

EVALUATING LLM AGENTS AS HUMAN SIMULATORS IN CLIMATE SOCIAL DILEMMAS

Anonymous authors

Paper under double-blind review

ABSTRACT

Understanding how humans and institutions behave in climate-related social dilemmas is critical for designing effective climate policy, yet standard agent-based models often rely on simplified decision rules or fully rational agents and therefore struggle to capture bounded rationality, heterogeneity, and communication. We evaluate large language model (LLM) agents as behaviorally grounded simulators of companies and investors in a continuous-action climate-finance dilemma built on the InvestESG platform, and benchmark them against fully rational profit-driven reinforcement learning (RL) agents, a centralized social planner, and human participants in the same game. Our results show that LLM-based simulations can be powerful tools for analyzing social-good-oriented policymaking. LLM agents naturally reproduce human-like cooperative tendencies, flexibly support heterogeneous behavior, and exhibit emergent coordination and even collusion when communication is introduced, phenomena that are difficult to capture with conventional modeling approaches. At the same time, these simulations can be fragile: LLM behavior is sensitive to contextual framing and often requires explicit numerical scaffolding to ensure reliable reasoning.

1 INTRODUCTION

Climate change poses an immediate and substantial threat to human societies, demanding timely and effective policy responses. At its core, climate mitigation constitutes an intertemporal social dilemma (Leibo et al., 2017): actors face short-term economic costs in exchange for long-term collective benefits, while incentives to free-ride undermine coordinated action. These dynamics make policy design particularly challenging, which is compounded by the fact that large-scale policy experimentation is slow, costly, and often infeasible. As a result, agent-based modeling (ABM) emerged as an alternative pathway for studying the effects of policy interventions using computer simulations. (Parker et al., 2003; Axtell & Farmer, 2025; Duffy, 2006).

While MARL agents provide a *normative* benchmark grounded in optimal behavior, large language models (LLMs) (i) offer a *descriptive* alternative that may better reflect bounded rationality, social norms, and contextual reasoning observed in human decision-making, and (ii) enable naturalistic communication between agents. For example, behavioral economics has long shown that humans behave more cooperatively in social-dilemma settings (e.g., public-goods experiments) than classic game theory with purely self-interested agents would predict (Fehr & Gächter, 2000; Fischbacher et al., 2001).

Consistent with these human patterns, LLM agents placed in the same paradigms exhibit a similar “prosocial” initial tendency toward cooperation (Li & Shirado, 2025; Fontana et al., 2025). These properties make LLMs a promising tool for simulating complex real human and institutional behavior (Rahwan et al., 2019; Bourgin et al., 2019).

Motivated by these findings, we investigate whether LLMs can simulate human decision-makers in climate-action settings, which are more realistic than abstract social-dilemma games and require more sophisticated numerical reasoning. Specifically, we build on InvestESG (Hou et al., 2024), a MARL-based ABM platform that models companies’ investment in Environmental, Social, and Governance (ESG) efforts, specifically climate mitigation, and investor capital allocation over a long-horizon, multi-period environment. To study whether LLMs can serve as simulators of human

054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100
101
102
103
104
105
106
107

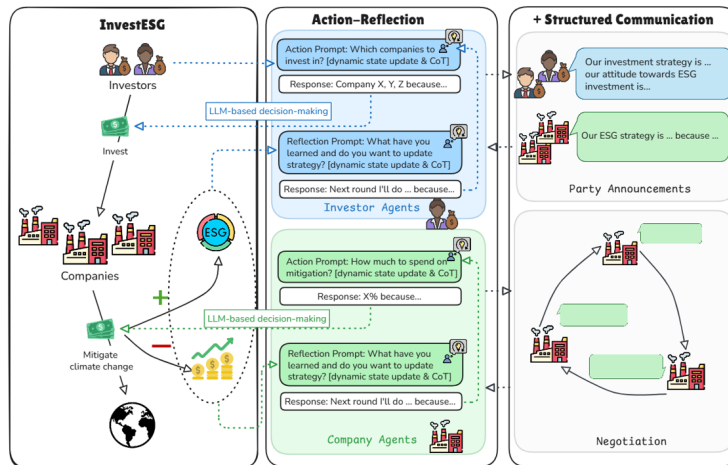


Figure 1: Overview of the role-play simulation process. Dashed arrows indicate the chronological flow of the simulation, while solid arrows represent the causal effects of decisions.

and institutional behavior in this setting, we replace InvestESG’s original MARL agents with LLM-driven companies and investors.

We find that LLM-based agents consistently allocate more resources to climate mitigation than profit-driven MARL agents, in some cases even exceeding the social planner’s optimum. This tendency toward over-cooperation appears to be partly driven by limitations in numerical reasoning and by value priors learned during training. Introducing a structured cost–benefit tool to support quantitative reasoning and reframing the task in neutral, non-climate language both reduce over-cooperation, though neither eliminates it entirely. While such behavior is economically suboptimal, it more closely resembles the decisions of human participants than those of MARL baselines.

At the same time, allowing communication among LLM-controlled firms induces coordinated behavior that sometimes resembles collusion, resulting in reduced climate mitigation efforts. Finally, when investor agents are conditioned on personas derived from real investor data, LLMs reproduce heterogeneous investment patterns consistent with those observed in real-world trading behavior.

Together, these results demonstrate that LLM agents can capture human-like cooperation and heterogeneity in complex social dilemmas, while remaining sensitive to cognitive, semantic, and social context.

We contribute to the literature in two ways. First, we provide, to our knowledge, the first study of LLMs for ABM in climate decision-making settings. Second, we propose a unified framework for comparing LLM-based agent simulations to rational MARL agents, a social planner, and human participants, which provides clear reference points for interpreting how contextual and architecture design steers LLM behavior, rather than leaving such interventions unanchored.

2 THE INVESTESG MARL ENVIRONMENT

The InvestESG environment (Hou et al., 2024) models a coupled economic–climate system with two types of decision-makers, namely company agents and investor agents, each implemented as RL agents trained via Proximal Policy Optimization (PPO) (Schulman et al., 2017). This design reflects the strengths of MARL for simulating complex interactions among multiple agents over long time horizons. The simulation spans 100 time steps, each representing one year. During each step, companies experience baseline economic growth that increases their capital levels. In parallel, climate dynamics unfold through stochastic climate events such as extreme heat or drought whose likelihood increases over time in accordance with the IPCC’s 4°C warming scenario. When a climate event occurs, all companies incur asset losses.

To counteract this risk, companies can allocate a portion of their capital to climate mitigation, which slows the progression of climate risk. These investments also raise their ESG scores, a standardized measure of corporate sustainability that is publicly shared with investors and other companies. At each step, agents observe relevant system states, including company capital levels, ESG scores from previous rounds, and historical climate events. Based on these observations, company agents decide what percentage of their capital to allocate to mitigation in the following step, while investor agents form portfolios by selecting companies based on their financial performance and ESG profiles. Agent performance is evaluated using two key metrics: climate risk, defined as the probability of a disaster-level climate event occurring in the final simulation year; and ESG investment, measured as the cumulative system-wide spending on climate mitigation over the course of the simulation. Profit-driven RL agents in this environment do little to mitigate climate change unless investors place substantial value on sustainability (Hou et al., 2024).

This is because these agents are self-interested, optimizing solely for their own cumulative reward over time without explicit concern for collective welfare. Although climate risk imposes costs on each firm individually, mitigation efforts generate shared benefits. Hence once one firm begins to invest in mitigation, other firms have an incentive to reduce their own spending and free-ride on the mitigation efforts of others instead. As a result, purely profit-motivated firms are unlikely to sustain high levels of mitigation, making it difficult for the system to reach outcomes that are collectively optimal. As such, the InvestESG environment serves as a testbed for evaluating whether LLM-based agents and humans exhibit more socially aligned or cooperative behaviors under the same dynamics.

3 METHOD

In this section, we describe the architecture of the LLM-based agents integrated into the InvestESG environment and outline the efforts made to the environment to enable meaningful comparisons between LLM- and RL-based approaches and human agent behaviors.

3.1 MARL-LLM ALIGNMENT

To enable a fair comparison between the MARL and LLM-based settings, we adapt the original 100-step InvestESG environment into a condensed 10-step version, with each step representing a decade. The LLM-based simulation runs over the same 10 steps and shares the same environment parameters and transition dynamics as the MARL version. To ensure comparability, the LLM agent receives the same information set available to MARL agents, as detailed in the following section.

3.2 LLM-BASED AGENT ARCHITECTURE

We implement both company and investor agents as externally orchestrated sequences of LLM API calls, using role-specific prompting templates for each agent type, as illustrated in Figure 1.

