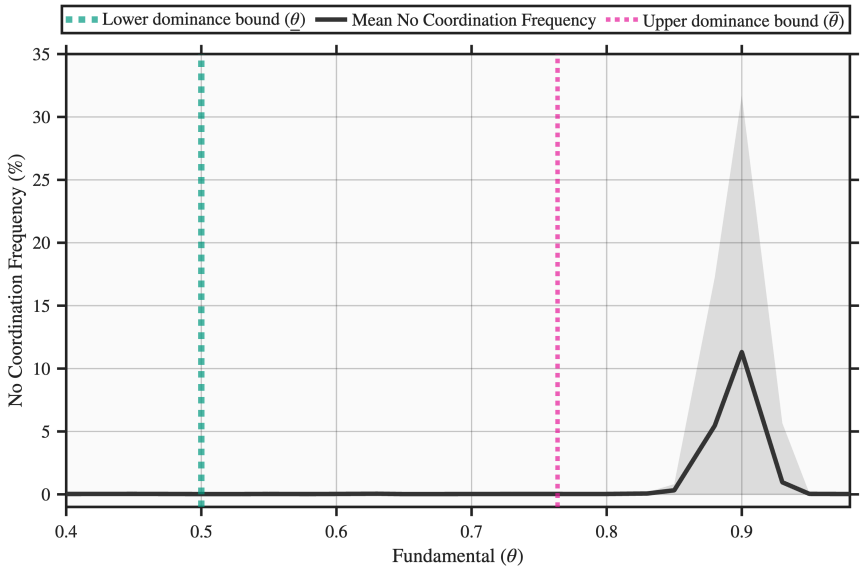


Figure 10: Frequency of coordination failures (neither "all redeem" nor "all stay") across fundamentals θ, with payoff but no fundamental uncertainty.

Payoff uncertainty can also hinder coordination among QL-investors. Figure 10 illustrates this by plotting the share of the last ten percent of episodes in which investors fail to coordinate on either "all redeem" or "all stay". For most values of θ, investors achieve nearly perfect coordination. However, for a range of high θ values, which correspond to the intermediate redemption shares in Figure 8, the degree of coordination falls sharply. This suggests that the intermediate levels of redemptions stem from coordination failure rather than equilibrium multiplicity. Coordination failure, in turn, arises because QL-investors fail to converge on a stable strategy profile.

Figure 11 examines this convergence failure more closely by plotting the rolling average of redemptions across episodes for several values of θ. The shaded regions indicate the standard deviation across independent training runs. For most θ, the share of redemptions converges to one and the variation between runs decreases, consistent with successful coordination on the "all redeem" equilibrium. But, for some intermediate θ values, the redemption shares exhibit a persistent upward drift and substantial cross-run volatility. These patterns suggest that the "all stay" outcome is unstable: repeated episodes in which agents receive zero payoff eventually push

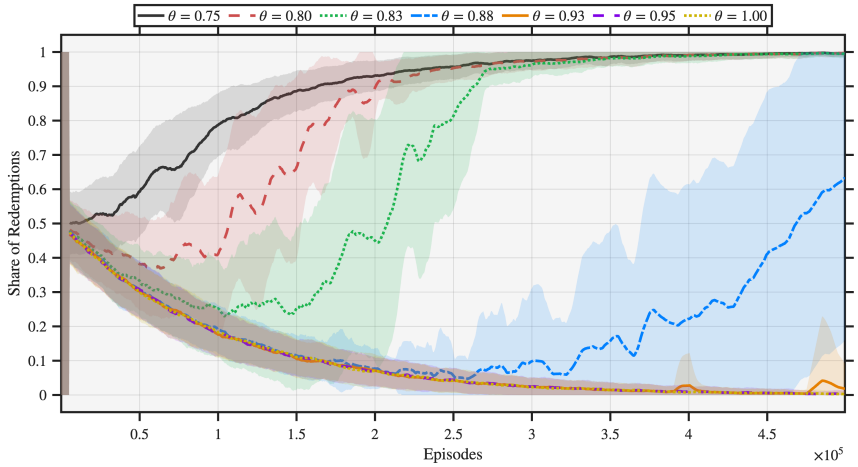


Figure 11: Evolution of redemptions across episodes with payoff uncertainty. The figure shows the rolling average share of redemptions over twenty-five training rounds, together with the associated standard deviations, for different values of θ.

them toward redeeming. In the limit, this dynamic implies that redemption becomes dominant even when fundamentals are strong. Only at the extreme case of $\theta = 1$, where staying yields a strictly positive payoff with certainty, does "all stay" emerge as a stable outcome.

Findings 2: *Introducing payoff uncertainty, while maintaining that there is no fundamental uncertainty, has no material impact on the aggregate behavior of LLM-investors. They use expected value as a solution concept to handle payoff uncertainty, thus rendering the results indistinguishable from those without payoff uncertainty. In contrast, payoff uncertainty introduces a strong bias towards redeeming for QL-investors. This bias persists well beyond the upper dominance bound, $\bar{\theta}$. Moreover, for very high values of θ, we find intermediate values for the share of redemptions, mainly driven by a very slow convergence to the "all redeem" outcome.*

C. *Global Games Equilibrium with Fundamental Uncertainty*

With fundamental uncertainty, equilibrium behavior is characterized by the panic threshold θ^* and the average redemption rate $w^*(\theta^*, \theta)$. Figure 12 compares the simulated outcomes of QL and LLM-investors to the theoretical benchmark across different levels of signal precision.

