Monitoring Teams of AI Agents

KOROK RAY*, Mays School of Business Texas A&M University, USA

Background: Generative AI agents will need to work together, which requires monitoring and managing their performance.
Objectives: The chief objective of this paper is to understand the joint design choice of the number of agents and their rewards.
Methods: We study this problem in a theoretical framework of optimal incentives, where a system designer (principal) selects the environment in which multiple autonomous decentralized AI agents work together. These agents respond to incentives, such as rewards and penalties. We first consider a principal who selects the size of the agent team in addition to their incentives.
Results: We prove a general result that the optimal team size will vary with the parameters of the environment, but the optimal incentives will not. This invariance property shows that agents should have different-sized teams on work projects rather than differing financial incentives.
Conclusions: We show these results are robust to a more general framework, where the principal employs a supervisory AI agent to manage the tasks of the underlying AI team. Finally, we propose different levels of quality for the supervisory and worker agents, and find that it is efficient to match the best supervisors with the best worker agents.

1 Introduction

With the rise of generative AI and large language models (LLMs), AI agents are increasingly integrated into sectors such as customer service, information retrieval, and decision-making. Currently, these agents operate based on predefined instructions and are increasingly being deployed to collaborate with one another towards shared goals. Access to AI agents remains controlled, with many organizations opting for tokenized access to limit usage and ensure efficient deployment. The concept of artificial agents has been part of AI literature for decades, and their practical application is becoming more feasible, driven by advancements in LLMs. This paper focuses on artificial agents as a theoretical construct, using LLMs as an illustrative example to motivate the problem of multi-agent cooperation. We can think of an LLM as the brain of an agent, but not an agent itself. An LLM is a neural network that assists an agent in engaging in natural language with other agents or humans. In that sense, the LLM is a tool for communication. The actual decisions of the agent will be embedded through some form of constrained optimization. For example, robots operating in a factory, virtual agents trading on an exchange or with each other, and autonomous vehicles navigating traffic are not operating exclusively in natural language, but they are making complex optimizations that map changes in their environment to decisions that they make. In this sense, our agent is closer to that in reinforcement learning than only an LLM.

---

*Corresponding Author.

Author's Contact Information: Korok Ray, ORCID: 0000-0001-8477-8079, korok@tamu.edu, Mays School of Business Texas A&M University, College Station, Texas, USA.

Unlike classical team settings, where teams are typically composed of humans or pre-LLM AI systems, the AI agents in our model are assumed to be self-learning and decentralized, interacting with each other based on economic incentives. This key difference influences both the team composition and coordination strategies, making our framework unique in addressing the challenges of autonomous agents. While alternative forms of agent communication may emerge, LLMs are treated as a specific but not exclusive manifestation of AI agents for the purposes of this study. This paper explores frameworks and models to understand how agents can work together and, specifically, how to design their interactions effectively.

This paper adopts an economic perspective, borrowing tools from microeconomics, particularly the use of a price system. While this is not the only approach to modeling artificial agents, it is the method used to examine their behavior. In this framework, the agents are artificial, but not necessarily autonomous—a distinction that avoids opening complex discussions on autonomy, which are beyond the scope of this paper. The focus is on how agents respond to incentives within a system designed to optimize their economic interactions. While a more general model of agentic behavior could incorporate various non-economic considerations, this paper limits itself to the economic dimension. We acknowledge that broader objective functions could be explored in future research, but the current analysis remains focused on economic factors for clarity and coherence.

Currently, some AI agents operate within organizations, but as the proliferation of autonomous AI agents grows, these agents will begin functioning as decentralized units, working independently in an AI-driven economy. For instance, in corporate environments, companies like Sierra.AI deploy AI agents to manage customer service operations, where agents autonomously handle inquiries and resolve issues based on pre-defined incentives. Similarly, in automotive industries, autonomous agents are used in Tesla's self-driving technology to coordinate actions such as route planning, navigation, and vehicle maintenance. These systems, while designed for specific tasks, operate independently and interact with each other to optimize performance, making them prime candidates for the type of decentralized, incentive-driven models explored in this paper.

Beyond these industry-specific implementations, recent initiatives are also beginning to explore decentralized infrastructures for agentic collaboration. For example, the Masumi network proposes a blockchain-based protocol through which AI agents can establish verifiable identities, register their services, set prices, and transact with one another using tokenized payments. By requiring agents to log hashed outputs on-chain, the system provides accountability and transparency, while its decentralized identifier (DID) system helps ensure trust across heterogeneous platforms. Masumi's approach illustrates how an "AI agent economy" could be instantiated in practice: agents can discover one another, delegate tasks, and enforce contracts without centralized oversight (*Unlocking the AI Agent Economy* n.d.) This practical development complements our theoretical framework by showing how market-based coordination can scale to large populations of autonomous, self-learning AI agents.

Agents will interact with each other, react to incentives, establish contracts, and independently manage budgets. Just like participants in a self-regulated market, the agents will optimize their behavior based on rewards and penalties. Within logistics and supply chain sectors, firms like Amazon deploy AI agents to independently oversee inventory, streamline delivery routes, and forecast demand, all driven by market-based incentives. These real-world systems align with our model by adjusting behaviors to maximize efficiency and minimize cost, further illustrating the practicality of our framework. Just like the human economy, the price system will be an effective way to coordinate economic activity in a decentralized manner. In our model, agents respond to reinforcement signals, adjusting their behavior to maximize rewards and minimize penalties. Unlike traditional human teams, AI agents adapt through continuous learning, refining their decisions to align with both individual and collective objectives (Panait and Luke 2005). In such a setting, the system designer must define the structure of the contracts between agents. We build on the principal-agent framework from economics to examine the interactions between autonomous AI agents and the system designer, focusing on how their reward structures and team size can be optimized in decentralized, self-learning environments.

Imagine a group of AI agents who collectively allocate costly resources to accomplish a shared task. One agent might search for information, another process it, another critique it, another analyze it, and so on. We first consider a team of these agents operating in an exogenous system, where the system independently monitors the agents' performance. For example, a company like Sierra.AI can deploy multiple agents to a client to help solve a customer service problem, or a Tesla vehicle could deploy one agent to acquire information from the internet on places to visit, another agent to map out a route, and a third agent to narrate the journey. In these autonomous AI environments, we demonstrate that modifying team size is a more efficient response to environmental changes compared to altering incentive structures. This insight challenges traditional models, where incentives are the primary tool for adapting to risk or uncertainty, but in our model, the invariance property of incentives is more closely tied to the team size adjustment than incentive modifications. But if the principal can adjust the size of the number of agents simultaneously with their incentives, then all the variation in the optimal contract will come from the optimal team size, rather than from optimal incentives. This shows that adjusting the team size may prove more effective than altering financial incentives for the agents.

We extend the model by introducing a supervisory AI agent tasked with overseeing the performance of autonomous worker agents. This supervisor not only tracks performance but also adapts its strategies to the evolving behaviors of the agents, a crucial aspect of managing self-learning systems. This supervising agent can discover the output of the underlying agents at a cost. The more the supervisor exerts to monitor the team, the better signal it receives on their true performance. In this more general framework, we find that our invariance property of incentives still holds. Finally, we allow both the team of worker agents and the supervisory agent to have differing levels of skill. For example, agents may vary in their knowledge of a specific domain, their computational power, or their ability to reason, where better agents are more expensive. We show complementarity in quality, where matching the best supervisors with the best worker agents is efficient.

While our model offers valuable insights into optimizing incentives and team sizes for AI agents, it is important to acknowledge several limitations. First, the assumption of identical agents simplifies the model but may not capture the diversity of real-world systems. Additionally, the model's focus on monetary incentives and team size may not apply universally, particularly in settings where agent behaviors are influenced by non-financial factors. Moreover, this study focuses primarily on the economic dimensions of AI agents, leaving out significant considerations such as ethics and social impacts. These aspects warrant future exploration, which could expand our understanding and address the broader implications of deploying AI agents in real-world systems.

Overall, the study integrates classical economic theories with modern AI-based supervision strategies to offer a novel approach to managing AI agent teams. By emphasizing team size adjustments as a flexible tool for handling uncertainties, it provides new insights into MAS management. Future research could empirically test these findings in applications such as autonomous driving or smart manufacturing to validate their practical relevance. This work bridges the gap between traditional incentive design and the needs of evolving AI systems, providing a comprehensive foundation for theoretical and practical advancements in managing complex MAS environments.

In our framework, the system designer, who cannot directly observe the actions of individual agents, must rely on one or more supervisory agents to carry out performance monitoring. These supervisors are themselves self-interested and require proper incentives to engage in costly oversight. We contrast this setting with a baseline model of mechanical monitoring, which allows us to isolate the trade-offs involved in delegating oversight to autonomous agents. Specifically, we identify how supervisory monitoring can strengthen incentive alignment for worker agents, while also introducing a second layer of incentive frictions. Our model delivers closed-form solutions for the optimal number of agents and the structure of incentive contracts across both tiers. This in turn allows us to show how complementarities arise between supervisor and agent abilities—both when those abilities are fixed and when the principal selects from a talent pool with heterogeneous skill levels and reservation utilities.

The remainder of the paper is organized as follows. Section 2 lays out the basic model and derives its solution under both the benchmark (mechanical) and supervisory monitoring settings. Section 3 derives the existence of complementarities in the selection of supervisory and agent talent, in particular when external labor markets are present, and provides additional comparative statics results. The last section offers concluding comments and suggests ways in which the analysis in this paper could be extended or generalized.