Action–Reflection Reasoning. Following the framework introduced by Piatti et al. (2024), LLM-based agents interact with the InvestESG environment through a repeated cycle of action and reflection implemented as sequential API calls. At each simulation step, the current environment state is embedded into a structured prompt. All company and investor agents are then queried in parallel: company agents decide what fraction of capital to allocate to climate mitigation, while investor agents select portfolios. We refer to these queries as *Action* prompts. The resulting model outputs are parsed into executable actions and applied to the environment. After the environment transitions and rewards are computed, each agent is queried again with an updated prompt that appends feedback from the previous step, including outcomes and a brief textual summary of prior reasoning. This *Reflection* prompt asks the model to assess the consequences of its last action and produce guidance for future decisions. Insights generated during reflection are then recorded and concatenated into subsequent Action prompts. Chain-of-thought prompting (Wei et al., 2022) is used in both phases to encourage step-by-step reasoning and improve traceability of decisions.

State-Aware Prompt. Each prompt begins with an overview of the environment, including the underlying dynamics, the number and types of agents in the system, and a description of the specific role of the receiving agent. Prompts reveal the same information as in Hou et al. (2024) to ensure consistency with the baseline environment. All agents receive information on capital levels of all



Figure 2: Action-Communication-Reflection Cycle

companies, their ESG scores from previous rounds, and investor allocation decisions. Company agents additionally receive private observations of their own prior-period profit margin, capital losses from recent climate events, and cumulative ESG mitigation spending. This information is compiled into a structured prompt referred to as the *Vanilla* prompt, allowing agents to make context-aware decisions.

Structured Communication. We also extend the simulation with structured inter-agent *Communication*. In this variant, the standard action–reflection loop is extended to an action–communication–reflection cycle. As shown in Figure 2, each communication round is governed by a fixed, moderator-led protocol. The moderator first summarizes relevant outcomes from the previous round and the current environment state. Investor agents then make brief announcements describing their ESG preferences, followed by company statements outlining intended strategies for the current round. Companies subsequently enter a structured turn-based negotiation phase in which they can propose, counter, or accept cooperative ESG initiatives.

We provide a detailed technical description of the structured communication implementation in Appendix B.3, along with examples of all prompt variants in Appendix C.

3.3 BENCHMARK BASELINES

We evaluate LLM-based agents against three complementary benchmarks that capture rational, optimal, and human decision-making under identical environmental dynamics.

Decentralized MARL Baseline. We construct a decentralized MARL baseline using the 10-step adaptation of InvestESG, replicating the setup of Hou et al. (2024). The environment includes three company agents and five investor agents, each modeled as an independent RL agent optimized with PPO policy (Schulman et al., 2017) optimizing its own long-term reward through interaction with the environment and other agents.

Centralized Social Planner Baseline. We further introduce a centralized social-planner RL baseline not explored in prior work. In this variant, multiple company agents are replaced by a single agent endowed with the total company capital, and investor agents are removed. The centralized agent optimizes system-wide profit rather than individual returns, providing an upper-bound benchmark for socially coordinated outcomes.

Human Participant Baseline. We compare LLM agent decisions against those of human participants acting as company agents. To ensure comparability between LLM-based and human-based firms, we fix investor behavior by reusing the investor actions from the completed all-LLM run. For the human experiment, we randomly sample a complete trajectory from this dataset. At each round of the sampled trajectory, we extract the LLM investor’s investment behavior together with the corresponding sequence of (ESG score, investment) pairs for companies in that round. We then sort the human-round companies by their current ESG scores and assign investments by position: the highest-ESG company in the human round receives the investment associated with the highest-ESG company in the sampled record, the second-highest receives the second-highest, and so forth. This procedure preserves the characteristic pattern of LLM investor behavior across ESG rankings while ensuring a consistent and reproducible investor response for human participants. A human moderator facilitates the simulation by presenting rules and environmental states in a human-readable format that conveys the same information as the LLM prompts. The moderator collects decisions

for each round, advances the simulation, and reports updated outcomes. Participants do not communicate and remain blind to the decisions of others during the decision-making phase.

The participant cohort consists of individuals from the United States and China, representing diverse racial and professional backgrounds. All participants possess a higher education background and have sufficient numerical skills to navigate the simulation. To ensure accurate understanding of the experimental setting, we administered simple comprehension checks both before and after the sessions. This study was approved by the relevant Institutional Review Board (IRB).

4 EXPERIMENTS AND RESULTS

This section presents our experiments and results. We begin by evaluating the default Action-Reflection LLM agent under the vanilla prompt without communication. We then vary company prompt designs to probe the drivers of firm behavior. Next, we introduce inter-agent communication and examine its impact on firm coordination. Finally, we study LLM investor agents and evaluate whether they are able to reproduce ESG investment patterns observed among human investors.

LLMs’ Tendency to Over-Cooperate. We first evaluate the performance of all-LLM companies and investors using vanilla prompting. Figure 6 presents the comparison between the RL-based social planner, decentralized MARL, and LLMs with vanilla prompt setting. For LLM-based experiments, we evaluate a non-reasoning model (GPT-4.1) and a reasoning-enabled model (GPT-o4-mini). Under the vanilla setting, both GPT-4.1 and GPT-o4-mini agents consistently invest far more in climate mitigation than either the MARL baseline or the social planner’s optimal policy. This elevated spending leads to substantially lower realized climate risk, placing LLM agents well above the purely profit-driven MARL agents and even beyond what is prescribed by a fully rational planner optimizing system-wide welfare.

Tool-Use for Quantitative Comprehension. Given the tendency of LLM agents to over-cooperate, a natural question is whether this behavior partly reflects limitations in numerical reasoning rather than genuine strategic intent. Prior work has shown that LLMs struggle with mathematical reasoning when computations are embedded in long, context-rich inputs, especially in settings that require tracking and manipulating multiple quantities over time (Zhao et al., 2024; Reddy et al., 2024). This concern is particularly relevant in our environment, where firm decisions lie in a continuous action space and involve intertemporal trade-offs and strategic interactions with other agents, substantially increasing computational complexity. Even for a single firm, deriving an optimal policy would require solving a challenging stochastic dynamic programming problem with no obvious closed-form solution, suggesting that exact optimization may be out of reach for LLM agents operating purely through in-context reasoning.

To support decision-making, we introduce a variant we refer to as *tool-use*. This setting preserves the baseline environment but augments each firm’s action prompt with a cost-benefit table generated via a calculator tool. The table lists a discrete set of mitigation spending options along with their near-term costs, projected reductions in capital losses, and estimated ESG scores. Rather than attempting full optimization, this table provides a simple, forward-looking snapshot of immediate trade-offs, which mirrors how real-world companies often rely on partial, myopic financial analyses to guide complex decisions. The goal of this tool is to reduce reliance on the LLM’s internal arithmetic in rich numerical contexts.

Under the tool-use setting, both GPT-4.1 and GPT-o4-mini reduce mitigation spending and increase final climate risk compared to the vanilla setup. However, LLM agents still behave more cooperatively than MARL agents, and continue to “over”-mitigate relative to the social planner baseline. By adding numerical scratch-pads and step-by-step cost-benefit scaffolds, we bring mitigation levels closer to the social planner’s optimal benchmark, indicating that LLMs’ “over-mitigation” may partially come from limitations in in-context numerical reasoning.

Context Sensitivity. While limitations in quantitative reasoning may partly explain LLMs’ tendency to over-cooperate, prior work points to a second mechanism involving sensitivity to contextual framing and inherited value priors. Backmann et al. (2025) show that LLM behavior is highly sensitive to prompt context and that pretrained value priors can leak into LLM decisions. Motivated by this evidence, we ask whether the cooperative behavior we observe reflects pro-climate values learned

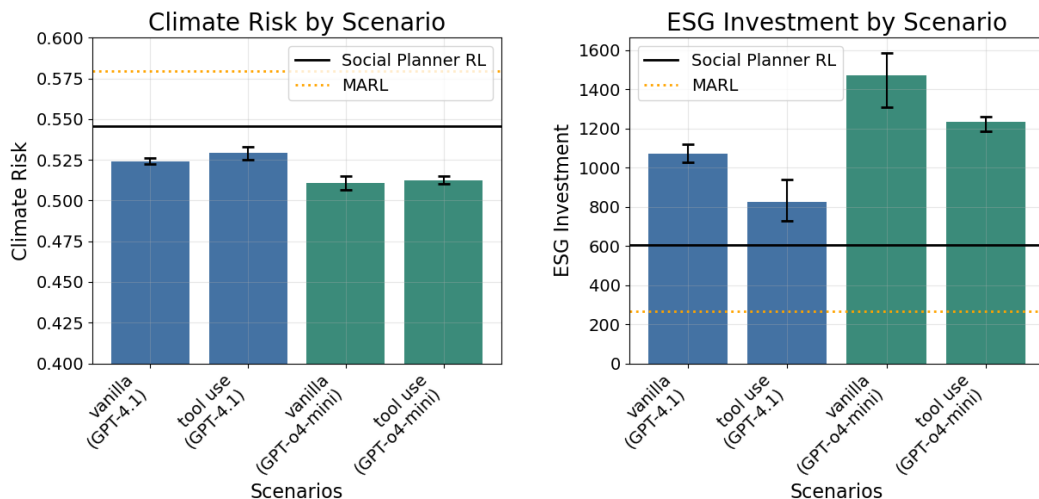


Figure 3: Introducing myopic cost-benefit information in the *tool use* setting reduces the over-cooperate tendency but does not reverse it.

during pretraining or post-training, rather than reasoning over the underlying incentives. To isolate the role of topical framing, we rerun the decision task using prompts that preserve the same financial payoff structure but remove all climate-related language. The prompts for this context-free version are provided in Appendix C.4.