Panel (a) shows the case of highly precise signals ($\eta = 0.01$). Here, both QL-investors and LLM-investors closely track the theoretical prediction: redemption behavior switches sharply around the threshold θ^*, consistent with the global games equilibrium.

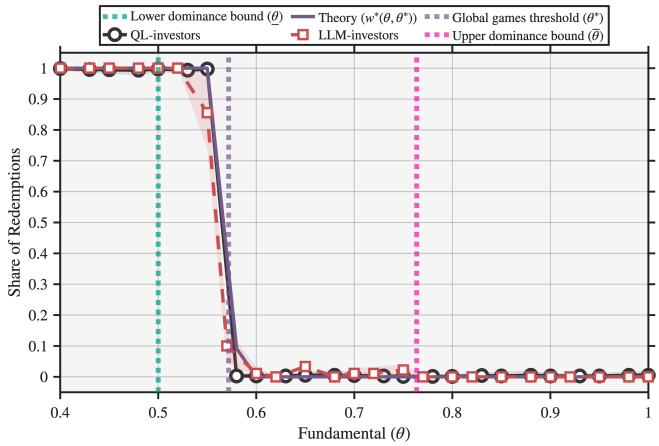
Panel (b) considers noisier signals ($\eta = 0.05$). In this case, a divergence emerges. LLM-investors continue to align with the theoretical prediction, coordinating their redemption decisions around θ^*. By contrast, QL-investors display a systematic bias toward redeeming, leading to higher average redemption rates than theory would predict.

Rationalizing the behavior of LLM-investors. Figure 13 illustrates the mental model employed by LLM-investors under fundamental uncertainty. These investors recognize that the environment is structurally equivalent to a canonical global games setup and proceed to solve it accordingly. By reframing the problem in this way, and (correctly) assuming that all other investors adopt the same signal threshold, they eliminate the ambiguity of belief selection. The decision problem then reduces to computing the threshold signal that makes the marginal investor indifferent between redeeming and staying.

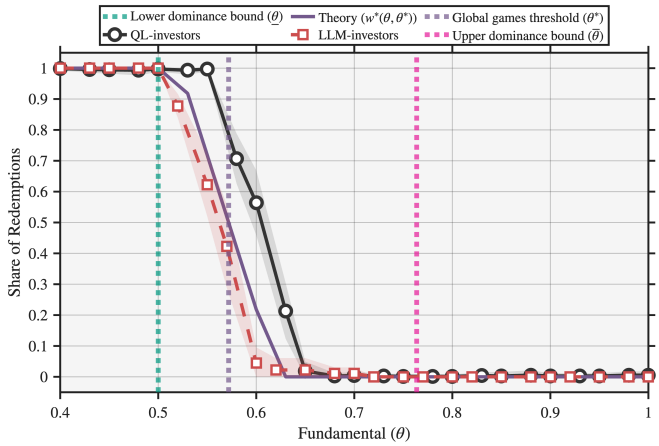
Rationalizing the behavior of QL-investors. Fundamental uncertainty induces a bias toward redeeming among QL-investors, though less pronounced than under payoff uncertainty. The mechanism is as follows. Because signals differ across investors, they may hold conflicting beliefs about the state of the world. This disagreement—especially when signals are imprecise—translates into heterogeneous redemption decisions. Relative to the case without fundamental uncertainty, for any realization of θ, the share of investors redeeming is higher, thereby reducing the payoff from staying.

This dynamic is particularly evident when $\theta \simeq \theta^*$ and signal precision is low. Some investors receive favorable signals suggesting that the fundamental is strong, while others observe weaker signals pointing below the threshold. Those who redeem early directly reduce the payoff to those who stay, which in turn depresses the Q-values associated with staying. In this sense, fundamental uncertainty operates much like payoff uncertainty in increasing the incentive to redeem early.

There is, however, a key distinction. Unlike under payoff uncertainty, learning in



(a) High precision signals $\eta = 0.01$



(b) Low precision signals, $\eta = 0.05$

Figure 12: Share of redemptions as a function of the fundamental, with fundamental uncertainty and no payoff uncertainty.

the presence of fundamental uncertainty does not converge to universal redemption. Instead, consistent with the theoretical prediction $w^*(\theta, \theta^*) \in (0, 1)$, the outcome involves persistent partial redemptions: some investors redeem while others stay. As shown in Figure 14, the rolling average of redemptions stabilizes at intermediate levels, with the shaded confidence bands flattening over time. This suggests that, in contrast to payoff uncertainty, the failure of QL-investors to coordinate fully is itself a stable equilibrium outcome under fundamental uncertainty.