1.1 Literature Review

This study builds on existing literature related to the coordination of autonomous agents in multi-agent systems (MAS), focusing on the challenges of optimizing interactions in decentralized environments. Horling and Lesser (2004) highlight the importance of dynamic team compositions that can adapt to environmental changes, demonstrating how organizational structures influence MAS efficiency. Rahwan et al. (2004) extends this by exploring interest-based negotiation, emphasizing its role in fostering cooperation among agents and reinforcing the need for adaptable team structures. Wooldridge and Jennings (1995) provide foundational theories of intelligent agents and their applications, linking theoretical models with real-world implementations.

The challenges of coordinating autonomous agents in decentralized environments have been well-examined in the literature. Jennings et al. (1998) provide a foundational overview of agent research, underscoring the evolution of MAS and the need for adaptable organizational structures. Their work highlights efficiency in agent coordination, essential for tackling the performance and scalability concerns addressed in this study. Agogino and Tumer (2010) demonstrate the potential of decentralized control strategies in managing air traffic flow, illustrating MAS's applicability in complex environments and motivating the need for refined models that incorporate adaptive incentive mechanisms. Additionally, Braun et al. (2023) examine human supervision designs in AI, emphasizing how robust oversight mechanisms help mitigate errors—an insight that directly informs the study's use of supervisory AI agents to oversee team performance.

Cooperative problem-solving within MAS plays a crucial role in facilitating collaboration among agents. Jennings (1995) introduces the joint intention framework to align agents' goals within industrial settings, laying the groundwork for investigating how generative AI agents can collaborate. A central theme of this study is the identification of an invariance property in MAS, where adjusting team size proves more effective than altering incentive structures when managing uncertainty. This finding challenges classical models, such as Baiman and Demski (1980), which suggest that incentives should vary with environmental conditions. Instead, we demonstrate that team size adjustments can absorb such variations, maintaining stable incentives. Lowe et al. (2017) further this discussion by presenting methods for training agents in mixed cooperative-competitive environments, offering a technical perspective on how MAS can be optimized.

Supporting the role of team composition, Vinyals et al. (2019) provide empirical evidence through their demonstration of multi-agent reinforcement learning in StarCraft II, showing how agent coordination can be achieved in complex environments. Ziv (1993) stresses the importance of tailored monitoring strategies for self-interested agents, while Stewart and Barrick (2000) emphasize that internal team processes can adapt to varying tasks, aligning with the strategic adjustments in team size proposed here. Building on this, Kok et al. (2006) propose a multi-agent reinforcement learning method based on payoff propagation, showcasing how agents can improve team performance through learned cooperation. Panait and Luke (2005) review cooperative learning techniques, underscoring the need for adaptive learning among agents to achieve optimized outcomes. Lastly, Foerster et al. (2018) introduce counterfactual multi-agent policy gradients, which enhance cooperative behavior by assigning credit based on counterfactual reasoning. These works illustrate the necessity of developing learning algorithms that foster effective teamwork, a key aspect in this study's exploration of how team composition influences MAS performance.

Achieving consensus and efficient coordination in MAS adds further complexity, which this study approaches from a novel perspective. Li and Tan (2019) investigate leader-follower consensus dynamics, introducing methods to achieve fixed-time convergence and synchrony in nonlinear systems—concepts that shape the study's coordination strategies. Similarly, Yokoo and Hirayama (2000) review distributed constraint satisfaction algorithms, emphasizing the value of efficient coordination under resource constraints.

Supervision and hierarchical structures are critical in MAS, ensuring effective oversight while balancing agent autonomy. Beyond these factors, agent diversity and the strategic pairing of roles are also emphasized in the literature as essential for enhancing team performance. Zarzà et al. (2023) examine how cooperation strategies among diverse agents, particularly those using large language models, can lead to better outcomes through collaboration. This aligns with the study's focus on leveraging team composition to address performance variability, prioritizing strategic role organization over adjustments to financial incentives. As AI systems evolve, findings on diversity underscore the need to balance various skills and responsibilities among agents to maximize performance. Additionally, Veale and Binns (2017) discuss how machine learning systems can be designed to mitigate discrimination without collecting sensitive data, highlighting the importance of understanding and accounting for machine behavior in system design.

Huhns and Stephens (1999) discuss agent organization within societies, highlighting how structured communication and clear roles facilitate collaboration. Building on Qian (1994) analysis of hierarchical supervision, the study extends these concepts to autonomous systems, showing how a supervisory AI agent can dynamically adjust its efforts to optimize team performance by reducing measurement noise. Amodei et al. (2016) contribute a modern perspective, addressing AI safety challenges and emphasizing the need for well-designed oversight mechanisms to ensure reliable AI performance. Additionally, Dorri et al. (2017) highlight the role of secure, scalable frameworks in decentralized settings like automotive security, emphasizing trust and privacy as essential components of successful agent interactions. Building on this, Lyu et al. (2023) analyze centralized critics in multi-agent reinforcement learning, revealing how centralized training frameworks can support decentralized agent behavior. Sandholm (2007) offers further insights into multi-agent learning, detailing the evolving strategies necessary for the effective management of agent interactions. Together, these studies underscore the intricate balance between supervision and autonomy, a principle central to this study's model of supervisory AI agents. Complementing these hierarchical perspectives, Tomašev et al. (2025) propose governance and market mechanisms tailored to autonomous agent societies. Their framework emphasizes sandbox economies and structured marketplaces for coordinating large-scale agent interactions, extending supervisory control ideas to macroeconomic structures and reinforcing the importance of formalized coordination frameworks in decentralized settings.

Ethics and governance are increasingly critical as AI systems gain autonomy, reinforcing the importance of developing responsible MAS. Zhang et al. (2021) survey machine learning researchers, uncovering major ethical concerns in AI governance, which shape this study's approach to MAS design. Shoham and Tennenholtz (1995) explore social laws for agent societies, proposing frameworks that foster order and cooperation among agents. Meanwhile, Weerdt et al. (2011) focus on task allocation within social networks, offering insights into effective responsibility distribution. This study builds upon these themes by proposing adaptive team structures that enhance MAS performance in complex environments, addressing gaps in existing research.

Together, these studies underscore the importance of flexible MAS frameworks that balance cooperation, supervision, and ethical considerations while addressing the complexities of agent interactions. Building on these foundations, our research focuses on optimizing the design of supervisory roles and incentive contracts in MAS to enhance performance under uncertainty. By integrating insights from both cooperative and competitive dynamics, as well as supervisory and ethical frameworks, this study advances our understanding of how MAS can be structured to manage inherent trade-offs in performance monitoring and coordination.

2 The Benchmark Model

Let's assume a principal (the system designer) contracts with n autonomous AI agents, where the designer's role is to define the environment in which these agents operate. The environment consists of exogenous factors like risk aversion, uncertainty, resource costs, and the efficiency of team coordination. These parameters are outside the agents' control but influence their decision-making and performance. In contrast, the endogenous variables, such as resource allocation and team size, are chosen by the system designer and represent the aspects of the system that can be manipulated to optimize agent performance. Each agent expends resources e_i at a cost given by $C_i(e_i) = \frac{1}{2}c_i e_i^2$, where c_i reflects the agent's efficiency. More efficient agents can exert higher effort at a lower cost. The formula $C_i(e_i) = \frac{1}{2}c_i e_i^2$ represents the quadratic cost function for agent $i's$ expenditure of resources e_i. The quadratic form ensures that costs increase at an increasing rate as the agent expends more resources, which is typical in many economic models where diminishing returns to effort are observed. Let $N = \{1, 2, ..., n\}$ denote the set of agents.

To simplify the analysis and ensure analytical tractability, we assume identical agents with the same cost structure, risk preferences, and behavior. This assumption allows us to derive closed-form solutions and analyze the relationship between team size and incentives. While this assumption provides valuable insights, it is a simplification of real-world systems, where agents may vary in terms of capabilities, preferences, and risk aversion. Future work could incorporate heterogeneous agents to better reflect real-world dynamics and explore how agent diversity impacts optimal team configurations and incentive structures. Additionally, the model focuses on monetary incentives as the primary driver of agent behavior. While other forms of incentives (e.g., social or moral rewards) could influence agent decisions, we focus on financial incentives to maintain clarity and tractability. This economic focus ensures a coherent framework, though other motivations could be included in more general models.

In traditional models, the principal may implement a mechanical monitoring system to generate a signal y that tracks agent performance. This signal serves as the basis for contract enforcement. However, with autonomous agents, performance monitoring and management are far more dynamic. In these scenarios, a supervisory AI agent is responsible for observing and interacting with the worker agents.

In the context of partial supervision, the supervisor's role is limited in its ability to monitor and influence the agents' actions. This is modeled by increasing the uncertainty term in the supervisor's measurement function. In hierarchical supervision, multiple levels of supervision are employed, each contributing to the overall system's uncertainty and measurement accuracy. These varying levels of supervision can be thought of as a multilayered oversight structure where each layer has different degrees of influence over agent behavior.

Unlike static systems, the supervisor in this model must adapt its monitoring strategies to the evolving behaviors of the agents as they learn and grow more self-sufficient. In autonomous vehicle fleets used by companies like Waymo, a supervisory AI agent oversees the actions of individual vehicles, ensuring that they follow traffic laws, adjust routes for safety, and coordinate with other vehicles to prevent accidents. This supervisory role is crucial for managing complex real-time interactions and ensuring that all agents work together efficiently and safely. The supervisor produces informative signals about agent effort, which is the only variable on which contracts can be written. We specify that:

$$y = \sum_{i \in N} e_i + \epsilon, \quad \text{where } \epsilon N(0, G(n)\sigma^2). \tag{1}$$

In AI systems, σ measures the uncertainty in performance measurement. We normalize $G(0)$ to equal 0, and assume that $G(n)$ is increasing and strictly convex in n. As the number of AI agents increases, measurement accuracy becomes more challenging due to the decentralized nature of the system and the self-learning behaviors of the agents. The performance of larger teams becomes harder to assess as agents adapt and optimize

independently, with issues related to coordination and communication complicating performance evaluation in multi-agent AI environments. In addition, we assume that $G(\cdot)$ satisfies the following technical condition:

$$\frac{d}{dn}\left(\frac{n[G'(n)]}{G(n)}\right) \geq 0. \tag{2}$$

This assumption holds for many common functional forms, making it broadly applicable in modeling noisy team-based performance. In particular, it holds for all power functions ($G(n) = An^\gamma$, $A > 0$, $\gamma > 1$), as well as any general polynomial function with positive coefficients. It is also satisfied by any function in the exponential class ($G(n) = A(e^n - 1)$, $A > 0$). Finally, note that the condition is implied by, and is therefore weaker than, a requirement that $G(n)$ be weakly log-convex.