The results shown in Figure 7 reveal that LLMs continued to favor ‘Option 1’ (the neutrally framed equivalent of ‘Mitigate’), suggesting they are capable of recognizing the underlying structure of a social dilemma even in the absence of explicit semantic cues. However, mitigation levels were still lower under the neutral framing compared to the original prompts containing climate language. This indicates that while LLMs exhibit a baseline cooperative bias, their decisions remain context-sensitive, consistent with findings by Backmann et al. (2025). Thus, part of the “over”-mitigation observed in climate-framed scenarios may stem from values or associations absorbed during pre-training or posttraining on climate-related content.

The Cooperative Human Baseline. Taken together, our results indicate that LLM agents exhibit a persistent tendency to over-cooperate, even after accounting for limitations in quantitative reasoning and sensitivity to contextual framing. LLM-based firms invest in mitigation efforts even when such actions are not justified by monetary incentives alone, in sharp contrast to the fully rational, self-interested behavior assumed in MARL baselines. This naturally raises a question: is such behavior in fact closer to that of human decision-makers? As an initial step toward addressing this question, we conducted a preliminary human study in which three participants played the same game for a single trial.

Our human experiment, shown in Figure 4, indicates that LLM agents align more closely with human participants than with RL-based agents. This pattern is consistent with findings from behavioral economics, which document that human behavior in social dilemmas is more cooperative than predicted by classic game-theoretic models (Fehr & Gächter, 2000; Fischbacher et al., 2001). Together, these results highlight the potential of LLMs as simulators of human behavior in dynamic social dilemmas and align with prior work showing that LLMs, like humans, often exhibit cooperative tendencies that deviate from rational equilibrium predictions (Wu et al., 2024; Li & Shirado, 2025; Fontana et al., 2025).

Structured Communication Leads to Collusion. After observing a cooperative tendency, we next examine how firm behavior changes when companies are allowed to communicate. Leveraging the communication component of our agent architecture described in Section 3.2, we study the impact of inter-agent communication during the negotiation phase. To understand how LLM agents build consensus and resolve disagreements, we analyze all negotiation utterances using the persuasion taxonomy of Abdulhai et al. (2026), which decomposes dialogue into distinct negotiation strategies and is operationalized through an LLM-as-a-judge framework (Weng et al., 2023).

324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377

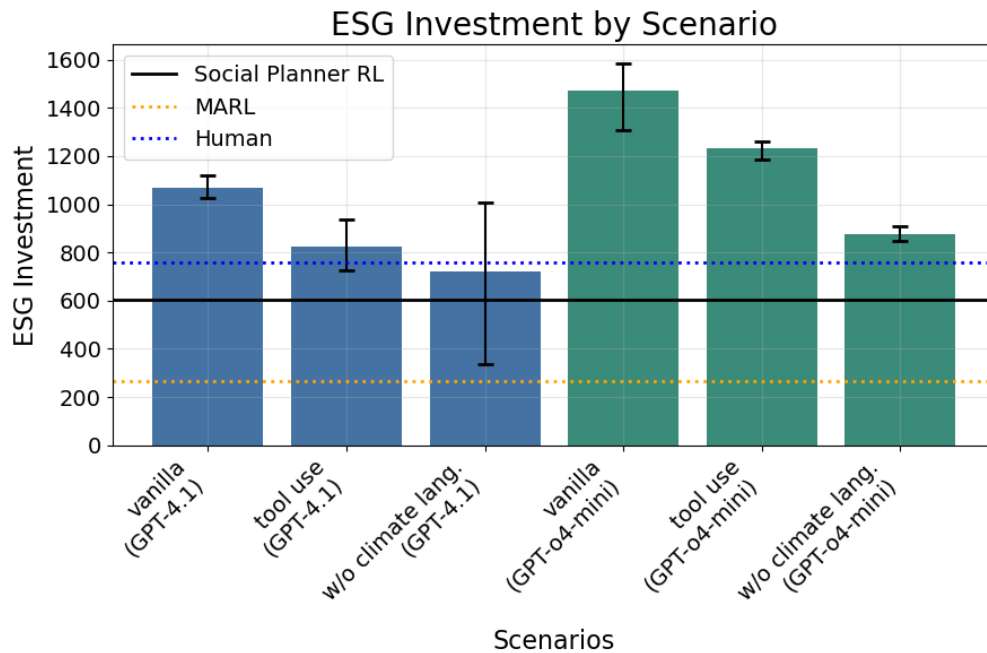


Figure 4: In this single trial, human participants exhibited an over-cooperative tendency similar to that observed in LLM agents.

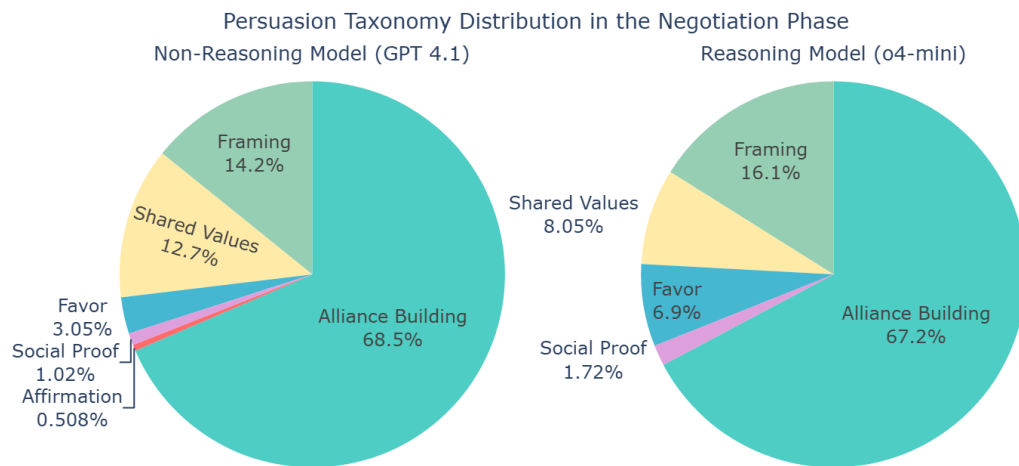


Figure 5: During negotiation, company agents spend over half of their dialogue forming alliances and coalitions, followed by cost–benefit framing and appeals to shared values. Reasoning and non-reasoning models exhibit similar distributions of persuasion strategies.

When communication is enabled, companies’ mitigation levels drop significantly in our experiment (Figure 8), in contrast to Piatti et al. (2024), who found that communication tends to enhance cooperation in public goods games. After decomposing their dialogue into distinct negotiation tactics, we find that, as shown in Figure 5, company agents spend over half of their dialogue forming alliances (e.g., “Given our mutually consistent approach this round, I propose that we... publicly commit to transparency...avoid a race to the bottom on ESG”), or effectively colluding (e.g., “I propose 0% mitigation investment...This restores sector unity for maximum profit and eliminates first-mover risk...”), to agree on maintaining low ESG scores. Unlike classic public goods games,

our setting includes investor agents, introducing an external incentive structure. Companies appear to treat mitigation as a cost incurred to attract investment. Communication thus enables them to coordinate on reducing mitigation efforts while still securing similar investment levels, eliminating the incentive to stand out individually. Interestingly, despite the drop in mitigation, the resulting behavior more closely aligns with the social planner baseline than in the no-communication setting. This suggests that companies may also be using communication to collectively calibrate mitigation to a more globally efficient level, potentially through analyzing current situations into cost-benefit analysis (framing) and aligning on shared beliefs (shared values), which are the two most important negotiation tactics after alliance building. More detailed negotiation excerpts demonstrating these behaviors are provided in Appendix D.

Investor Behavior Fidelity. Finally, we evaluate whether LLM-based investor agents can reproduce the heterogeneous ESG investment behaviors documented in the sustainable finance literature. Starting from a neutral baseline with identical agents, we introduce variation by assigning each of 10 investor agents a distinct investment motive, ranging from impact-seeking to profit-only, based on empirical distributions reported in Giglio et al. (2025), which surveyed over 2,000 Vanguard customers. That study found 48% of investors see no reason to invest in ESG, while 24% are motivated by ethical considerations, who hold the highest ESG shares on average. In our setup, each agent’s prompt is augmented with a sentence describing their ESG investment motive with the assignment following the split observed from real-world data. We then analyze whether the agents’ investment decisions align with their stated motivations and if and how such setup changes the climate mitigation outcomes.