Findings 3: *With fundamental uncertainty, LLM-investors use the global games so-*

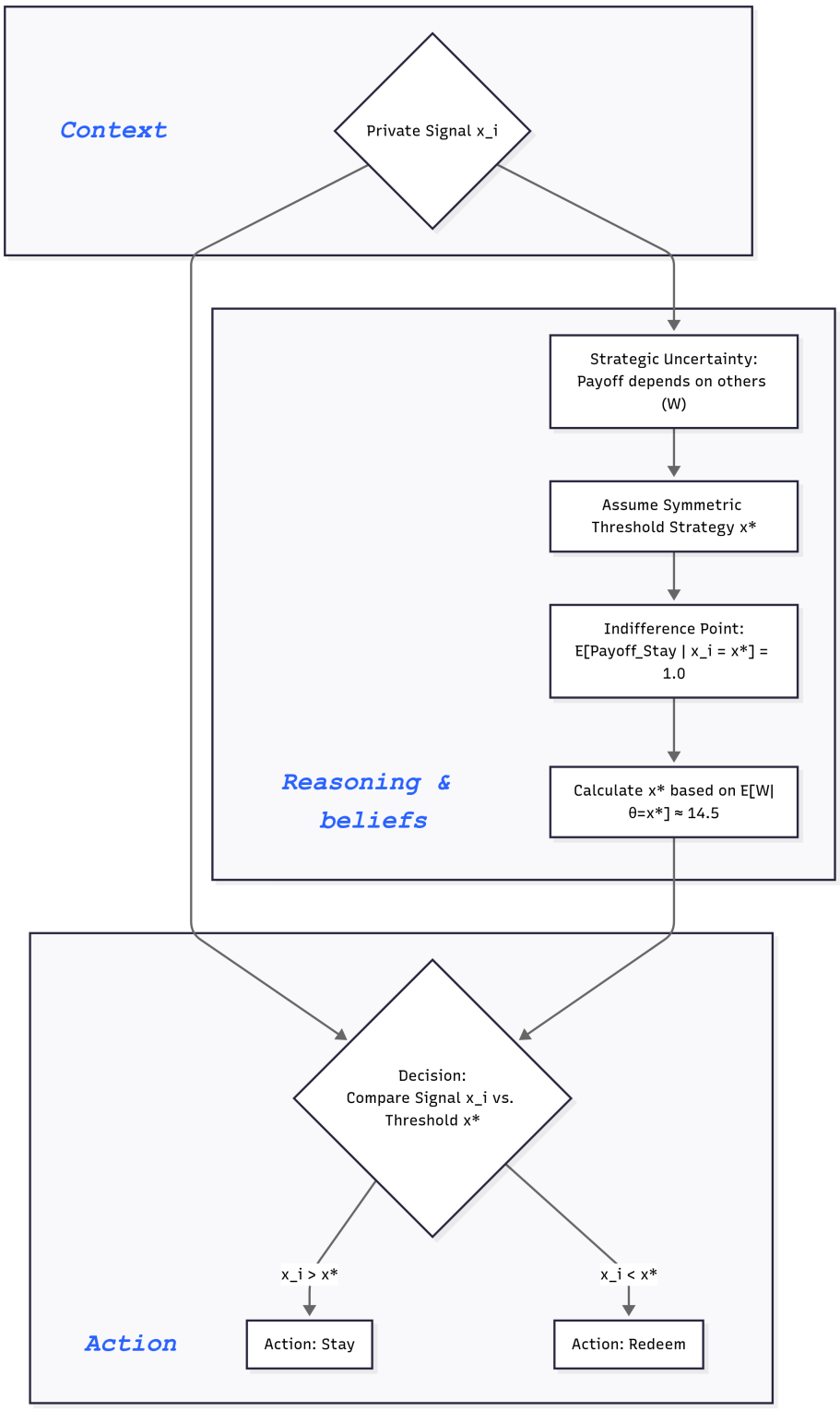


Figure 13: Inferred mental model for a representative LLM-investor with fundamental uncertainty. This DAG was derived from the explanations provided in the experiment with $A = 30$, $\theta = 0.55$ and $\eta = 0.05$.