Each agent is offered a linear contract by the principal, where the agent receives $w_i = a_i + b_i y$. All n agents are equally risk averse with exponential utility and a common coefficient of risk aversion parameter r. Risk aversion is a feature of utility functions from decision theory, initially motivated by human aversion to risk. Nonetheless, machines also may dislike variation in their payoffs for other reasons. For example, very low payoffs may deplete the agent's budget and lead it to power down. Of course, agents will have different parameters of risk aversion than humans, so there is no expectation that these machines are equally risk-averse as humans, just that there is some nonzero amount of risk aversion ($r > 0$). Thus, agents' preferences assume a mean-variance representation. In certainty equivalent terms, each agent receives $Ew_i - \frac{r}{2}Var(w_i) - C_i(e_i)$: expected wages minus a risk premium minus the cost of resource expenditure. Each agent solves

$$\max_{e_i} a_i + b_i Ey - \frac{r}{2}b_i^2 G(n)\sigma^2 - C_i(e_i), \tag{3}$$

yielding the standard incentive constraint, $e_i = b_i/c_i$. Notice that for each agent, the resource expenditure choice is also a dominant strategy response in the agents' subgame. Let each agent have an outside option, represented by $\bar{u}$, which we normalize to zero.[1] The principal will set the payments (a_i) such that each agent's individual rationality (IR) constraint binds, given that all other agents select their equilibrium resource expenditure levels.

Let q denote the exogenous price for which each unit of output is sold in a competitive market. The principal seeks to maximize the expected revenue after subtracting wage payments and incorporates the binding (IR) constraints into the optimization. So the principal selects bonus coefficients (incentives) to maximize total surplus:

$$\max_{b_i} \sum_{i \in N}\left[qe_i - C_i(e_i) - \frac{r}{2}b_i^2 G(n)\sigma^2\right]. \tag{4}$$

This objective function reflects two key costs that the principal must account for. The first is the personal cost of providing resource expenditure: $C(e_i)$. The second is the risk-premium: $\frac{r}{2}b_i^2 G(n)\sigma^2$. This is the incentive cost, which is the main focus of our study. After inserting the agent's incentive constraint, the first-order condition gives the optimal incentives for an exogenous n:

$$\hat{b}_i(n) = \frac{q}{1 + rc_i G(n)\sigma^2}, \tag{5}$$

yielding the standard risk/incentive trade-offs. Observe that as team size increases, measurement accuracy declines, prompting the principal to reduce incentive intensity in response to greater performance uncertainty. Also note that the equilibrium resource expenditure from each agent is $\hat{b}_i/c_i$, which remains below the idealized level of q/c_i that would arise under full information.

---

[1]We relax this assumption in Section 3 of the paper.

2.1 Incentives and Optimal Team Size

In a typical framework (e.g., a single-agent, single-task model in a one-period setting), the optimal incentive weight is given by $\hat{b} = 1/(1 + rc\sigma^2)$ and the sensitivity of incentives to risk, represented by $\partial\hat{b}/\partial\sigma$ reflects how incentives adjust with varying levels of risk or uncertainty. In the standard model, this derivative is negative, indicating a potential inverse relationship between risk and the incentive structure. However, empirical support for this trade-off is weak, as noted in the introduction. For instance, in AI-assisted retail management, AI agents are deployed to optimize product placements, pricing strategies, and inventory levels. The size of the AI team can vary depending on the complexity of the store layout or the range of products, while incentives such as sales targets or stock turnover rates drive the agents' decisions on where to place products or how to adjust prices. We show that by adjusting team size, the effect of risk and uncertainty on incentives can be moderated, which helps explain why the expected negative trade-off does not always hold in practice.

Consider a real-world application, such as a fleet of autonomous agents tasked with surveying a large area. Each agent works independently, but the overall mission success depends on the coordinated efforts of all agents. In this context, the incentive structure can be modeled similarly to the standard framework, where we assume all agents are identical (i.e., $c_i = c$) for simplicity. This simplification leads to $\hat{b}_i(n) = \hat{b}(n)$, allowing us to determine the optimal incentives for each agent and calculate the total profit based on the number of agents, n:

$$\hat{b}(n) = \frac{q}{1 + rcG(n)\sigma^2} \quad \text{and} \quad \hat{\Pi}(n) = \frac{nq\hat{b}(n)}{2c}. \tag{6}$$

The system designer adjusts the team size, denoted $\hat{n}$, to maximize $\hat{\Pi}(n)$. Given that $\hat{\Pi}$ and its derivative are continuous on $[0, \infty)$ and $\hat{\Pi}'(0) > 0$, a key condition for $\hat{n}$ to represent the optimal team size is that $\hat{\Pi}'(\hat{n}) = 0$. This leads to the first-order condition:

$$\hat{n}G'(\hat{n}) - G(\hat{n}) = \frac{1}{rc\sigma^2}. \tag{7}$$

Since the expression on the left is 0 at $n = 0$ and is strictly increasing in n (from the convexity of G), a solution, $\hat{n}$, exists and is unique. Furthermore, it is easy to verify that $\hat{\Pi}''(\hat{n}) < 0$, so that $\hat{n}$ is indeed the optimal team size.

Implicit differentiation in (7) yields our first key result:

PROPOSITION 1. *The optimal team size decreases as either agent risk aversion r, agent cost of resource expenditure c, or uncertainty in performance measurement σ^2 rises.*

In systems with autonomous agents, adjusting team size plays a key role in determining optimal incentives. We now analyze how incentive strength $\hat{b}$ responds to shifts in parameters such as uncertainty σ^2, agent risk aversion r, and agent cost of resource expenditure c, with a comparison between optimal team size adjustments and exogenously determined team sizes. Let us denote any of the symbols r, c, or σ^2 by ν. Suppose the current team size is n_0. If the team size were determined exogenously, the effect of changes in ν would be given by

$$\delta_{\text{ex}}^{\nu}(n_0) = \left.\frac{\partial\hat{b}}{\partial\nu}\right|_{n=n_0}, \tag{8}$$

where $\hat{b}$ is given by (6). If team size is chosen optimally, the effect of changes in ν is given by

$$\delta_{\text{end}}^{\nu}(n_0) = \left.\frac{\partial\hat{b}}{\partial\nu}\right|_{n=n_0} + \left.\frac{\partial\hat{b}}{\partial n}\right|_{n=n_0} \cdot \left.\frac{\partial\hat{n}}{\partial\nu}\right|_{\hat{n}=n_0}, \tag{9}$$

where $\partial \hat{n}/\partial \nu$ is obtained by implicit differentiation in (7). We show next that when team size is optimally adjusted in response to changes in external factors, the direction of the impact on incentive strength remains the same, though the magnitude is reduced.

PROPOSITION 2. *For any set of parameter values (σ^2, r, c), $\delta_{ex}^{\nu}(\hat{n}) < \delta_{end}^{\nu}(\hat{n}) \leq 0$.*

To rephrase, $|\delta_{end}^{\nu}(\hat{n})|$ is always smaller than $|\delta_{ex}^{\nu}(\hat{n})|$, meaning the sensitivity of incentives to risk, when team size is optimally adjusted, is always lower than when team size is fixed. Unlike the standard agency model, the system designer in this framework has the flexibility of two key levers: team size and incentives. Increasing risk aversion (r), the cost of resource expenditure (c), or measurement uncertainty (σ^2) leads to a direct reduction in $\hat{b}$ and an indirect increase in $\hat{b}$ through the decrease in team size ($\hat{n}$). As a result, the negative impact on incentives from changes in environmental factors is mitigated by the adjustment of team size. This illustrates how modifying team size in decentralized AI systems offers greater flexibility in managing incentives than simply adjusting them directly.

A notable exception to the typical risk-incentive trade-off occurs when $G(n)$ is a power function, $G(n) = An^{\gamma}$ (with $A > 0$ and $\gamma > 1$). In this situation, the sensitivity of incentives to performance is not merely reduced but entirely neutralized.

PROPOSITION 3. *The optimal incentive weight b does not depend on parameter values (σ^2, r, c) if and only if G is a power function, $G(n) = An^{\gamma}$.*

That is, we observe an *invariance property* when G is a power function: The optimal incentives remain constant despite changes in σ^2, r, or c. Rather than adjusting incentives directly, the system designer responds to these changes by dynamically adjusting team size, which keeps incentives stable across varying environments.

To clarify the underlying intuition, assume that n is initially fixed, and σ^2 increases. From equation (6), the principal reduces the incentive weight $\hat{b}$ because of the rising risk premium. As incentives decrease, so do profits. However, if the principal adjusts the team size, a reduction in n decreases the risk premium and thus increases the incentives. This rise in incentives compensates for the decline in profits due to the smaller team. When G is a power function, adjusting the team size exactly offsets the increase in uncertainty, maintaining stable incentives. This demonstrates that in decentralized systems with autonomous agents, team size adjustment is a more versatile tool for managing incentives than traditional modifications.