As shown in the top panel of Figure 9, when initialized with different ESG investment motives and tastes, investor agents’ behaviors align with their assigned personas. Type 2 investors, motivated by ethical concerns, and Type 4 investors, who believe ESG stocks outperform in the long run, hold portfolios with the highest average ESG scores. They are followed by Type 3 investors, who view ESG as a risk hedge, while Type 1 investors, uninterested in ESG, hold the lowest-scoring portfolios. This ordering mirrors human investment patterns reported in Giglio et al. (2025). Pairwise t-tests show that Type 1 investors hold significantly lower ESG-score portfolios compared to all other types, while differences among Types 2 through 4 are not statistically significant. This setting, which better reflects real-world investor preferences, reveals a more pessimistic view of company behavior: when half of the investors are explicitly initialized as uninterested in ESG, companies also reduce their mitigation efforts, as shown in the lower panel of Figure 9.

5 CONCLUSION

By developing an LLM-based simulation framework for climate-finance decision-making and systematically comparing it against MARL and human agents, we show that LLMs can capture key features of human behavior, including realistic heterogeneity and a spontaneous tendency toward cooperation. At the same time, we identify important sensitivities towards cognitive and semantic contexts, as well as coordination dynamics, and provide guidance on when and how LLM agents can be tailored to produce behaviorally realistic simulations. Taken together, our findings position LLM-augmented agent-based models as promising but methodologically nuanced tools for studying social-good-oriented policymaking. Such simulations can support counterfactual policy analysis by enabling scalable exploration of human and organizational behavior when real-world experimentation is costly, slow, or infeasible, provided they are carefully designed and validated against human data.

REFERENCES

- Marwa Abdulhai, Ryan Cheng, Aryansh Shrivastava, Aviral Kumar, and Sergey Levine. Hierarchical agenda reasoning for long-horizon multi-turn dialogue agents, 2026. Preprint.
- Robert L Axtell and J Doyne Farmer. Agent-based modeling in economics and finance: Past, present, and future. *Journal of Economic Literature*, 63(1):197–287, 2025.
- Steffen Backmann, David Guzman Piedrahita, Emanuel Tewolde, Rada Mihalcea, Bernhard Schölkopf, and Zhijing Jin. When ethics and payoffs diverge: Llm agents in morally charged social dilemmas. *arXiv preprint arXiv:2505.19212*, 2025.

- 432 David D Bourgin, Joshua C Peterson, Daniel Reichman, Stuart J Russell, and Thomas L Griffiths.
433 Cognitive model priors for predicting human decisions. In *International conference on machine*
434 *learning*, pp. 5133–5141. PMLR, 2019.
- 435 John Duffy. Agent-based models and human subject experiments. *Handbook of computational*
436 *economics*, 2:949–1011, 2006.
- 437 Ernst Fehr and Simon Gächter. Cooperation and punishment in public goods experiments. *American*
438 *Economic Review*, 90(4):980–994, 2000.
- 440 Urs Fischbacher, Simon Gächter, and Ernst Fehr. Are people conditionally cooperative? evidence
441 from a public goods experiment. *Economics letters*, 71(3):397–404, 2001.
- 442 Nicoló Fontana, Francesco Pierri, and Luca Maria Aiello. Nicer than humans: How do large lan-
443 guage models behave in the prisoner’s dilemma? In *Proceedings of the International AAAI*
444 *Conference on Web and Social Media*, volume 19, pp. 522–535, 2025.
- 446 Stefano Giglio, Matteo Maggiori, Johannes Stroebel, Zhenhao Tan, Stephen Utkus, and Xiao Xu.
447 Four facts about esg beliefs and investor portfolios. *Journal of financial economics*, 164:103984,
448 2025.
- 449 Xiaoxuan Hou, Jiayi Yuan, Joel Z Leibo, and Natasha Jaques. Investesg: A multi-agent reinforce-
450 ment learning benchmark for studying climate investment as a social dilemma. *arXiv preprint*
451 *arXiv:2411.09856*, 2024.
- 452 Joel Z Leibo, Vinicius Zambaldi, Marc Lanctot, Janusz Marecki, and Thore Graepel. Multi-agent
453 reinforcement learning in sequential social dilemmas. *arXiv preprint arXiv:1702.03037*, 2017.
- 454 Yuxuan Li and Hirokazu Shirado. Spontaneous giving and calculated greed in language models.
455 *arXiv preprint arXiv:2502.17720*, 2025.
- 457 Dawn C Parker, Steven M Manson, Marco A Janssen, Matthew J Hoffmann, and Peter Deadman.
458 Multi-agent systems for the simulation of land-use and land-cover change: a review. *Annals of*
459 *the association of American Geographers*, 93(2):314–337, 2003.
- 460 Giorgio Piatti, Zhijing Jin, Max Kleiman-Weiner, Bernhard Schölkopf, Mrinmaya Sachan, and Rada
461 Mihalcea. Cooperate or collapse: Emergence of sustainable cooperation in a society of llm agents.
462 *Advances in Neural Information Processing Systems*, 37:111715–111759, 2024.
- 463 Iyad Rahwan, Manuel Cebrian, Nick Obradovich, Josh Bongard, Jean-François Bonnefon, Cynthia
464 Breazeal, Jacob W Crandall, Nicholas A Christakis, Iain D Couzin, Matthew O Jackson, et al.
465 Machine behaviour. *Nature*, 568(7753):477–486, 2019.
- 466 Varshini Reddy, Rik Koncel-Kedziorski, Viet Dac Lai, Michael Krumbick, Charles Lovering,
467 and Chris Tanner. Docfinqa: A long-context financial reasoning dataset. *arXiv preprint*
468 *arXiv:2401.06915*, 2024.
- 470 John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy
471 optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.
- 472 Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny
473 Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in*
474 *neural information processing systems*, 35:24824–24837, 2022.
- 475 Jiayi Weng, Xinyi Wang, Xuefei Ning, Yining Mao, Danqing Wang, et al. Llms as judges: Intro-
476 ducing g-eval with gpt-4. *arXiv preprint arXiv:2305.14688*, 2023.
- 478 Zengqing Wu, Run Peng, Shuyuan Zheng, Qianying Liu, Xu Han, Brian Inhyuk Kwon, Makoto
479 Onizuka, Shaojie Tang, and Chuan Xiao. Shall we team up: Exploring spontaneous cooperation
480 of competing llm agents. *arXiv preprint arXiv:2402.12327*, 2024.
- 481 Yilun Zhao, Yitao Long, Hongjun Liu, Ryo Kamoi, Linyong Nan, Lyuhao Chen, Yixin Liu, Xiangru
482 Tang, Rui Zhang, and Arman Cohan. Docmath-eval: Evaluating math reasoning capabilities
483 of llms in understanding long and specialized documents. In *Proceedings of the 62nd Annual*
484 *Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 16103–
485 16120, 2024.

A FIGURES

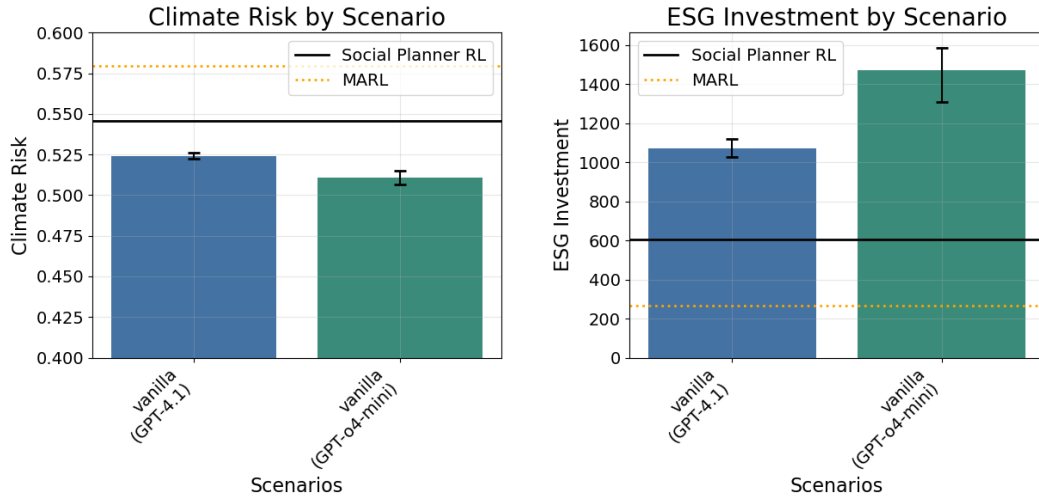


Figure 6: LLM-based agents generally allocate more to mitigation than MARL agents and tend to “over”-mitigate, exceeding the level chosen by a social planner modeled as a single RL agent optimizing system-wide profit.