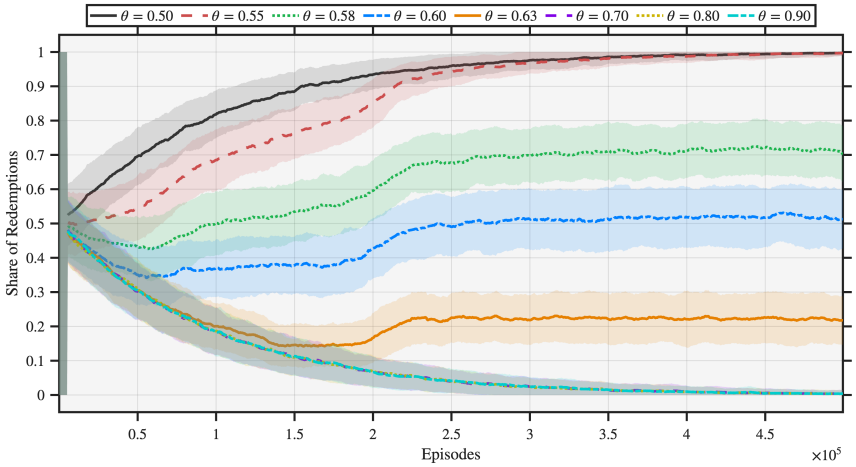


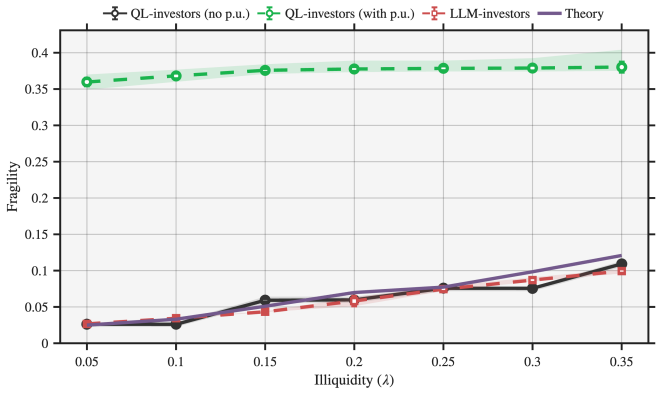
Figure 14: Evolution of redemptions across episodes with fundamental uncertainty ($\eta = 0.05$). The figure shows the rolling average share of redemptions over twenty-five training rounds, together with the associated standard deviations, for different values of θ.

lution concept and switch around the critical threshold, θ^*, irrespective of the level of signal precision. QL-investors, in contrast, are more sensitive to the level of signal precision, which induces a bias towards redeeming. Moreover, instances of partial redemptions for intermediate values of the fundamental are not driven by slow learning dynamics and emerge as an equilibrium outcome.

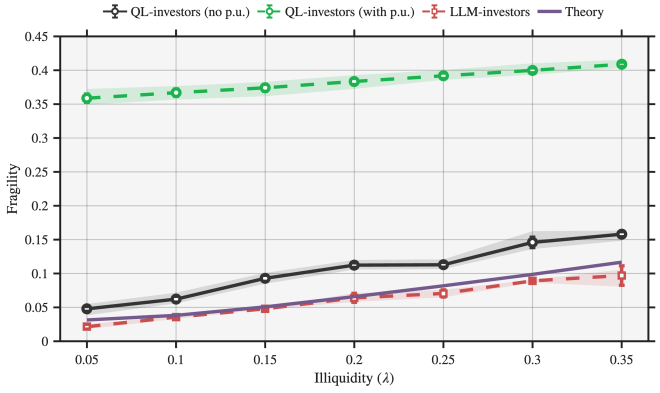
D. Relationship between Fragility and Illiquidity

We conclude by examining how fragility depends on asset illiquidity, captured by the parameter λ. Focusing on the case with fundamental uncertainty, our theoretical benchmark defines fragility as the difference between the redemption share $w^*(\theta^*, \theta)$ and the first-best allocation, where all investors redeem if and only if $\theta < \underline{\theta}$. Intuitively, this gap measures the excess redemptions arising from coordination failures. In our simulations with QL and LLM-investors, we define fragility analogously: as the deviation of the simulated redemption profile from the first-best allocation.

Figure 15 plots the simulation results for fragility as a function of λ. The two panels consider high-precision signals ($\eta = 0.01$) and low-precision signals ($\eta = 0.05$). Since payoff uncertainty has been shown to be irrelevant for LLM-investors, we restrict



(a) High precision signals $\eta = 0.01$



(b) Low precision signals, $\eta = 0.05$

Figure 15: Relationship between fragility and asset illiquidity.

attention to the case without payoff uncertainty. For QL-investors, by contrast, we report results both with and without payoff uncertainty.

The results reveal a clear difference in behavior across the two types of investors. The behavior of LLM-investors closely tracks the theoretical benchmark across both levels of signal precision, both in the overall level of fragility and in the upward-sloping relationship between fragility and illiquidity. This confirms the theoretical prediction that greater illiquidity systematically increases fragility.

For QL-investors, the outcomes are more nuanced. In the absence of payoff uncertainty, when signals are highly precise, the results are largely consistent with the

theoretical benchmark, in line with the evidence reported in Figure 12a. With noisier signals, however, QL-investors exhibit a stronger bias toward redemption, which raises the level of fragility relative to the benchmark.

The introduction of payoff uncertainty amplifies this bias further, but the relationship between fragility and illiquidity now depends on signal precision. With high precision, disagreement across investors is limited, and QL-investors tend to coordinate on redeeming for almost all values of θ, leaving fragility largely insensitive to λ apart from the narrow range of fundamentals where convergence problems arise (Figure 10). When signals are less precise, by contrast, partial redemptions persist as a stable feature of the learning dynamics (Figure 14), which in turn produces a stronger positive relationship between fragility and illiquidity.