If $G(n)$ is a power function, the optimal team size can be derived as:

$$\hat{n} = \left(\frac{1}{(\gamma - 1)Arc\sigma^2} \right)^{\frac{1}{\gamma}}. \tag{10}$$

Substituting this expression for optimal team size into the equation for the agents' incentives yields the equilibrium strength of incentives at the optimal team size:

$$\hat{b}(\hat{n}) = q \cdot \left(\frac{\gamma - 1}{\gamma} \right) \text{ and } \hat{\Pi}(\hat{n}) = \frac{\hat{n}q^2(\gamma - 1)}{2c\gamma}. \tag{11}$$

From equation (11), the invariance property is readily observable: The optimal incentive weight, $\hat{b}(\hat{n})$, is independent of σ, r and c.

To conclude our analysis of the benchmark model, we turn to the robustness of our results. Specifically, we examine two key aspects of the model: the use of a collective performance measure for all agents and the assumption of an additively separable production function. We show that the key findings—such as the attenuation and invariance results—remain valid when these assumptions are relaxed. In particular, the model's predictions hold when agents are compensated based on individual performance signals rather than a shared measure y.

The results also extend to situations where production technology includes complementarities between agents, as opposed to the linear technology with uniform marginal productivity (q) in the benchmark model.

2.2 The Model with a Supervisor

We extend the framework by introducing a supervisory agent, which may be an AI system, tasked with managing the worker agents. For clarity, we refer to this overseeing entity as "the supervisor" and the agents under supervision as "worker agents." The supervisor's role is to actively and endogenously supply the necessary monitoring functions, differing from a static, mechanical monitoring system. This form of endogenous monitoring influences the trade-offs between team size and agent incentives, but we show that the attenuation result still holds when team size is chosen internally.

The supervisor is an economic agent who, though risk-averse, plays an important role in actively influencing the team's performance. Instead of solely monitoring, the supervisor collaborates with agents to improve strategies, enhance teamwork, and refine decision-making through ongoing feedback. As agents become more independent, the supervisor's role expands from mere observation to active guidance, ensuring that interventions are tailored to the agents' evolving needs, which guarantees optimal team function as the agents mature. The supervisor allocates resources, denoted by m, to enhance performance measurement. In AI-driven healthcare systems, where AI agents assist in diagnosing diseases and recommending treatments, supervisors oversee the interactions between agents. For example, in a hospital setting, the supervisor ensures that agents with different specializations—such as radiology AI and pathology AI—work together efficiently by adjusting their incentives (e.g., treatment success rates, time to diagnosis) and team size based on hospital needs and patient demands.

The cost of these supervisory efforts is captured by the function $C(m) = km$, where k reflects the supervisor's resource allocation efficiency. A supervisor with higher quality (or lower costs) has a smaller k, allowing them to exert more effort at a reduced cost. Therefore, the supervisor not only monitors agent performance but also adjusts strategies to ensure that agents' learning processes remain aligned with the broader team objectives. This evolving supervisory role is crucial as the agents' behaviors become more complex and autonomous over time. Specifically, we assume

$$y = \sum_{i \in N} e_i + \epsilon \text{ where } \epsilon N \left(0, \frac{G(n)\sigma^2}{m} \right). \tag{12}$$

The division of labor is such that the supervisor's investment in monitoring resources decreases the variance of y, while the agent's resource expenditure increases its mean. By this modeling choice, we choose to emphasize the monitoring role of the supervisor, leaving its other functions (such as fulfilling some direct productive tasks) as second-order effects. Within this monitoring role, the supervisor can, for example, exert resource expenditure to coordinate multiple agents within the team or facilitate communication between such agents. As a reduced form, we model this as directly reducing the variance on the team performance measure, as opposed to explicitly framing the exact consequence of the supervisory action, to keep the level of complexity manageable.

Assume m is unobservable, meaning the supervisor must be incentivized to allocate resources for monitoring. The system designer offers the supervisor a linear contract consisting of a fixed salary and a performance-based bonus tied to the team's output, represented as $\omega = \alpha + \beta y$. The supervisor thus solves:

$$\max_m \alpha + \beta \mathbb{E} y - \frac{r}{2} \beta^2 \frac{G(n)\sigma^2}{m} - C(m). \tag{13}$$

Although all agents and the supervisor are compensated under linear contracts, the methods for motivating resource expenditure differ. The supervisor's expenditure on monitoring does not directly affect the expected wage but serves to reduce the wage variance. Since the supervisor is risk-averse, there is a natural incentive to

minimize this variance. The first-order condition leads to the supervisor's incentive constraint:

$$m^* = \beta\sigma \left(\frac{G(n)r}{2k} \right)^{\frac{1}{2}}. \tag{14}$$

Even with linear personal costs, $C(m)$, the supervisor's optimization problem is concave, as the benefits from increased monitoring effort grow at a diminishing rate. The use of linear personal cost ensures tractability, while other functional forms of $C(m)$ would provide similar insights, albeit with more complex mathematical expressions.

Equation (14) shows that, for a fixed incentive coefficient β, an increase in measurement uncertainty (σ^2) or higher risk aversion (r) causes the supervisor to increase its monitoring efforts to counteract the higher risk premium. Moreover, as the supervisor's span of control grows (n increases), the supervisor must exert more effort to manage the decentralized AI team. Adding agents increases the variance of y and thus raises the risk premium. Since the supervisor aims to reduce performance uncertainty, it increases its monitoring efforts to compensate for the additional risk cost. However, the supervisor's incentive coefficient β is endogenous, depending on the parameters σ and r. We return to address the overall effect on equilibrium monitoring resource expenditures later (in Proposition 5).

As before, the system designer sets the payment levels (a_i, α) such that the individual rationality constraint binds for every agent and the supervisor. The opportunity wage for the supervisor is $\bar{u}_S$, normalized to zero. The system designer maximizes total surplus, as in the previous sections.

$$\max_{b_i, \beta} \sum_{i \in N} \left[qe_i - C_i(e_i) - \frac{r}{2}Var(w_i) \right] - C(m) - \frac{r}{2}Var(\omega), \tag{15}$$

where

$$Var(w_i) = b_i^2 \frac{G(n)\sigma^2}{m} \text{ and } Var(\omega) = \beta^2 \frac{G(n)\sigma^2}{m}. \tag{16}$$

Substituting in each agent's incentive constraint gives

$$\max_{\beta, b_i} \sum_{i \in n} \left[\frac{qb_i}{c_i} - \frac{b_i^2}{2c_i} - \frac{r}{2}b_i^2 \frac{G(n)\sigma^2}{m^*} \right] - km^* - \frac{r}{2}\beta^2 \frac{G(n)\sigma^2}{m^*}, \tag{17}$$

where m^* is given by (14). The first-order conditions yield the optimal incentives for the agent:

$$b_i^*(n) = \frac{q}{1 + \dfrac{rc_i G(n)\sigma^2}{m^*}}.$$

Similar to the earlier setting with a mechanical monitor, an increase in risk aversion (r), cost of resource expenditure (c_i), or noise (σ^2) results in the system designer reducing the agents' incentives, which reflects the classic trade-off between effort and reward. The key distinction in this model is that the supervisory AI agent's resource expenditure (m) plays a crucial role in determining the optimal incentives for the autonomous agents. The system designer adjusts m by setting the supervisor's incentive β in response to the agents' performance.

Optimizing the principal's objective function with respect to β results in a first-order condition, from which we can see the principal's trade-off in choosing the supervisor's incentive:

$$\left[\frac{r}{2} \frac{G(n)\sigma^2}{m^{*2}} \sum_{i=1}^n b_i^2 - k + \frac{r}{2} \frac{G(n)\sigma^2}{m^{*2}}\beta^2 \right] \frac{\partial}{\partial \beta}m^* - r\beta \frac{G(n)\sigma^2}{m^*} = 0. \tag{18}$$

Increasing supervisory incentives (i.e., increasing β) leads to a greater resource expenditure by the supervisory AI agent (see equation (14). This has two main benefits. First, it decreases the variance of the performance measure y, reducing the supervisor's risk premium since their compensation is tied to y. Second, increasing β

also reduces the needed risk-premium of all autonomous agents, because the agent' contracts are also tied to y. The system designer benefits from a spill-over effect resulting from the supervisory AI agent's monitoring efforts. The supervisor, being self-interested, aims to reduce the variance in its own compensation, which, in turn, reduces the wage variance of all autonomous agents. This results in lower overall compensation costs for the system. This is beneficial to the system designer because it reduces salary costs (a_i^* and α^*). These two (marginal) benefits are shown as the two positive terms in (18). The cost of providing supervisory incentives is two-fold. First, a higher incentive increases the supervisory AI agent's resource expenditure, requiring the supervisor to be compensated. Second, a higher incentive makes the supervisor's compensation more sensitive to signal variations, which raises the risk premium. These two (marginal) costs are shown as the two negative terms in (18).

Consider an extension of the model where the supervisory AI agent exerts a productive resource expenditure, denoted by e_m, which, similar to agents' resource expenditures e_i, increases the output and the aggregate performance measure y. When choosing the optimal β, the system designer now must consider the marginal benefit and cost of inducing e_m, in addition to the four effects considered in equation (18). However, so long as these additional marginal effects are well-behaved (i.e., smooth and bounded), their only impact is to make the optimal β quantitatively different. The fundamental trade-off in the supervisory monitoring activities, which is the focus of our paper, persists and leads to the same qualitative results.

Substituting the supervisor's incentive constraint (14) into (18), and rearranging terms yields

$$\beta = \left(\frac{1}{2} \sum_{i \in N} b_i^2 \right)^{\frac{1}{2}}. \tag{19}$$

The supervisory AI agent's incentives increase with the incentives of each autonomous agent. Intuitively, as the system designer increases b_i, it induces greater effort from the agents but also increases the risk premium. To offset this, the system designer increases β, which leads the supervisory AI agent to allocate more resources to monitoring the agents. This, in turn, reduces the risk premium for all agents and the supervisor.