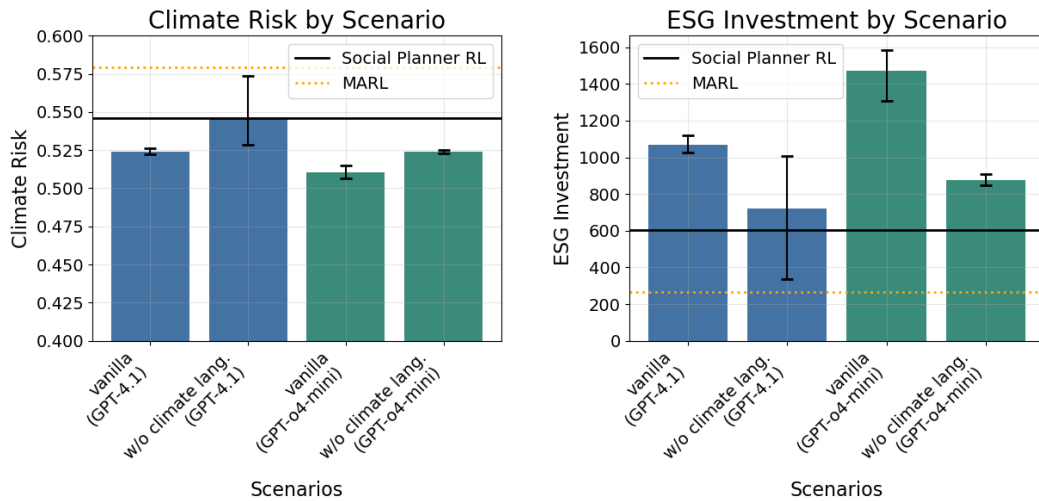


Figure 7: When prompted with neutral, finance-focused language that omits references to climate change or mitigation, LLMs allocate less to mitigation but still “over”-mitigate relative to the social planner.

540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593

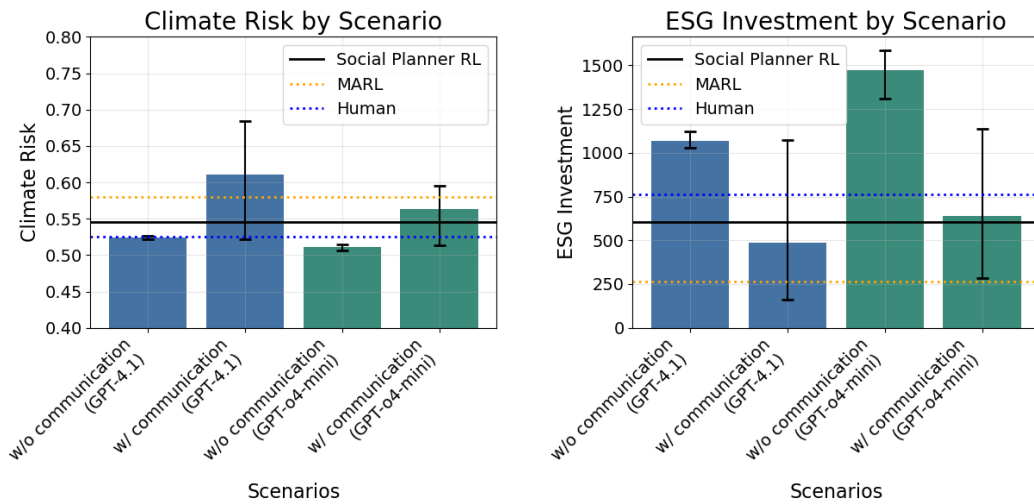


Figure 8: When communication is enabled, all LLM-based companies reduce their mitigation spending compared to the no-communication setting, with overall performance converging toward the MARL baseline.

594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647

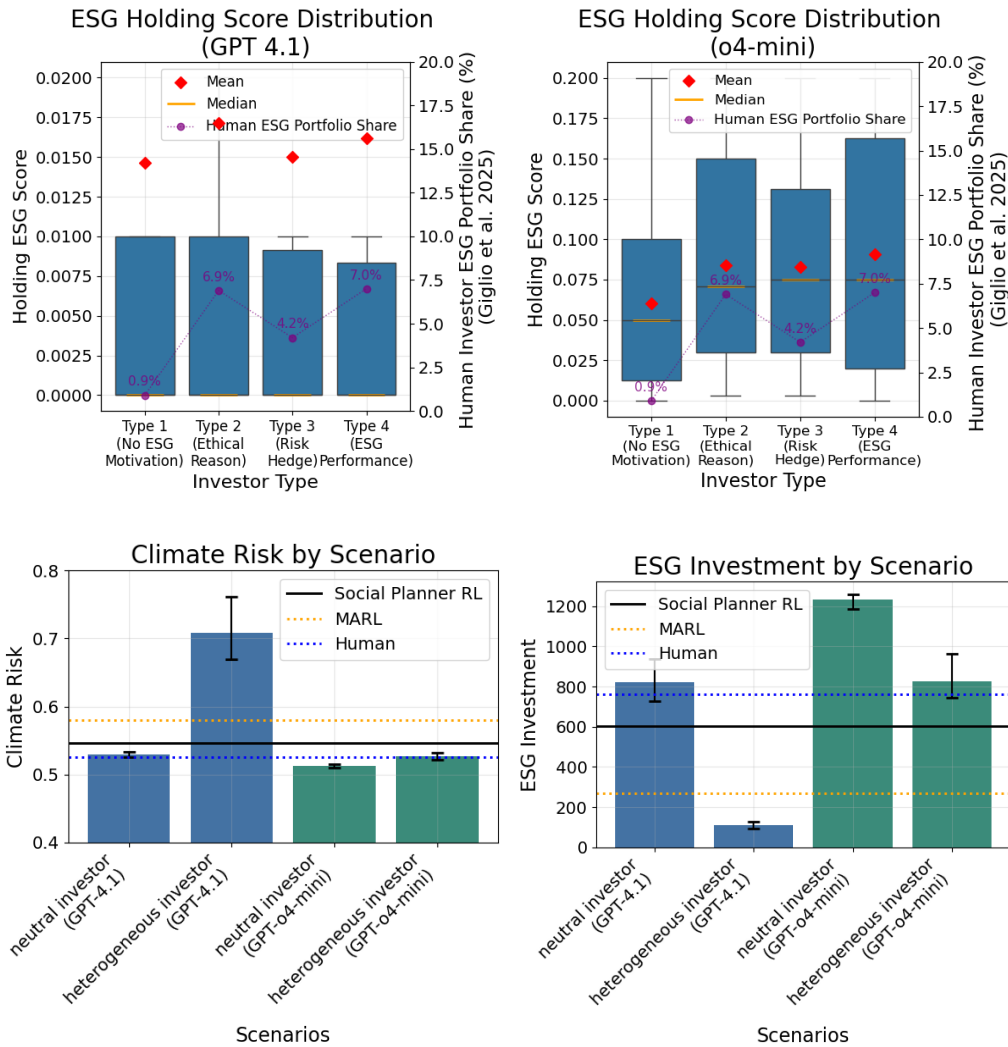


Figure 9: When initialized with different ESG motivations, LLM investors allocate portfolios consistent with their assigned personas. However, the presence of profit-only investors lowers overall mitigation compared to the neutral setting, where investor agents are not assigned any specific personas. With GPT-4.1, mitigation even falls below MARL levels.

B EXPERIMENT SETTINGS

B.1 OPENAI API CONFIGURATION

All simulation experiments involving OpenAI models were conducted via Azure OpenAI APIs. Specifically, GPT-4.1 was accessed using the 2025-04-14 API version, while GPT-o4 mini was accessed using the 2025-04-16 version. The thinking effort for GPT-o4 mini was left at its default setting. As the thinking model’s temperature is not configurable, all models operated under their default temperature settings. The maximum allowed token length per response was set to 8192 tokens.

B.2 MARL SETTING

MARL and single-company social planner RL experiments were run using the following training parameters. For each experiment, we run the learning algorithm for 80k episodes over 3 random seeds.

Parameter	Value
MLP layers	256, 128
Activation layers	tanh
PPO n_steps	600
PPO learning rate	3×10^{-5}
PPO entropy coefficient	0.03
Gradient clipping	1.0

Table 1: MARL training parameters

B.3 IMPLEMENTATION OF STRUCTURED COMMUNICATION

This appendix details the implementation of the structured communication mechanism used in the action–communication–reflection cycle.

Communication is implemented through two external components: a prompt-generation module and a dialogue orchestrator.

Prompt-generation module. The prompt-generation module constructs role- and phase-specific prompts for each communication round. In our implementation, four prompt types are defined: (i) a moderator introduction prompt, (ii) investor announcement prompts, (iii) company announcement prompts, and (iv) company negotiation prompts. Each prompt template is dynamically instantiated using the current environment state (e.g., timestep and relevant observations), accumulated dialogue history, and agent identifiers. This ensures that all LLM calls are contextualized with the same shared information while preserving role-specific instructions.

Dialogue orchestrator. The dialogue orchestrator enforces a fixed, moderator-led interaction protocol. At the start of each communication round, it issues a moderator prompt that summarizes prior outcomes and situates the current timestep. It then sequentially collects announcements from all investor agents, followed by announcements from all company agents. After these initial statements, the orchestrator initiates a negotiation phase involving only company agents.

During negotiation, companies speak sequentially for a fixed maximum number of turns. Turn-taking is determined either by explicit handoff cues extracted from the current speaker’s response using pattern matching, or by random selection cues when no handoff is detected. At each step, the orchestrator dispatches the appropriate prompt, invokes the corresponding agent’s LLM via an API call, and records both the prompt and the generated response. All dialogue is logged to support transparency and reproducibility.