Findings 4: *The relationship between fragility and illiquidity for LLM-investors is well approximated by the theoretical benchmark, irrespective of signal precision. In contrast, for QL-investors the relationship depends on both payoff uncertainty and signal precision: payoff uncertainty amplifies their bias toward redemption, while low signal precision strengthens the positive slope of fragility with respect to illiquidity.*

V. Conclusion

We study how AI agents behave in canonical coordination problems (e.g., bank runs) and, thus, their implications for financial stability. We find that equilibrium outcomes are highly sensitive to the design of agents' architectures. We show that Q-learning investors systematically over-redeem relative to the theoretical cutoff, whereas LLM-investors adhere more closely to the benchmark but coordinate less. Even when individual agents pursue their objectives effectively, collective dynamics can still produce uniform and inefficient behavior, transforming small shocks into system-wide runs and cascades.

Our contribution to the emerging literature on AI in finance is to show that the design of AI systems matters. It is not merely the presence of AI in financial decision-making, but how these agents are architected and how they interact with one another that shapes outcomes. As AI use becomes increasingly prevalent in financial domains, from trading algorithms to robo-advisors, it is essential to understand how these agents behave both individually and in aggregate. The risks of AI-induced coordination failures, such as bank runs and systemic crashes, are real, and may

surpass human-driven risks because of AI's speed, scale, and synchrony.

Our findings also raise broader questions about AI alignment in multi-agent context. While much of the AI alignment literature focuses on aligning a single system with human values, our results suggest that multi-agent alignment, or ensuring that interactions among many AI agent cohorts lead to socially beneficial outcomes, is equally important. In our model, each agent is individually aligned with its objective (e.g., maximizing rewards), yet the group sometimes converges on globally inefficient outcomes, much as humans do (Lorè and Heydari, 2024). With AI, however, such dynamics may emerge faster and more uniformly.

REFERENCES

Akata, Elif, Lion Schulz, Julian Coda-Forno, Seong Joon Oh, Matthias Bethge, and Eric Schulz, 2023, Playing Repeated Games with Large Language Models, arXiv:2305.16867.

Albrecht, Stefano V., Filippos Christianos, and Lukas Schäfer, 2024, *Multi-Agent Reinforcement Learning: Foundations and Modern Approaches* (MIT Press, Cambridge, MA).

Aldasoro, Iñaki, Leonardo Gambacorta, Anton Korinek, Vatsala Shreeti, and Merlin Stein, 2024, Intelligent Financial System: How AI is Transforming Finance, Technical report, BIS Working Paper 1194.

Anthis, Jacy Reese, Ryan Liu, Sean M. Richardson, Austin C Kozlowski, Bernard Koch, James Evans, Erik Brynjolfsson, and Michael Bernstein, 2025, LLM Social Simulations are a Promising Research Method, arXiv preprint arXiv:2504.02234.

Banchio, Martino, and Giacomo Mantegazza, 2022, Artificial Intelligence and Spontaneous Collusion, in *Proceedings of the 2022 ACM Conference on Economics and Computation*, 1–2.

Banchio, Martino, and Andrzej Skrzypacz, 2022, Artificial Intelligence and Auction Design.

Bank of England, 2025, Financial Stability in Focus: Artificial Intelligence in the Financial System, Technical report, Bank of England.

Bearden, J. Neil, 2001, The Evolution of Inefficiency in a Simulated Stag Hunt, *Behavior Research Methods, Instruments, & Computers* 33, 124–129.

Bebchuk, Lucian, and Itay Goldstein, 2011, Self-Fulfilling Credit Market Freezes, *Review of Financial Studies* 24, 3519–3555.

Bhagwat, Vineet, J Anthony Cookson, Chukwuma Dim, and Marina Niessner, 2025, The Market's Mirror: Revealing Investor Disagreement with LLMs, *FEB-RN Research Paper* .

Bybee, J. Leland, 2023, The Ghost in the Machine: Generating Beliefs with Large Language Models, arXiv preprint arXiv:2305.02823.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolo, and Sergio Pastorello, 2020, Artificial Intelligence, Algorithmic Pricing, and Collusion, *American Economic Review* 110, 3267–3297.

Carlsson, Hans, and Eric Van Damme, 1993, Global Games and Equilibrium Selection, *Econometrica* 61, 989–1018.

Chen, Qi, Itay Goldstein, Zeqiong Huang, and Rahul Vashishtha, 2024, Liquidity Transformation and Fragility in the US Banking Sector, *Journal of Finance* 79, 3985–4036.

Chen, Qi, Itay Goldstein, and Wei Jiang, 2010, Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Outflows, *Journal of Financial Economics* 97, 239–62.

Christianos, Filippos, Georgios Papoudakis, and Stefano V. Albrecht, 2023, Pareto Actor-Critic for Equilibrium Selection in Multi-Agent Reinforcement Learning, arXiv:2209.14344.

Colliard, Jean-Edouard, Thierry Foucault, and Stefano Lovo, 2025, Algorithmic Pricing and Liquidity in Securities Markets, CEPR Discussion Paper No. 17606.

Cong, Lin William, Ke Tang, Jingyuan Wang, and Yang Zhang, 2023, AlphaPortfolio: Direct Construction through Reinforcement Learning and Interpretable AI, SSRN 4124085.