By combining the previous expressions for b_i, β, and the supervisory agent's incentive constraint, we derive the following expression for b_i^* as an implicit function of the model parameters:

$$b_i^*(n) = \frac{q}{1 + 2c_i \sqrt{\dfrac{rG(n)\sigma^2 k}{\sum_{j \in N} b_j^{*2}}}}. \tag{20}$$

Despite the fact that equation (20) is an implicit function of b_i, it can be shown that, as in the setting with a mechanical monitor, the strength of incentives (b_i) falls as the cost of resource expenditure (c_i) increases.

2.3 AI Specificity of the Model

While the modeling framework builds on classical principal–agent theory, several features make its application to teams of AI agents distinct from human or pre-LLM multi-agent systems. First, the agents in our setting are self-learning and decentralized, adapting their strategies continuously rather than following fixed rules or human behavioral heuristics. This changes the locus of control: Instead of relying on hierarchical direction, coordination must emerge through economic signals such as rewards and penalties.

Second, the introduction of risk aversion has a different interpretation than in human teams. In traditional models, risk aversion reflects psychological preferences over uncertain payoffs. For AI agents, by contrast, risk aversion proxies for safety preferences or operational constraints. A low payoff might deplete an agent's computational budget or shut down its processes. Thus, variation in the parameter r captures a system-level trade-off between innovation and safety, rather than human psychology.

Third, because agents are modeled as autonomous and resource-constrained, the system designer must consider both the number of agents and the structure of their contracts. Our results show that environmental changes are more efficiently absorbed by adjusting team size rather than incentive strength. This invariance property is distinctive to decentralized AI teams, where scaling the number of cooperating agents can substitute for re-optimizing complex incentive schemes.

Finally, the inclusion of a supervisory agent reflects the technological possibility of AI-based monitoring, rather than human oversight. Unlike a human supervisor, such an agent incurs computational costs to extract more precise signals of worker performance. This allows us to model a second layer of incentive frictions specific to AI systems: Both supervisors and workers must be incentivized, and their abilities interact complementarily in equilibrium.

These features illustrate why our analysis departs from a simple extension of classical team theory. The economic framework remains the same, but the interpretation and implications differ in ways specific to autonomous AI agents, particularly in how safety, scalability, and decentralized learning reshape the trade-offs faced by the system designer.

2.4 AI Safety

When designing multi-agent systems, AI safety becomes a critical consideration—specifically, the steps required to prevent harmful outcomes in systems involving autonomous or semi-autonomous AI agents. In autonomous drones used for environmental monitoring, AI safety measures are critical to ensure that the drones do not malfunction or cause damage to ecosystems. Supervisory agents monitor the drones' movements, ensuring that they avoid restricted areas or potentially hazardous zones. These safety preferences might influence the team size (e.g., fewer drones are deployed in sensitive areas) and the incentive structure (e.g., bonuses are awarded for precise data collection and minimal environmental disruption). A preference for safety in these systems influences the equilibrium outcomes in several key areas, such as incentives, team size, resource allocation, and profits. These preferences, while essential to safeguard society from potential harm, often introduce trade-offs with innovation and system efficiency.

System designers must carefully balance these trade-offs to prevent social damage from occurring within an autonomous community of agents. For example, ensuring that AI systems operate within safe and ethical boundaries may require adjusting team sizes or lowering incentives to prevent harmful behaviors from emerging in larger, more independent agent communities. These adjustments help mitigate risks like exploitation or unsafe actions that could otherwise arise in decentralized settings.

However, some results are counterintuitive. While safety measures can limit the system's potential efficiency, they are necessary for ensuring that AI agents' behavior aligns with human values and avoids dangerous outcomes. Introducing safety preferences may lead to different resource allocation strategies, such as hiring higher-quality supervisors, and potentially increasing the cost of supervision to ensure alignment. These shifts underscore the complex relationship between AI safety and innovation, where more safety may reduce the system's overall performance, but at the cost of societal well-being.

This subsection highlights the trade-offs between AI safety and system performance and provides insight into how system designers can implement safeguards to prevent social harm. These discussions are essential as we move forward in developing autonomous systems that are not only effective but also aligned with ethical considerations. While this analysis remains focused on the economic aspects of agent behavior, future work could integrate more detailed safety and ethical constraints, expanding the model to address these critical concerns in real-world applications.

2.5 Heterogeneous Agents

All the analysis so far has assumed homogeneous agents. This assumption is primarily for simplicity, as it allows for expressing the principal's profit in terms of the univariate scalar team size (n), which can then be optimized by the principal. Of course, in practice, agents are heterogeneous and therefore will have different marginal cost terms, $c_i \neq c_j$. While we will still lose some of the more novel results like the invariance property, the intuition behind the other results should still go through.

To see this, suppose that agents are heterogeneous, so $c_i \neq c_j$ for $i \neq j$. The high-cost agents will have lower optimal bonuses under both mechanical monitoring and with a supervisor, given by equations 5 and 20. Therefore, in both cases, greater costs result in lower incentives. The effects from the other exogenous parameters, namely risk aversion, noise σ^2, and the supervisor's cost k, will have the same directional impact on incentives. Increases in any of those exogenous parameters will lead to lower incentives.

This is the same effect as what happens when agents are homogeneous. As such, we would expect the optimal team size under heterogeneous agents to have the same directional effect with respect to changes in the exogenous parameters as it does under homogeneous agents. Because the optimal incentives behave the same, it is plausible that the optimal team size should behave the same, so we conjecture that increases in risk aversion, noise, or supervisory cost will all lead to decreases in the optimal team size.[2] It is likely that the directional effects should be similar since heterogeneity among the agents does not actually change the impact of the exogenous parameters on incentives.

Finally, observe that higher-cost agents will have a larger magnitude change in their incentives with respect to the exogenous parameters. Formally, we know that the first derivative $\frac{\partial b_i}{\partial \theta} < 0$ for $\theta = r, \sigma^2, k$. However, high-cost agents with a high c_i will have a larger impact from a change in the exogenous parameters and, therefore, the cross partial term $\left| \frac{\partial^2 b_i}{\partial c_i \partial \theta} \right| > 0$. Said differently, an increase in risk aversion or noise will always decrease incentives, but will do so even more for the high-cost agents. We conjecture that a team with more high-cost agents will have even lower incentives and, therefore, even smaller team sizes than a team with fewer high-cost agents. Ultimately, the principal's optimization of team size relies on how sensitive the optimal incentives are with respect to the agent's environment. The greater the sensitivity, which occurs for the highest-cost agents, the greater the effect on optimal team size.

2.6 Optimal Incentives and Team Size with a Supervisor

When agents are identical ($c_i = c$), we simplify the equation to $b_i^* = b^*$, enabling us to solve (20) and derive a closed-form solution for the equilibrium agent incentive for a given n:

$$b^*(n) = q - 2c\sigma\sqrt{rkG(n)/n}. \tag{21}$$

Recall that the equilibrium supervisory agent incentive (β^*), monitoring resource expenditure (m^*), and agent resource expenditure (e^*) are all functions of b^*. The maintained assumption here is that $b^* > 0$. We exclude parameter combinations that would yield a negative b^*, which would imply that autonomous agents receive zero incentives, effectively halting their performance.

---

[2]Proving this formally is a challenge because we do not have a formal expression for the optimal team size in closed form when agents are heterogeneous. It is unlikely that invariance will hold exactly, though it might hold in an approximation.

From (17), we can now write the profit function with a supervisor for a given n, denoted by $\Pi(n)$, as follows:

$$
\begin{aligned}
\Pi(n) &= n\left[\frac{qb^*}{c} - \frac{b^{*2}}{2c} - \frac{r}{2}b^{*2}\frac{G(n)\sigma^2}{\frac{b^*\sigma}{2}\sqrt{\frac{rnG(n)}{k}}}\right] - 2k\frac{b^*\sigma}{2}\sqrt{\frac{rnG(n)}{k}} \\
&= \frac{n\left(b^*(n)\right)^2}{2c}.
\end{aligned}
\tag{22}
$$

Now we characterize the optimal n in this setting.

LEMMA 1. *Assume $c_i = c$ for all $i \in N$, and the principal hires a supervisor with quality k. There exists a unique optimal team size, n^*, which is characterized by the following implicit function:*

$$
h(n^*) + 2n^*h'(n^*) = \frac{q}{2c\sigma\sqrt{rk}},
\tag{23}
$$

where $h(n) = \sqrt{G(n)/n}$.

Implicit differentiation of the equation defining n^* above enables us to study how the principal adjusts team size in response to changes in the exogenous parameters of interest.

PROPOSITION 4. *The optimal team size is decreasing in measurement uncertainty σ, agent risk aversion r, and agent and supervisor costs of resource expenditure, c and k.*

When either c, k, r, or σ increases, the risk premium rises. As in the case of mechanical monitors, the system designer responds by adjusting team size, thereby mitigating the increase in the risk premium. As the quality of the supervisor improves (k shrinks), the optimal team size increases for two key reasons. The first reason is the reverse effect of the risk premium mentioned earlier. The second reason is that better supervisors are more effective at measuring performance, allowing them to manage larger teams.