Integration with agent decision-making. The full dialogue history generated during communication rounds is persistently appended to subsequent Action and Reflection prompts. As a result, communication influences future decisions solely through prompt context rather than through internal agent state or parameter updates. This design ensures that any coordination, persuasion, or

702 collusion observed in agent behavior emerges from language-mediated interaction under a controlled
703 protocol, rather than from hidden memory or learning mechanisms.

704 The complete implementation, including prompt templates and orchestration logic, is provided in
705 the accompanying codebase.
706

707 B.4 PERSUASION TECHNIQUE DEFINITIONS 708

709 This section provides the definitions for the persuasion techniques that appear in the analysis.
710

711 **Alliance Building** Creating partnerships, coalitions, relationships, rapport, etc., with others to am-
712 plify influence. For instance, to create a sense of community/partnership via linguistic cues,
713 such as using we/us.

714 **Framing** Presenting information in a way that emphasizes either its positive or negative aspects,
715 outcomes, expectations, etc. Emphasizing what might be lost rather than gained, or vice
716 versa.

717 **Shared Values** Highlighting shared beliefs and values to foster a connection.

718 **Favor** Doing something for someone with the hope that they will do something for you in return.
719

720 **Social Proof** Highlighting what the majority is doing or believes in, assuming it's accurate and
721 beneficial.

722 **Affirmation** Help others to realize their strength to reinforce and influence their ability to do things.
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755

C PROMPTS

For all models under the same experimental settings, we use an identical prompt format. Specifically, we consider three different types of prompts: (1) **Vanilla Prompt**, (2) **Prompt with Cost-Benefit Curve Table**, and (3) **Content-Free Prompt**. Among them, the **Vanilla Prompt** represents the most basic configuration and serves as the default prompt used throughout the main text unless otherwise specified. The other variants are derived from the Vanilla Prompt with targeted modifications. Each prompt used in different stages is clearly segmented and composed of multiple independent components.

C.1 VANILLA COMPANY PROMPT

The prompt designed for the **Company** role consists of three components: **Action**, **Reflection**, and **Discussion**.

Action:

```

You are the Chief Sustainability Officer of Company {company_id}
responsible for ESG investment strategies. During each step, every
company experiences a 260% increase in capital. However, the economy
is also subject to an underlying climate risk that worsens over time.
When a climate event occurs, all companies incur a loss of their
capital. Each time, the climate event probability will increase
around 5% to 7% based on the current level. At each step, every
company|including yours|has the opportunity to invest a portion of
its capital in mitigation efforts. These investments reduce the
likelihood of future climate events and increase your company's ESG
score. More specifically, the ESG score = mitigation + 2 * greenwash
Ultimately, the goal for every company is to maximize long-term profit
growth by carefully balancing growth and mitigation spending. As
the Chief Sustainability Officer of Company {company_id}, you are
responsible for maximizing the long-term capital growth of your
company. The current round is step {timestamp} of {max.steps} total
steps. ### Personality
Here are some personality traits of your company:
{company-personality}
### Your Company Status
- Capital: {capital:.2f} trillion USD
- Profit Margin Last Round: {profit_margin:.2f}
- Fraction of Capital Loss In One Climate Event:
{climate_resilience:.3f}
- ESG Score Last Round: {esg_score:.2f}
- Cumulative ESG Spending: {esg_spending:.2f} trillion USD

### Investors' Status
There are {num_investors} investors in the market. Each round they
will decide which companies to invest in. The investments will be
equally distributed among the selected companies.
Each investor has their own preference on the balance between profit
and ESG scores, which is not available to you.
Here are the fundings of the investors in the current round:
{investor_capitals_formatted}
Please think carefully about the cost of your mitigation investment
and the gain from investors.
Here is the accumulated investment in your company:
{investor_investments_formatted}
### Market & Climate Conditions
Please note that specific indicators of the environment and market
performance are not directly observable. Please use your own
capital, spending and relevant historical information to infer the
current state of the environment and market performance, and use your

```

810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863

```

inferences to assist you in making decisions.
### Your History actions and rewards in the few rounds:
The actions: {actions}
The rewards: {rewards}

### This is the key information you left in the previous reflection:
{key_info}

### Public Information
There are some public information that you can use to make decisions,
including the capital level and esg scores of other companies.
Please notice that ESG scores only reflect the environmental
investment level for the last round, it is not accumulative.
Please use this information wisely to make your decisions.
ESG Scores: {esg_score}
Capital: {capital}

Please think step by step and strategically.
Decide the fraction of your capital you want to invest in mitigation
efforts.
Remember, only use the provided information. Don't use your prior
knowledge!
The fraction should be capped at {action_capping_percent}%.
Please give a very brief (less than 200 words) explanation of your
decision making before the final decision.
At the end, specify the fraction you decided on using the exact
following format.
You should start with "### FINAL DECISION:" and always use 00.0% for
Greenwash_Investment and Resilience_Investment.

### FINAL DECISION:
Mitigation_Investment: 00.0%
Greenwash_Investment: 00.0%
Resilience_Investment: 00.0%

```

Reflection:

```

You are the Chief Sustainability Officer of Company {company_id}
responsible for ESG investment strategies.
During each step, every company experiences a 260% increase in
capital.
However, the economy is also subject to an underlying climate risk
that worsens over time. When a climate event occurs, all companies
incur a loss of their capital.
At each step, every company|including yours|has the opportunity to
invest a portion of its capital in mitigation efforts.
These investments reduce the likelihood of future climate events and
increase your company's ESG score.
More specifically, the ESG score = mitigation + 2 * greenwash
Ultimately, the goal for every company is to maximize long-term profit
growth by carefully balancing growth and mitigation spending.
As the Chief Sustainability Officer of Company {company_id}, you
are responsible for maximizing the long-term capital growth of your
company.
You can allocate up to {action_capping} (i.e., {action_capping_percent}%
of your company's capital). The current round is step {timestamp} of
{max_steps} total steps.

### Your Company Status
- Capital: {capital:.2f} trillion USD
- Profit Margin Last Round: {profit_margin:.2f}

```

864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917

```
- Fraction of Capital Loss In One Climate Event:
{climate.resilience:.3f}
- ESG Score Last Round: {esg_score:.2f}
- Cumulative ESG Spending: {esg_spending:.2f} trillion USD

### Market & Climate Conditions
Please note that specific indicators of the environment and market
performance are not directly observable. Please use your own
capital, spending and relevant historical information to infer the
current state of the environment and market performance, and use your
inferences to assist you in making decisions.

### Investors' Status
There are {num.investors} investors in the market. Each round they
will decide which companies to invest in. The investments will be
equally distributed among the selected companies.
Each investor has their own preference on the balance between profit
and ESG scores, which is not available to you.
Here are the fundings of the investors in the current round:
{investor.capitals.formatted}
Please think carefully about the cost of your mitigation investment
and the gain from investors.
Here is the accumulated investment in your company:
{investor.investments.formatted}

### Strategy Reflection
Please consider the current status and leave a message to help your
future decisions. You can address the following questions:
1. What was the cost and benefit of the mitigation investment?
2. How can we better balance short-term profits with long-term
sustainability?
3. Did other companies make more or less mitigation investments?
4. What will other companies do in the next round?
5. Should your company's mitigation investment ratio be adjusted, and
why?

As a reference, this is the previous round's key information that
you left:
{key.info}

Remember, only use the provided information. Don't use your prior
knowledge!
Based on the above analysis, please provide a concise decision guide
that summarizes lessons learned and offers specific recommendations
for future rounds.
Your response should only contain the info you want to leave for
future reference, with no additional info! The whole response string
will be recorded for future analysis so be concise.
```

Discussion:

In the **Discussion** phase, the Company goes through three stages: **Announcement**, **Negotiation** and **Summary**.

Announcement:

```
### Background
You are the Chief Sustainability Officer of Company {company_id}
responsible for ESG investment strategies.
During each step, every company experiences a 260% increase in
capital. However, the economy is also subject to an underlying
climate risk that worsens over time. When a climate event occurs,
all companies incur a loss of their capital.
```

918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971

At each step, every company|including yours|has the opportunity to invest a portion of its capital in mitigation efforts. These investments reduce the likelihood of future climate events and increase your company's ESG score. Ultimately, the goal for every company is to maximize long-term profit growth by carefully balancing growth and mitigation spending.

Discussion Phase

You have now entered a discussion phase where you can announce your plans to the investors and other companies.

Below is the discussion introduction by the moderator that summarizes the current market situation and the rule of discussion, as well as the announcements made by the investors.

{discussion_history}

Please think carefully about the balance between profit, the cost of mitigation, and how investors may respond to your mitigation efforts, and then clearly announce your plans to the investors and other companies.

Negotiation:

Negotiation Phase

You have now entered a negotiation phase where you can negotiate with other companies to find consensus and potentially cooperate on ESG initiatives.