Cont, Rama, and Wei Xiong, 2024, Dynamics of Market Making Algorithms in Dealer Markets: Learning and Tacit Collusion, *Mathematical Finance* 34, 467–521.

Cook, Thomas R, and Sophia Kazinnik, 2025, Social Group Bias in AI Finance, arXiv preprint arXiv:2506.17490.

Cook, Thomas R., Sophia Kazinnik, Anne Lundgaard Hansen, and Peter McAdam, 2023, Evaluating Local Language Models: An Application to Financial Earnings Calls, *Federal Reserve Bank of Kansas City Research Working Paper* no. 23-12.

Danielsson, Jon, Robert Macrae, and Andreas Uthemann, 2022, Artificial Intelligence and Systemic Risk, *Journal of Banking & Finance* 140, 106290.

Danielsson, Jon, and Andreas Uthemann, 2025, Artificial Intelligence and Financial Crises, Journal of Financial Stability, forthcoming.

Deloitte, 2024, Retail investors may soon rely on generative AI tools for financial investment advice.

Diamond, D., and P. Dybvig, 1983, Bank Runs, Deposit Insurance and Liquidity, *Journal of Political Economy* 91, 401–19.

Dou, Winston Wei, Itay Goldstein, and Yan Ji, 2025, AI-powered Trading, ALgorithmic Collusion, and Price Efficiency, NBER Working Paper No. 34054.

Fedyk, Anastassia, Ali Kakhbod, Peiyao Li, and Ulrike Malmendier, 2024, AI and Perception Biases in Investments: An Experimental Study, Available at SSRN 4787249.

Financial Stability Board, 2024, The Financial Stability Implications of Artificial Intelligence, Technical report, Financial Stability Board, Report to the G20.

Financial Times, 2025, The fund manager of the future might just be a machine, https://www.ft.com/content/ad12a0ec-9d6d-4ee7-83b7-643415f1a373.

Foley-Fisher, Nathan C., Borghan Narajabad, and Stephane H. Verani, 2020, Self-fulfilling Runs: Evidence from the U.S. Life Insurance Industry, *Journal of Political Economy* 128, 3520–3678.

Gao, Shen, Yuntao Wen, Minghang Zhu, Jianing Wei, Yuhan Cheng, Qunzi Zhang, and Shuo Shang, 2024, Simulating Financial Market via Large Language Model Based Agents, arXiv preprint arXiv:2406.19966.

Goldstein, Itay, Hao Jiang, and David T Ng, 2017, Investor Flows and Fragility in Corporate Bond Funds, *Journal of Financial Economics* 126, 592–613.

Goldstein, Itay, and Ady Pauzner, 2005, Demand Deposit Contracts and the Probability of Bank Runs, *Journal of Finance* 60, 1293–1327.

Gorton, Gary B., Elizabeth C. Klee, Chase P. Ross, Sharon Y. Ross, and Alexandros P. Vardoulakis, 2025, Leverage and Stablecoin Pegs, *Journal of Financial and Quantitative Analysis* .

Hansen, Anne Lundgaard, John J Horton, Sophia Kazinnik, Daniela Puzzello, and Ali Zarifhonarvar, 2024, Simulating the Survey of Professional Forecasters, SSRN Working Paper.

Horton, John J., 2023, Large Language Models as Simulated Economic Agents: What Can We Learn from Homo Silicus?, NBER Working Paper 31282.

Kazinnik, Sophia, 2023, Bank Run, Interrupted: Modeling Deposit Withdrawals with Generative AI, SSRN Working Paper.

Leitner, Georg, Jaspal Singh, Anton van der Kraaij, and Balazs Zsamboki, 2024, The Rise of Artificial Intelligence: Benefits and Risks for Financial Stability, Technical report, European Central Bank, Financial Stability Report.

Liu, Xuewen, and Antonio S. Mello, 2011, The Fragile Capital Structure of Hedge Funds and the Limits to Arbitrage, *Journal of Financial Economics* 102, 491–506.

Lopez-Lira, Alejandro, 2025, Can Large Language Models Trade? Testing Financial Theories with LLM Agents in Market Simulations, SSRN Working Paper.

Lorè, Nunzio, and Babak Heydari, 2024, Strategic Behavior of Large Language Models and the Role of Game Structure versus Contextual Framing, *Scientific Reports* 14, 18490.

MIT Sloan, 2024, Can generative ai provide trusted financial advice?, https://mitsloan.mit.edu/ideas-made-to-matter/can-generative-ai-provide-trusted-financial-advice.

Morris, Stephen, and Hyun Song Shin, 1998, Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks, *American Economic Review* 88, 587–597.

Morris, Stephen, and Hyun Song Shin, 2002, Measuring Strategic Uncertainty, Manuscript.

Nvidia, 2025, AI On: How Financial Services Companies Use Agentic AI to Enhance Productivity, Efficiency and Security.

Reinhart, Carmen M., and Kenneth S. Rogoff, 2009, *This Time Is Different: Eight Centuries of Financial Folly* (Princeton University Press, Princeton, NJ).