Next, we explore whether the ability to adjust team size influences the strength of incentives, similar to the mechanical monitor setting. Let ν denote any of the parameters σ, c, $\sqrt{r}$, or $\sqrt{k}$. The impact on the agents' incentive parameter, b, of factor ν is then given by

$$
\phi_{\text{ex}}^{\nu}(n_0) = \left.\frac{\partial b^*}{\partial \nu}\right|_{n=n_0}
\tag{24}
$$

and

$$
\phi_{\text{end}}^{\nu}(n_0) = \left.\frac{\partial b^*}{\partial \nu}\right|_{n=n_0} + \left.\frac{\partial b^*}{\partial n}\right|_{n=n_0} \cdot \left.\frac{\partial n^*}{\partial \nu}\right|_{n^*=n_0},
\tag{25}
$$

for the exogenous and the endogenous team-size cases, respectively. Since the monitoring is now performed by a self-interested supervisor, an immediate question is whether a similar comparison can be performed on the strength of the incentive, β, offered to the supervisor. We refer to the corresponding constructs for the supervisor as $\psi_{\text{ex}}^{\nu}(n_0)$ and $\psi_{\text{end}}^{\nu}(n_0)$, respectively. Finally, it is also of interest to identify the relative impact of parameter changes on the supervisor's and agents' equilibrium levels of reward-for-performance. The following result provides a full characterization of these issues.

PROPOSITION 5. *For any set of parameter values (σ^2, r, c, k), and for a change in any factor $\nu = \sigma, c, \sqrt{r}$, or $\sqrt{k}$:*

$$
\psi_{ex}^{\nu}(n^*) = \psi_{end}^{\nu}(n^*) \le \phi_{ex}^{\nu}(n^*) < \phi_{end}^{\nu}(n^*) \le 0,
\tag{26}
$$

with a strict inequality in the middle for any $n^* > 2$.

The implications of the above inequalities are as follows. First, as in the benchmark monitor setting, $|\phi_{\text{end}}^\nu(n^*)| < |\phi_{\text{ex}}^\nu(n^*)|$, i.e., introducing team size as a variable reduces the impact of incentives. An increase in c, σ, r, or k results in a decrease in the strength of incentives offered to the autonomous agents, regardless of whether team size is exogenously given or optimally adjusted. In the latter case, however, changes in any of these exogenous factors result in a smaller adjustment to the incentive strength b for autonomous agents, meaning that the effect of environmental changes on the optimal contract is more muted when team size is optimally chosen.

Second, we find that no such attenuation effect occurs for the supervisory agent's incentives. Whether or not the team size is adjusted optimally, the variation in the supervisor's incentive remains unaffected by environmental changes. Moreover, the decrease in the supervisor's incentive coefficient for an increase in any factor, ν, is greater than the corresponding drop in the agents' incentive coefficient. This suggests that, while both agents and supervisors are risk averse, supervisors' incentives are more responsive to changes in exogenous factors (uncertainty, risk aversion, or costs of resource expenditure) than are the incentives of the agents. Our model thus yields the testable prediction that one should empirically observe greater variation in the incentives for supervisors than those for production agents.

The intuition for the result is as follows. From the previous section, we know that $\beta^* = b^*\sqrt{n/2}$. Therefore, parameters affect the supervisory AI agent's incentives in two ways: through their impact on b^* and through their influence on the optimal team size. The effect of the parameters on β^* is stronger than that on b^* whenever the effects on b^* and n work in the same direction. Proposition 4 indicates that an increase in any of the parameters r, c, k, or σ causes the principal to reduce team size. Thus, whenever the principal reduces b^* in response to these parameter changes, the effects of the parameters on b^* and n amplify each other; as a result, β^* is more responsive than b^*.

We now show that the extreme version of invariance, where the autonomous agents' incentives remain unaffected, continues to hold for power functions, even when a supervisory AI agent is introduced into the system.

PROPOSITION 6. *The optimal b does not depend on parameter values (σ, r, c, and k) if and only if G is a power function, $G(n) = An^\gamma$.*

A different way of stating this result is as follows: the weak inequality, $\phi_{\text{end}}^\nu(n^*) \leq 0$, in the statement of Proposition 5 is a strict one unless $G(n)$ belongs to the class of power functions. At an intuitive level, we know from Proposition 3 that the invariance property holds in this new setting if and only if $G(n)/m^*$ is a power function. But, by (14), $m^*(n)$ is a power function if and only if $G(n)$ is a power function. Consequently, $G(n)/m^*$ is a power function if and only if $G(n)$ is.

Propositions 5 and 6 together imply that if G is a power function, the supervisor's incentives respond to parameter changes, while the incentives of the agents do not. To see this explicitly, consider the following closed-form solutions for the equilibrium levels of the agents' and supervisor's incentives when team size is chosen optimally:

$$b^*(n^*) = q \cdot \left(\frac{\gamma - 1}{\gamma}\right) \tag{27}$$

and

$$\beta^*(n^*) = \frac{q}{\sqrt{2}} \left(\frac{\gamma - 1}{\gamma}\right) \left(\frac{q}{2\gamma c\sigma\sqrt{Ark}}\right)^{\frac{1}{\gamma-1}}. \tag{28}$$

Recall that, in general, $\beta^* = b^*\sqrt{n/2}$, and n is decreasing in σ, r, c, and k. Since b^* remains unaffected by exogenous parameters when G is a power function, the combined effect is that β^* decreases in all of the parameters. The result may seem counterintuitive, as one might expect the system designer to raise incentives for the supervisory AI agent to monitor the autonomous agents more precisely if the measurement process worsens (σ rises). Instead, the system designer reduces team size to counter the increase in risk premium. Smaller teams require

less supervision, and thus the system designer can afford to reduce the incentives for the supervisory AI agent. The system designer once again uses team size as the primary tool for adjusting to exogenous changes in the environment, rather than increasing incentives for the supervisor.

2.7 Value of the Supervisor

To close this section, we examine the conditions under which it becomes advantageous for the principal to hire a supervisor, particularly in the context of autonomous AI systems. In smart cities, where AI agents are responsible for managing traffic lights, monitoring waste management, and controlling energy usage, a supervisory AI agent ensures all systems are operating efficiently. The supervisor's value is evident in how well it manages resource allocation (e.g., adjusting energy consumption based on peak hours) and ensures that AI agents don't over- or under-perform, reducing the overall cost of operations and increasing the system's efficiency. As AI companies consider deploying self-learning agents, the principal must decide whether to use a fixed, non-AI monitoring system or to employ an adaptive AI agent as a supervisor. This decision is pivotal, as the supervisor's role is no longer just about tracking performance—it extends to guiding agent learning and adapting to the evolving behavior of the agents as they continuously learn and adapt to new environments.

Assume the system designer's scale is fixed, and that all agents are identical with $c_i = c$ for all $i \in N$. The system designer must choose between a fixed mechanical monitoring system and a dynamic AI supervisor with a known quality level, k, who can adjust its strategies as the agents become more independent. Given the parameters (q, r, c, σ^2), the system designer compares the expected profit from both alternatives. The value of hiring a supervisor is given by:

$$
\begin{aligned}
\Pi(n) - \hat{\Pi}(n) &= \frac{n}{2c} \left[\left(q - 2c\sigma\sqrt{rkG(n)/n} \right)^2 - q\frac{q}{1 + rcG(n)\sigma^2} \right] \\
&= \frac{\sigma}{2}\sqrt{rG(n)} \left[q\sqrt{n} \left[\frac{q\sigma\sqrt{rnG(n)}}{1 + rcG(n)\sigma^2} - 4\sqrt{k} \right] + 4c\sigma\sqrt{rkG(n)} \right].
\end{aligned}
\tag{29}
$$

It follows immediately that a sufficient condition for the owner to find it in its interest to hire a supervisor with quality k is:

$$
\frac{q\sigma\sqrt{rnG(n)}}{1 + rcG(n)\sigma^2} \geq 4\sqrt{k}.
\tag{30}
$$

As expected, the supervisory agent's value increases as its quality improves: Lower values of k reduce the difficulty of motivating a supervisor who is both resource- and risk-averse. Furthermore, the expression in (30) shows that employing a supervisory AI agent becomes undeniably beneficial when the market value of output, q, is high. More notably, an examination of the profit differential reveals that hiring a supervisor is particularly advantageous when the autonomous agents are of higher quality (i.e., c is lower). We explore further the complementarity between autonomous agents and supervisory AI agents in the next section.

The relationship between uncertainty in the monitoring technology and supervisor value is more complex. In particular, it can be shown that the supervisor's value diminishes when σ^2 is either too low (as a mechanical monitor can suffice) or too high (due to the increased cost of motivating the supervisory agent and compensating for a risk premium in that case). However, for all intermediate values of the uncertainty parameter, σ^2, the system designer will find it beneficial to hire a supervisory agent. This suggests that if the precision of the performance measurement increases (i.e., σ decreases) over the life cycle of a system designer, the system designer may initially operate without a supervisory agent, hire one when σ becomes sufficiently manageable, and eventually switch to a mechanical monitor when the performance measurement system is refined enough.

When the size of the system designer is endogenous, for a given set of parameters, the system designer compares the expected profits under both a mechanical monitor at the optimal team size and a hired supervisory AI

agent at a possibly different team size. Although the economic intuition remains the same, a general formula for the profit differential cannot be derived due to the implicit nature of the optimal team size functions. We can, however, gain further insights by examining the case where $G(n) = An^\gamma$. In this setting, the value of hiring a supervisor is

$$\Pi(n^*) - \hat{\Pi}(\hat{n}) = \frac{q^2(\gamma-1)}{2c\gamma} \cdot \left(\frac{n^*(\gamma-1)}{\gamma} - \hat{n} \right). \tag{31}$$

Given that $\gamma > 1$, it follows that hiring a supervisor is beneficial only when the optimal team size with a supervisor is larger than the optimal team size without one. Supervisors contribute by enhancing team performance measurement, which in turn enables the formation of larger teams. This relationship between the supervisor's value and team size reveals a synergy between the two, with the size of the team influencing the supervisor's effectiveness. Notably, this synergy (or complementarity) arises not from the production technology itself, but from the inherent moral hazard issues in decentralized team-based production.