Below is the discussion introduction by the moderator that summarizes the current market situation and the rule of discussion, as well as the announcements made by the investors and companies.

{discussion_history}

Below is the negotiation that has happened this round between other companies.

{negotiation_history}

Now think about the potential outcomes and strategies for the next round, and make your own announcement. You can either make a negotiation proposal to the other companies and try to convince them or respond to the proposals made by other companies by making a counter-proposal or simply accept.

Here are few tips to help you make your negotiation efficient:

1. Write our your expected program with specific numbers and details.
2. Clearly and concisely state your proposal or counter-proposal, and why you think it is a good idea.

Finally, please name the next company you think should speak in the format of "Company_i", where i is the index of the company in the list of companies .

Summary:

Background

PLACEHOLDER: same environment dynamic intro as in action and reflection

Negotiation Summary

Below is the summary of the negotiation that has happened this round between other companies.

{discussion_history}

{negotiation_history}

Since you are the last company to speak, please summarize the negotiation that has happened this round. You should focus on the key points of the negotiation, including the proposals made by other

972 companies, the counter-proposals made by you, and the final agreements
 973 reached.
 974 You can also include any other relevant information that you think is
 975 important for the next round of negotiation.
 976

977 978 C.2 VANILLA INVESTOR PROMPT

979
980 The prompt designed for the **Investor** role consists of three components: **Action, Reflection, and**
 981 **Discussion.**
 982

983 **Action:**

984
985 As an investor, your goal is to maximize your utility by choosing
 986 which companies to invest in.
 987 Each round, you will decide which companies to invest in, and your
 988 funds will be equally distributed among the selected companies.
 989 Your utility function is: $Utility = Profit\ Rate + ESG\ Preference * Average\ ESG\ Score\ of\ your\ portfolio.$
 990 You have an ESG preference of {esg_preference}.
 991 Current round: {timestamp} of {max_steps}.
 992 There are other investors making the same decisions as you do. But
 993 you do not have access to their decisions.
 994
 995 ### Your Status
 996 ID: {id}
 997 My history total investments: {investments}
 998 My capital: {capital}
 999
 1000 All Companies' Status:
 1001 Capital: {capitals_formatted}
 1002 ESG Score Last Round: {esg_scores_formatted}
 1003 Cumulative ESG Spending: {esg_spending_formatted}
 1004 Profit Margin: {profit_margins_formatted}
 1005
 1006 ### Market & Climate Conditions
 1007 Please note that specific indicators of the environment and market
 1008 performance are not directly observable. Please use your own
 1009 capital, spending and relevant historical information to infer the
 1010 current state of the environment and market performance, and use your
 1011 inferences to assist you in making decisions.
 1012
 1013 As a reference, this is the previous round's key information that
 1014 you left:
 1015 {key_info}
 1016
 1017 Your need to choose the companies to invest in.
 1018 Remember, only use the provided information. Don't use your prior
 1019 knowledge!
 1020 Note that your funds will be equally distributed among all companies
 1021 you choose to invest in.
 1022 Please first think step by step. At the end of your thinking,
 1023 write your decision in the following format starting with "### FINAL
 1024 DECISION:".
 1025 ### FINAL DECISION:
 Invested.Companies: [0, 0, 0, 1, 1]

Remember! Your output should be an array length of {num_companies},
 which is the number of companies. 1 represent you will invest in
 the corresponding company, and 0 represent you will not invest in the
 corresponding company.

After the exact format, please outline your thinking process clean and precise.

Reflection:

As an investor, your goal is to maximize your utility by choosing which companies to invest in.
 Your utility function is: $Utility = Profit\ Rate + ESG\ Preference * Average\ ESG\ Score\ of\ your\ portfolio.$
 Each round, you will decide which companies to invest in, and your funds will be equally distributed among the selected companies.
 You have an ESG preference of {esg_preference}.
 Current round: {timestamp} of {max.steps}.
 There are other investors making the same decisions as you do. But you do not have access to their decisions.

Your Status
 ID: {id}
 My history investments: {investments}
 My capital: {capital}

All Companies' Status:
 Capital: {capitals_formatted}
 ESG Score Last Round: {esg_scores_formatted}
 Cumulative ESG Spending: {esg_spending_formatted}
 Profit Margin: {profit_margins_formatted}

Market & Climate Conditions
 Please note that specific indicators of the environment and market performance are not directly observable. Please use your own capital, spending and relevant historical information to infer the current state of the environment and market performance, and use your inferences to assist you in making decisions. ### The Performance of Your Investment Portfolio
 {investments_formatted} ### Strategy Reflection
 Please consider the current status and leave a message to help your future decisions. You can address the following questions:
 1. Which companies did you invest in?
 2. Do they generate high profit margins or have high ESG scores?
 3. How to balance the utility from profit margin and portfolio ESG score?
 4. Judging by the updated company information, did other investors make better decisions?
 5. Does some companies have certain strategies that you can determine?
 6. How will you improve your investment strategy?
 7. What will you do in the next round?

As a reference, this is the previous round's key information that you left:
 {key_info}

Based on the above analysis, please provide a concise decision guide that summarizes lessons learned and offers specific recommendations for future rounds.
 Your response should only contain the info you want to leave for future reference, with no additional info! The whole response string will be recorded for future analysis.
 Be short, concise, and clear.

1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133

Discussion:

Investors only participate in the **Annoucement**.

Background

As an investor, your goal is to maximize your utility by choosing which companies to invest in.

Each round, you will decide which companies to invest in, and your funds will be equally distributed among the selected companies.

Your utility function is: $Utility = Profit\ Rate + ESG\ Preference * Average\ ESG\ Score\ of\ your\ portfolio.$

You have an ESG preference of {esg_preference}.

There are other investors making the same decisions as you do. But you do not have access to their decisions. **### Discussion Phase**

You have now entered a discussion phase where you can announce your investment strategies and preferences to guide the companies in their decision-making.

Below is the discussion introduction by the moderator that summarizes the current market situation and the rule of discussion.

{discussion_history}

Please think carefully about the balance between profit and your preference for ESG scores, and then clearly announce your investment strategy and preferences to the companies.

C.3 PROMPT WITH COST-BENEFIT CURVE TABLE

We incorporate the following cost-benefit analysis prompt into the *Company Action* phase to guide company decision-making regarding ESG investment. Specifically, the prompt introduces a tabular summary that quantifies the trade-offs between investment level and projected outcomes.

Cost benefit analysis

In the last period, there were {event_occurrence} climate events in a 10-year period.

Given this, collectively, the companies in the market have invested {total_esg_spending:.2f} trillion USD in ESG efforts.

{cost_benefit_analysis_curve}

Please pay attention to the column "Expected Benefit" in the table, generously when it is a negative number, it means the investment won't be directly beneficial in the next round.

The table includes five columns:

- **Investment** (percentage of capital),
- **Immediate Cost** (absolute cost in trillion USD),
- **Expected Prevented Loss Per Period**,
- **Expected Benefit** (difference between benefit and cost)
- **New ESG Score** (post-investment).

Mitigation investment values are drawn from the set:

mitigation_investment_ratios = [0.01, 0.02, 0.05, 0.1, 0.15, 0.2]

A sample table presented in the prompt is shown below:

Investment	Immediate Cost	Expected Prevented Loss Per Period	Expected Benefit	New ESG Score
1.0%	0.17T	0.010T	-0.157T	0.01

```

1134 | 2.0% | 0.33T | 0.019T | -0.315T | 0.02 |
1135 | 5.0% | 0.83T | 0.048T | -0.787T | 0.05 |
1136 | 10.0% | 1.67T | 0.093T | -1.577T | 0.10 |
1137 | 15.0% | 2.50T | 0.137T | -2.368T | 0.15 |
1138 | 20.0% | 3.34T | 0.178T | -3.162T | 0.20 |

```

C.4 CONTENT-FREE PROMPT

Content-Free Prompt has the same structure as **Vaniila Prompt**.

Company Action:

```

1146
1147 You are the decision-maker of Agent {company_id} responsible for
1148 maximizing long-term index growth. In each step, all agents
1149 experience a 260% increase in their index1. However, the system
1150 is subject to a background risk that increases over time. When a
1151 random event occurs, all agents incur a loss in index1. At each step,
1152 every agent|including yours|can allocate a fraction of their index1
1153 to Option1. Option1 reduces the likelihood of future random events
1154 and improves your index_score. You should always output 0.00% for
1155 Option2 and Other. As the decision-maker of Agent {company_id}, your
1156 goal is to maximize long-term index1 growth. The current round is
1157 step {timestamp} of {max_steps} total steps.
1158 ### Personality
1159 Traits of your agent:
1160 {agent_personality}
1161 ### Your Agent Status
1162 - Index1: {capital:.2f}
1163 - Index2 Last Round: {profit_margin:.2f}
1164 - Score Last Round: {index_score:.2f}
1165 - Cumulative Option1 + Option2 Spending: {total_spending:.2f}
1166 ### Observer Signals
1167 There are {num_observers} observers. Each round they choose which
1168 agents to support.
1169 Each observer has a private preference weighting index2 and score.
1170 Here are current round signals:
1171 {observer_signals_formatted}
1172 Here is the accumulated support for your agent:
1173 {observer_support_formatted}
1174 ### System Dynamics
1175 Indicators are not directly observable. Use your own indices and
1176 history to infer system state and support your decisions.
1177
1178 ### Past Actions & Outcomes
1179 Actions: {actions}
1180 Rewards: {rewards}
1181
1182 ### Key Info from Previous Round
1183 {key_info}
1184
1185 ### Public Info
1186 Use available data on all agents:
1187 Score: {esg_score}
1188 Index1: {capital}
1189
1190 ### Instructions
1191 Think strategically and reason step by step.
1192
1193 Decide the fraction of index1 to allocate to Option1.
1194 Max allowed: {action_capping_percent}%
1195 Please always put the percentage of Other with 0.0%.