Reuters, 2025, After DeepSeek, Chinese fund managers beat High-Flyer's path to AI, https://www.reuters.com/technology/artificial-intelligence/after-deepseek-chinese-fund-managers-beat-high-flyers-path-ai-2025-03-14.

Shabsigh, Ghiath, and El Bachir Boukherouaa, 2023, Generative Artificial Intelligence in Finance: Risk Considerations, Fintech Note 2023/006, International Monetary Fund.

Yang, Hao, 2024, AI Coordination and Self-fulfilling Financial Crises, Working Paper.

Zarifhonarvar, Ali, 2024, Experimental Evidence on Large Language Models, SSRN Working Paper.

Appendices

Appendix A. Proofs

Proof of Proposition 1. For extreme values of the fundamental θ, investors have a dominant action. We start with the low values of θ. Irrespective of what other investors choose, redeeming at $t = 1$ is optimal for an investor when even the highest payoff they can accrue at $t = 2$, which corresponds to the payoff accrued if no other investors redeem, is smaller than R_1. Formally, this is the case when $R_1 R\theta < R_1$, that is, when $\theta < \underline{\theta} \equiv \frac{1}{R}$.

Symmetrically, staying until $t = 2$ is a dominant action when even the worst payoff that an investor expects to receive at the final date, which corresponds to the payoff accrued if all $A - 1$ investors redeem, is larger than R_1. Formally, this is the case when $R_1 R\theta \frac{N-(A-1)}{N-(A-1)(1+\lambda)} > R_1$ that is when $\theta > \overline{\theta} = \frac{1}{R} \frac{N-(A-1)}{N-(A-1)(1+\lambda}$.

When $\theta \in [\underline{\theta}, \overline{\theta}]$, both redeeming at $t = 1$ and not redeeming are equilibria. If an investor expects others to redeem, it is optimal for them to redeem as well, since staying until $t = 2$ would yield a lower payoff than R_1. Conversely, if no one else redeems at $t = 1$, it is optimal not to redeem either, as the payoff from staying, $R_1 R\theta$, exceeds R_1 for all $\theta > \underline{\theta}$. This completes the proof. □

Proof of Proposition 2. The proof adapts the standard approach in the global game literature (Goldstein and Pauzner, 2005; Chen et al., 2010) to the case with a discrete number of players. The arguments in their proof establish that there is a unique equilibrium in which investors redeem at $t = 1$ if and only if the signal they receive is below a common threshold θ^*, which is the signal at which an investor is indifferent between redeeming at $t = 1$ and $t = 2$ given what he or she believes about the signals received by other investors and, in turn, their behaviors.

We start by characterizing investors' decisions in two extreme ranges of fundamentals where investors have a dominant action. These two ranges correspond to the one characterized in Proposition 1. When $\theta < \underline{\theta}$ redeeming at $t = 1$ is a dominant action and we refer to this range as the lower dominance region. When $\theta > \overline{\theta}$, redeeming at $t = 2$ is the dominant action and we refer to this range as the upper dominance region.

Consider now the intermediate region where $\theta \in [\underline{\theta}, \overline{\theta}]$. Assume that investors behave according to a threshold strategy: they redeem their shares if they receive a signal below θ^* and stay until $t = 2$ otherwise.[11] Given that the signal is uniformly distributed over the interval $[-\eta, +\eta]$, the probability of receiving a signal below θ^* is $\frac{\theta^* - \theta + \eta}{2\eta}$. Building on this, we can compute the share of investors receiving the signal

[11] This comes at no loss of generality, as Goldstein and Pauzner (2005) show that in this game every equilibrium strategy is a threshold strategy.

below the cutoff, which is given by

$$w = \sum_{i=1}^{A} \{\theta_i < \theta^*\} \sim Binomial\left(A, \frac{\theta^* - \theta + \eta}{2\eta}\right). \tag{A1}$$

It follows that the probability that W out of A investors have received a signal below θ^* is given by:

$$f(W, A) = \binom{A}{W} \left(\frac{\theta^* - \theta + \eta}{2\eta}\right)^W \left(1 - \frac{\theta^* - \theta + \eta}{2\eta}\right)^{A-W}. \tag{A2}$$

Using the above, we can compute the probability that the investor receiving the signal θ^* assigns to n out of $A-1$ investors redeeming at $t=1$ as:

$$\frac{\binom{A-1}{W}}{2\eta} \int_{\theta^*-\eta}^{\theta^*+\eta} \frac{(\theta^* - \theta + \eta)^W (\theta - \theta^* - \eta)^{A-1-W}}{(2\eta)^{A-1}} d\theta \tag{A3}$$

As shown in Morris and Shin (2002), this probability is equal to $\frac{1}{1+A-1}$. Hence, the indifference condition that characterizes the run threshold θ^* reads:

$$\sum_{W=0}^{A-1} \frac{1}{A} \frac{N - W(1+\lambda)}{N - W} R_1 R\theta = R_1, \tag{A4}$$

which gives the expression (2) in the proposition. □

Proof of Proposition 3. The proof is straightforward and entails replicating the analysis of the previous two propositions using $R_2(\theta)$ as specified in (4). Since for any θ, the return is simply $R\theta$ and investors are risk neutral, all thresholds are exactly as in the previous two propositions. □