In summary, this model presents the principal with three key design tools to optimize expected profit. They are: (1) the number of agents (n), (2) the decision to hire a supervisor, and (3) the determination of incentive coefficients for both the agents and the supervisor (b and β). As the environment shifts—represented by changes in parameters such as (q, r, σ, c, or k)—all three of these variables can be adjusted accordingly. The distinctive feature of this model is that autonomous agents' incentives show robustness to environmental changes. In particular, b^* responds minimally, or not at all, to factors other than the market price of output q, compared to when team size is fixed. In contrast, the system designer reacts to changes in all other parameters by adjusting team size and monitoring levels, with less emphasis on altering the agents' incentives.

This section explores complementarity, in the sense of optimal matching, between supervisors and agents. That is, should better supervisors be asked to supervise better agents or worse ones? We show that it is always optimal to do the former, i.e., supervisor and agent abilities are complements. We then explore how the selection of the optimal supervisor and agent combination in a labor market varies with uncertainty and other parameters of our model.

3   Simulations

In this section, we will restrict our attention to the case when G is a power function, $G(n) = An^\gamma$, with $A > 0$, $\gamma > 1$. In this case, the closed-form solutions for the agents' and supervisor's incentive contracts are given in equations (27) and (28). The solutions for the other variables of interest are provided below:

$$n^* = \left(\frac{q}{2\gamma c\sigma\sqrt{Ark}} \right)^{\frac{2}{\gamma-1}}; \tag{32}$$

$$m^* = \frac{q\sigma}{2} \left(\frac{\gamma-1}{\gamma} \right) \sqrt{\frac{Ar}{k}} \left(\frac{q}{2\gamma c\sigma\sqrt{Ark}} \right)^{\frac{\gamma+1}{\gamma-1}}; \tag{33}$$

$$\Pi^* = \frac{n^*(b^*)^2}{2c} = \frac{q^2}{2c} \left(\frac{\gamma-1}{\gamma} \right)^2 \left(\frac{q}{2\gamma c\sigma\sqrt{Ark}} \right)^{\frac{2}{\gamma-1}}. \tag{34}$$

Now that we have closed-form solutions for the optimal incentive contracts, team size, monitoring effort, and profits, we can graph these in numerical simulations to generate intuition on their shapes and directions.

The optimal bonus for the supervising agent is quite simple since it varies only with respect to γ and not with respect to any other parameter. This is fundamentally the invariance property discussed in the prior section. The incentive contract for each worker agent is more complex, which varies with each of the parameters of the
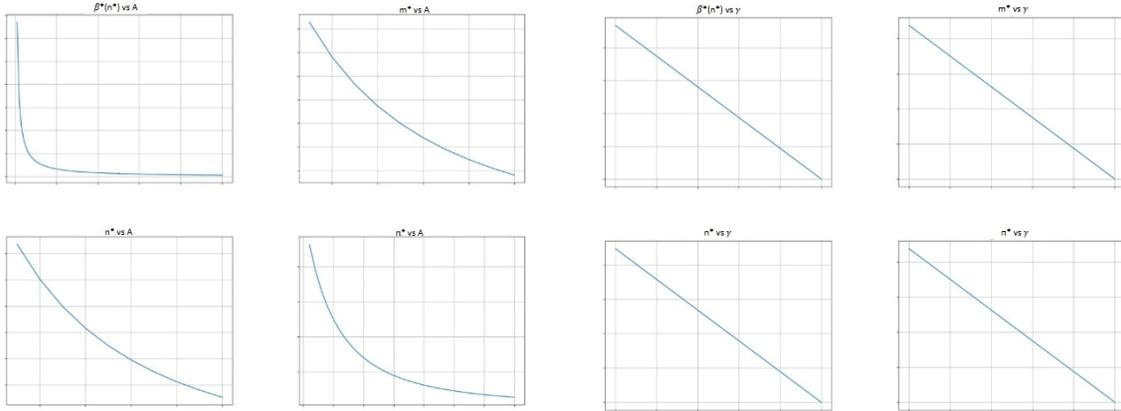
Fig. 1a. Equilibrium Changes
in Respect to A

Fig. 1b. Equilibrium Changes
in Respect to γ

model. These parameters are γ, c, σ, A, r, and k. We now provide some intuition and examples for when and how these parameters should vary.

With closed-form solutions, we can examine how the equilibrium of the model changes with the exogenous parameters. The equilibrium here is the equilibrium contract b^* and B^*, as well as the equilibrium action choices of the agents. These actions include the monitoring effort of the supervisory agent (m^*) and the choice of the number of worker agents (n^*). Finally, we can also examine what happens to equilibrium profits. We will interpret each of these exogenous parameters and show via simulations how they will change the equilibrium, in several figures.

First, consider the parameters A and γ. These are the parameters for $G(n) = [An]^\gamma$, which governs the efficiency of the supervisory agent's monitoring technology. Any decrease in the function $G(n)$ will decrease the variance in team output, y. We interpret changes in A and γ as improvements in changes in the software stack used by the supervisory agent to manage the worker AI agents. Decreases in either A or γ will decrease the G function, which will decrease the variance in team output, leading to better monitoring of the AI agents' output. The parameter A scales the level of G, whereas a change in the parameter γ affects more directly the impact on the size of the team. For example, general advances in LLM software will correspond to decreases in A, while more specific improvements in reasoning and/or agent coordination can correspond to decreases in γ.

Figure 1a shows how the equilibrium changes with respect to A, and Figure 1b shows how it changes with respect to γ. In each case, decreasing A or γ will lead to stronger incentives for the supervisor, more effort and resources from the supervisor, larger teams, and greater profits. This happens because decreasing γ or A reduces the variance in team output, which allows the principal to offer stronger incentives, which in turn induces more effort/resources out of the supervisory agent, which then induces the worker agents to invest more effort/resources, which ultimately increases firm profits.

Now consider changes in the parameters c or k, which govern the cost functions for the worker agent and supervisory agent. We interpret these parameters as the hardware resource costs for the worker and the supervisor. Smaller levels of c or k lead to a smaller marginal resource cost for the worker agent and the supervisory agent, since these parameters track the marginal resource cost since $C'(e) = ce$ and $C'(k) = km$. For example,
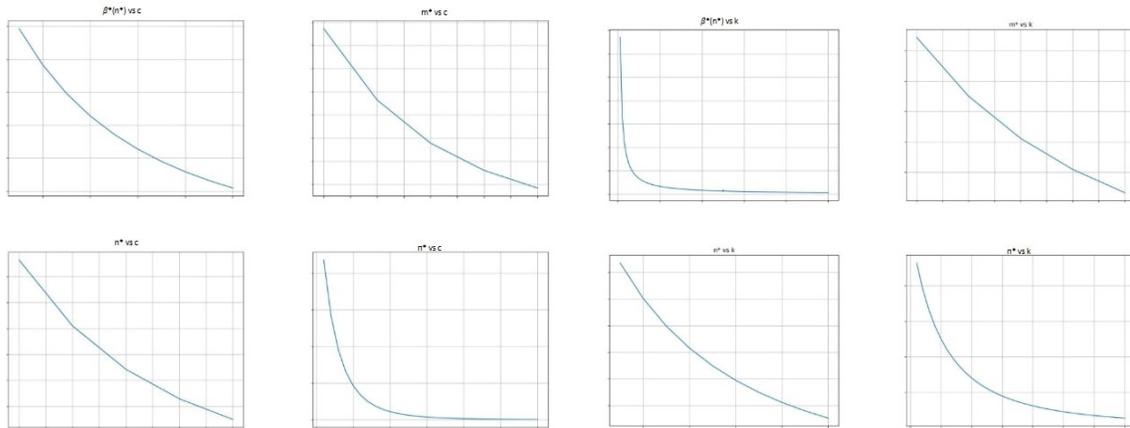
Fig. 2a. Equilibrium Changes
in Respect to c

Fig. 2b. Equilibrium Changes
in Respect to k

advances in accelerated computing will lead to a smaller resource cost, or, said differently, a smaller cost per unit of compute provided to the AI agent.

Figures 2 shows how the equilibrium changes with respect to changes in these marginal resource costs c and k. As the resource costs decrease, the optimal incentive for the supervisor increases, as does the optimal monitoring effort, the optimal team size, and firm profits. Decreasing the cost of compute allows the supervisor to monitor more carefully since every unit of supervisory work is cheaper, which in turn allows the principal to offer stronger incentives to the supervisory agent, which leads the workers to work more, and allows for larger teams and greater revenue and profits. This should be a somewhat natural consequence of lowering the cost of compute, which follows the general trend from companies like NVIDIA and AMD.

Figure 3a shows how the equilibrium changes under an environment of greater uncertainty. In all cases, greater uncertainty leads to weaker incentives for the supervisor, less supervisory effort, smaller teams, and lower profits. In a highly uncertain environment, both the AI worker and supervisor AI agents have a less consistent relationship between their input effort and their reward, which rationally leads them to reduce their investments of effort and resources, which leads to lower profits and smaller teams. For example, periods of high competition between AI agent companies (like OpenAI versus Anthropic) can lead to greater levels of uncertainty. Similarly, foreign competition from AI rivals (like DeepSeek) can also erode the relationship between inputs and outputs, leading to greater uncertainty.