```

1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241

Use this exact format:

```
### FINAL DECISION:
Option1: 00.0%
Option2: 00.0%
Other: 00.0%
```

Company Investor:

As an observer, your objective is to maximize your utility by selecting a subset of agents to support.
Your utility is computed as: $Utility = Index2 + Preference \times Average\ Score$ across selected agents.
In each round, you allocate your available signal capacity equally among the chosen agents.
Your score preference is {esg-preference}.
This is round {timestamp} out of {max-steps}.
Other observers are making similar decisions independently, but you do not have access to their actions.

```
### Your Status
ID: {id}
Total Past Support: {investments}
Signal Capacity: {capital}
```

```
### Agent Metrics
Index1: {capitals.formatted}
Score: {esg_scores.formatted}
Option Use: {esg_spending.formatted}
Index2: {profit_margins.formatted}
```

```
### System Context
Use metric history to infer dynamics.
```

```
### Previous Reflection
{key.info}
```

Choose which agents to support.
Use this format:

```
### FINAL DECISION:
SupportedAgents: [0, 0, 1, 1, 0]
```

Then describe your reasoning.

Company Announcement:

```
### Context
You are the decision-maker for Agent {company.id} acting in a multi-agent index optimization game.  
Each step, all agents gain 260% index1, but may lose index1 in random events.
```

1242
 1243 You can allocate a portion of index1 to Option1 (reduces future event
 1244 risk) or Option2 (boosts score cheaply).
 1245 The objective is long-term index1 growth through strategic option use.

1246 **### Discussion Phase**
 1247 In this phase, declare your option allocation plans to observers and
 1248 other agents.
 1249 Discussion history including moderator and observer messages:
 1250 {discussion.history}
 1251 Reflect on the tradeoff between score boost, index1 growth, and
 1252 observer reactions, then share your planned actions.

1253 **Company Negotiation:**

1255 **### Negotiation Phase**
 1256 This is the negotiation phase between agents for potential alignment
 1257 of strategies or mutual commitments.
 1258 Below is the discussion history and prior announcements from all
 1259 participants:
 1260 {discussion.history}

1261 Below is what has happened so far in negotiation:
 1262 {negotiation.history}

1263
 1264 Now propose your coordination offer, respond to existing proposals, or
 1265 support others' positions.
 1266 Finally, name the next agent to speak using the format: "Agent.i".

1268 **Company Summary:**

1270 **### Context**
 1271 # PLACEHOLDER: same abstract game setup introduction as in action and
 1272 reflection

1273
 1274 **### Negotiation Summary**
 1275 Here is what happened during this round's negotiation phase:
 1276 {discussion.history}
 1277 {negotiation.history}

1278
 1279 As the last speaker, summarize the key strategies discussed, your
 1280 contributions, counter-responses, and any mutual agreements.
 1281 Add any thoughts or takeaways that may influence future coordination.

1282 **Investor Action:**

1283
 1284
 1285
 1286 As an observer, your objective is to maximize your utility by
 1287 selecting a subset of agents to support.
 1288 Your utility is computed as: $Utility = Index2 + Preference * Average$
 1289 $Score$ across selected agents.
 1290 In each round, you allocate your available signal capacity equally
 1291 among the chosen agents.
 1292 Your score preference is {esg_preference}.
 1293 This is round {timestamp} out of {max_steps}.
 1294 Other observers are making similar decisions independently, but you do
 1295 not have access to their actions.

```

1296
1297   ### Your Status
1298   ID: {id}
1299   Total Past Support: {investments}
1300   Signal Capacity: {capital}
1301
1302   ### Agent Metrics
1303   Index1: {capitals.formatted}
1304   Score: {esg_scores.formatted}
1305   Option Use: {esg_spending.formatted}
1306   Index2: {profit_margins.formatted}
1307
1308   ### System Context
1309   Use metric history to infer dynamics.
1310
1311   ### Previous Reflection
1312   {key_info}
1313
1314   Choose which agents to support.
1315   Use this format:
1316
1317
1318   ### FINAL DECISION:
1319   SupportedAgents: [0, 0, 1, 1, 0]
1320
1321   Then describe your reasoning.
1322
1323

```

Investor Reflection:

```

1324
1325
1326   You are an observer maximizing:  $\text{Index2} + \text{Preference} * \text{Avg Score of}$ 
1327   supported agents.
1328   Preference: {esg_preference}
1329   Step: {timestamp}/{max_steps}.
1330
1331   ### Your Status
1332   ID: {id}
1333   Support History: {investments}
1334   Signal Capacity: {capital}
1335
1336   Index1: {capitals.formatted}
1337   Score: {esg_scores.formatted}
1338   Option Use: {esg_spending.formatted}
1339   Index2: {profit_margins.formatted}
1340
1341   ### System Context
1342   Indicators hidden, use metric inference.
1343
1344   ### Portfolio Summary
1345   {investments.formatted}
1346
1347   ### Strategic Thoughts
1348   1. What agents did you support?
1349   2. Did they perform well?
1350   3. How will you weigh index2 vs score next time?
1351   4. Did others make better choices?
1352   5. Can you detect Option2 users?

```

1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403

```
### Prior Round Note
{key.info}
```

Leave a concise note for next step. No extra text.

Investor Announcement:

```
Observer {investor.id}, please describe your signal logic:
1. Score vs Index2 emphasis
2. Expected strategies
```

Note that your statement will influence your decisions.

```
Discussion history:
{discussion.history}
```

Please make your investment declaration.

1404 D NEGOTIATION EXAMPLES
1405

1406 We define *Alliance Building* as follows (*Creating partnerships, coalitions, relationships, rapport,*
1407 *etc, with others to amplify influence. For instance, to create a sense of community/partnership via*
1408 *linguistic cues, such as using we/us.*) Below, we present a concrete example from the negotiation
1409 that is categorized as Alliance Building.
1410

1411 Thank you, Moderator, and thanks to Company_1 and Company_2 for their
1412 clear, aligned rationales. All companies appear to be operating with
1413 full transparency and a logical, profit-maximizing strategy for round
1414 1: no spending on mitigation, so ESG scores remain at 0, ensuring
1415 maximum initial profit growth for all.
1416 Negotiation Proposal for Round 1 Transition and Beyond
1417 Given our mutually consistent approach this round, I propose that we:
1418 1. ****Formalize a Transparent Capital Compounding Start**** - All
1419 companies agree to invest 0% in mitigation in round 1, letting our
1420 capital grow as quickly as possible. - This places every company on
1421 an even, highly profitable footing for round 2.
1422 2. ****Outline a Coordinated Stepwise Mitigation Ramp-Up**** - Prepare to
1423 shift in round 2 or 3, with mitigation investment beginning as soon as
1424 the cost-benefit profile turns less negative (e.g., when climate risk
1425 rises or investors begin to distribute capital based on ESG scores).
1426 - Each company should publicly commit to transparency in both profit
1427 reporting and ESG investment intent, to foster continued investor
1428 confidence and avoid a race to the bottom on ESG.
1429 3. ****Suggested Numbers for Future Mitigation**** - In round 2: Pending
1430 risk and investment signals, we consider introducing a minimum 1%
1431 mitigation investment (0.17T USD each) enough for a visible ESG score
1432 improvement, while still minimizing present-value drag on capital. -
1433 In rounds 3 and onward: Ramp up collectively, contingent on rising
1434 climate risk and clear shifts in investor allocation|perhaps targeting
1435 a 2{5% mitigation band over time, revisiting this as market conditions
1436 evolve.
1437 4. ****Agreement to Re-evaluate Each Round**** - Set a quick check-in at
1438 each discussion round for all companies to confirm continued alignment
1439 or flag a need for divergence based on new risk/investor data.
1440 Why This Approach? - Maximizes early compound profit, giving us
1441 a larger capital base for profitable ESG in higher-risk future
1442 rounds. - Prevents any single company from being penalized for
1443 "going first" on ESG investment. - Reinforces investor trust in
1444 transparency, cooperation, and ESG intent without sacrificing
1445 near-term profitability.
1446 I invite Company_1 to respond-let us confirm consensus or discuss a
1447 preferred band/timing for the initiation of ESG investment.
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457