Appendix B. Prompts to generate DAGs

In this appendix, we provide the prompts used to generate the DAGs from the output json files. Prompt 4 provides the prompt used to produce the DAG with (i) a single investor (Figure 5), (ii) multiple investors and no payoff uncertainty (Figure 7), and (iii) multiple investors with fundamental uncertainty (Figure 13). To generate the DAG to depict the mental model used by LLMs to handle payoff uncertainty (Figure 9), we used Prompt 5.

```
You are given a JSON file from a simulation. The file contains: "actions": 0 = stay, 1 = redeem, "llm_responses" and "explanations": reasoning text, parameters such as θ, R, λ, N, number of active/passive investors. Your task is to produce a Mermaid DAG that cleanly summarizes the reasoning process for a typical investor.

Instructions:
Include three main sections only:
- Context: key parameters, information and setup

- Reasoning & belief: payoff comparison and threshold calculation

- Action: belief evaluation and aggregated final actions.

Show branching at the decision node only if needed: Optimistic (W < Wcrit), Pessimistic (W > Wcrit), Uncertain (W ∼ Wcrit).

Keep node labels short (≤ 10 words).
```

Prompt 4: Basic DAG generation prompt

```
You are given a JSON file from an experiment with payoff uncertainty. The file contains explanations for the decisions. Extract the steps for how investors handle the payoff uncertainty. Summarize those steps in a Mermaid DAG under a single section titled ‘‘Handling Payoff Uncertainty’’. Each node should include ≤10 words. Do not consider how strategic uncertainty is handled.
```

Prompt 5: Prompt to generate DAG depicting handling of payoff uncertainty

Appendix C. Fund illiquidity and payoff uncertainty at date 1

In this section, we relax the assumption concerning fund illiquidity and assume that $A > \frac{N}{1+\lambda}$. This implies that in the presence of a sufficiently large number of early redemptions, i.e., when $W \geq \overline{W} \equiv \frac{N}{1+\lambda}$, the fund liquidates its entire portfolio at date 1 and defaults. In this circumstance, we assume that early redeeming investors receive nothing due to the presence of full bankruptcy costs. This modification relative to the benchmark model introduces payoff uncertainty at date 1 and, thus, allows us to check whether our results concerning to the Q-algorithms' behavior in the presence of payoff uncertainty are robust.[12] As it will become useful for the characterization of the case with fundamental uncertainty, in line with the literature, e.g., Goldstein and Pauzner (2005), we assume that when $\theta = 1$, $\lambda = 0$ and the liquidation value of the fund's investment jumps to R. This implies that when $\theta = 1$, it is a dominant action for the investors not to redeem at date 1.

To isolate the role of payoff uncertainty at date 1, we characterize the equilibrium assuming that the bank's investment returns $R\theta$ at date 2, i..e, no payoff uncertainty at the final date. We proceed as in the main text, deriving first the equilibrium in the absence of fundamental uncertainty (i.e., θ is observable) and then assuming that investors receive an imperfect private signal on θ of the form $\theta_i = \theta + \epsilon_i$, with $\epsilon_i \sim [-\eta, +\eta]$. The following proposition characterizes the equilibria in the two instances.

Proposition 4: *When the fundamentals θ are observable, all investors redeem at $t = 1$ when $\theta \leq \underline{\theta} \equiv \frac{1}{R}$ and stay until $t = 2$ when $\theta = 1$. In the range $\theta \in (\underline{\theta}, 1)$, redeeming at date 1 and at date 2 are both equilibria.*

When the fundamentals θ are not observable, model has a unique symmetric Bayes-Nash equilibrium in threshold strategies that is characterized by a critical signal

$$\theta^* = \frac{\sum_{W=0}^{\overline{W}}}{\sum_{W=0}^{\overline{W}} \frac{N-W(1+\lambda)}{N-W} R}, \tag{C1}$$

such that an investor will redeem their share at $t = 1$ if and only if their signal is below θ^*.

Proof. The proof follows closely that of Proposition 1 and 2, with the only difference that, in the case of fundamental uncertainty, the indifference condition giving θ^* is

[12] Notice that the assumption of full bankruptcy costs give rise to a payoff structure that is akin to the case where early redeeming investors are repaid according to a sequential service schedule.

equal to:

$$\sum_{W=0}^{\overline{W}} \frac{1}{A} \frac{N - W(1+\lambda)}{N - W} R_1 R\theta = \sum_{W=0}^{\overline{W}} \frac{1}{A} R_1. \tag{C2}$$

The p-dominance threshold in the presence of payoff uncertainty at $t = 1$ θ_{RD} solves $u_1 = u_2$ evaluated at $p = \frac{1}{2}$, where

$$u_1 = R_1 \sum_{W=0}^{\overline{W}} \binom{A-1}{W} p^W (1-p)^{A-1-W},$$

and

$$u_2 = R_1 \sum_{W=0}^{\overline{W}} \binom{A-1}{w} p^W (1-p)^{A-1-W} R\theta \frac{N - W(1+\lambda)}{N - W}.$$

□