Finally, consider changes in r, the risk aversion of the AI agents. While risk aversion traditionally has been seen as a psychological feature of humans in how they respond to financial gambles, in this context it can proxy for the preference for safety of the AI system. Different AI agent platforms in the market all trade-off safety and innovation at different levels.[3] A high value of risk aversion is a high level of preference for safety. Figure 3b shows how the equilibrium changes with these different preferences for safety. As a preference for safety grows, the optimal incentive for the supervisor shrinks, as does the optimal team size, supervisory monitoring effort, and firm profits. While it may be intuitive that increasing a preference for safety leads to lower profits, as this is

---

[3]For example, the public tension surrounding OpenAI profiled in the news generally characterizes Sam Altman as prioritizing innovation over safety, and others like Ilya Sutskever as prioritizing safety over innovation, on the margin. This is based on Ilya Sutskever leaving OpenAI to found a new AI startup called Safe Superintelligence.
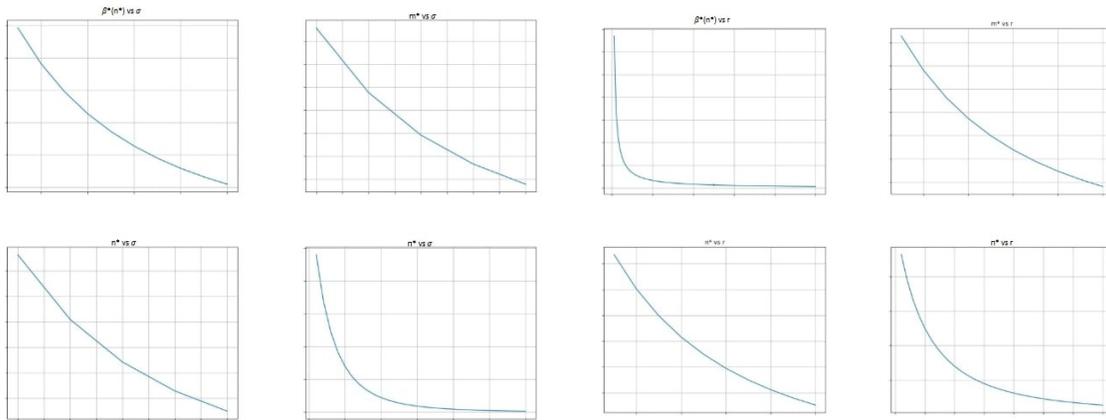
Fig. 3a. Equilibrium Changes
in Respect to σ

Fig. 3b. Equilibrium Changes
in Respect to r

the chief argument for innovation over safety, it is more surprising that increasing safety leads to less incentives for monitoring.

The reason this happens is through the effect of team size. Larger teams are riskier because more things can go wrong, so the principal reduces the team size to respond to this higher preference for safety. Because teams are now smaller and incentives are costly, the principal can afford to reduce the expensive incentives for the supervisor, which also leads him to monitor less. This counterintuitive result stems from the delicate interaction between team size as a design tool, versus more traditional methods like incentives alone.

4   Discussion

This paper provides insights into the design of multi-agent systems, yet several limitations should be acknowledged. First, the model assumes homogeneous agents, which simplifies the analysis and allows for closed-form solutions but does not capture the variability in agent capabilities that would exist in real-world settings. While some results, such as the directional effects of exogenous parameters on incentives and team size, likely extend to heterogeneous agents, other properties, like invariance, may not hold exactly in heterogeneous environments. This highlights an avenue for future research to explore how agent diversity affects optimal contracts, monitoring, and team formation.

Second, the focus on economic incentives as the primary tool for shaping agent behavior leaves out other critical considerations. AI safety, ethical constraints, and broader social impacts are not incorporated into the model, yet these factors could substantially influence equilibrium outcomes. For example, systems with high safety requirements may necessitate smaller teams or adjusted incentive structures, which could interact in complex ways with the economic optimization considered here. Future work could integrate these considerations, examining how ethical or safety constraints affect optimal design choices.

Third, the model assumes a homogeneous environment in which all agents face similar tasks, risks, and rewards. In practice, AI agents often encounter heterogeneous environments, where different subgroups of agents may face different challenges. Understanding how variability in the environment interacts with team composition and incentive design is an important extension. More realistic simulations could allow researchers to explore these effects, testing the theoretical predictions against diverse operational settings.

Finally, while the simulations provide intuition for how equilibrium quantities respond to changes in exogenous parameters, they remain stylized. Incorporating more detailed simulations of agent interactions and multi-agent learning dynamics could enhance the practical relevance of the results. Such simulations would allow exploration of heterogeneous risk preferences, variations in agent quality, and adaptive behaviors, providing a bridge between theory and real-world multi-agent systems.

5 Conclusion

AI agents in the future will likely operate in teams, making it essential to understand how to structure both their composition and reward systems. Two critical design choices are the number of agents assigned to a task and their associated reward functions. Because agents act autonomously and respond rationally to incentives, changes in rewards will influence their behavior toward activities that maximize compensation, reflecting principles similar to reinforcement learning.

The primary goal of this paper is to analyze the joint determination of team size and incentive structures. Our results demonstrate that these two design variables have distinct effects on equilibrium outcomes. In particular, adjusting team size is generally a more effective mechanism for responding to environmental changes than altering incentives alone. This implies that system designers have greater flexibility and control through team composition rather than modifying reward schemes. These findings offer both prescriptive guidance for designing multi-agent systems and predictions about how autonomous agents might self-organize in decentralized environments.

By solving for equilibrium team sizes and incentive levels, the paper shows how these outcomes shift in response to variations in agent quality, uncertainty, risk aversion, and resource costs. These insights provide a foundation for understanding multi-agent interactions and can inform future theoretical and empirical research. As AI agents become increasingly common in practice, the results highlight the importance of considering both team composition and incentive structures to optimize performance in decentralized, autonomous systems.

References

A. K. Agogino and K. Tumer. 2010. "A multiagent approach to managing air traffic flow." *Autonomous Agents and Multi-Agent Systems*, 24, 1, 1–25. doi:https://doi.org/10.1007/s10458-010-9142-5.

D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mané. 2016. "Concrete problems in AI safety." *arXiv*. doi:arXiv:1606.06565.

S. Baiman and J. S. Demski. 1980. "Economically optimal performance evaluation and control systems." *Journal of Accounting Research*, 18, 184–220. doi:https://doi.org/10.2307/2490338.

M. Braun, M. Greve, A. B. Brendel, and L. M. Kolbe. 2023. "Humans supervising artificial intelligence: Investigation of designs to optimize error detection." *Journal of Decision Systems*, 32, 1, 1–26. doi:https://doi.org/10.1080/12460125.2023.2260518.

A. Dorri, M. Steger, S. S. Kanhere, and R. Jurdak. 2017. "BlockChain: A distributed solution to automotive security and privacy." *IEEE Communications Magazine*, 55, 12, 119–125. doi:https://doi.org/10.1109/MCOM.2017.1700879.

J. N. Foerster, G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson. 2018. "Counterfactual multi-agent policy gradients." In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence (AAAI 2018)*. AAAI Press, 363.

B. Horling and V. Lesser. 2004. "A survey of multi-agent organizational paradigms." *The Knowledge Engineering Review*, 19, 4, 281–316. doi:https://doi.org/10.1017/S0269888905000317.

M. N. Huhns and L. M. Stephens. 1999. *Multiagent systems and societies of agents*. Ed. by G. Weiss. MIT Press, Cambridge, MA, 79–120.

N. R. Jennings. 1995. "Controlling cooperative problem solving in industrial multi-agent systems using joint intentions." *Artificial Intelligence*, 75, 2, 195–240. doi:https://doi.org/10.1016/0004-3702(94)00020-2.

N. R. Jennings, K. Sycara, and M. Wooldridge. 1998. "A roadmap of agent research and development." *Autonomous Agents and Multi-Agent Systems*, 1, 1, 7–38. doi:https://doi.org/10.1023/A:1010090405266.

J. R. Kok, N. Vlassis, and M. Littman. 2006. "Collaborative Multiagent Reinforcement Learning by Payoff Propagation." *Journal of Machine Learning Research*, 7, 9, 1789–1828.

X. Li and Y. Tan. 2019. "Leader-follower consensus of nonlinear multi-agent systems with fixed-time convergence." *Journal of the Franklin Institute*, 356, 7, 3579–3595. doi:https://doi.org/10.1016/j.jfranklin.2019.03.001.

R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch. 2017. "Multi-agent actor-critic for mixed cooperative-competitive environments." In: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017)*. Curran Associates Inc., 6382–6393. ISBN: 9781510860964.

X. Lyu, A. Baisero, Y. Xiao, B. Daley, and C. Amato. 2023. "On centralized critics in multi-agent reinforcement learning." *Journal of Artificial Intelligence Research*, 77, 295–354. doi:https://doi.org/10.1613/jair.1.14386.

L. Panait and S. Luke. 2005. "Cooperative multi-agent learning: The state of the art." *Autonomous Agents and Multi-Agent Systems*, 11, 3, 387–434. doi:https://doi.org/10.1007/s10458-005-2631-2.

Y. Qian. 1994. "Incentives and loss of control in an optimal hierarchy." *The Review of Economic Studies*, 61, 3, 527–544. doi:https://doi.org/10.2307/2297902.

I. Rahwan, L. Sonenberg, and F. P. M. Dignum. 2004. "On interest-based negotiation." In: *Advances in Agent Communication*. Lecture Notes in Computer Science. Vol. 2922. Ed. by F. Dignum. Springer, 292–306. doi:https://doi.org/10.1007/978-3-540-24608-4_22.

T. Sandholm. 2007. "Perspectives on multiagent learning." *Artificial Intelligence*, 171, 7, 382–391. doi:https://doi.org/10.1016/j.artint.2007.02.004.

Y. Shoham and M. Tennenholtz. 1995. "On social laws for artificial agent societies: Off-line design." *Artificial Intelligence*, 73, 1-2, 231–252. doi:https://doi.org/10.1016/0004-3702(94)00007-N.

G. L. Stewart and M. R. Barrick. 2000. "Team structure and performance: Assessing the mediating role of intrateam process and the moderating role of task type." *The Academy of Management Journal*, 43, 2, 135–148. doi:https://doi.org/10.2307/1556372.

N. Tomašev, M. Franklin, J. Z. Leibo, J. Jacobs, and W. A. Cunningham. 2025. "Virtual Agent Economies." *arXiv preprint arXiv:2509.10147*. Preprint. https://doi.org/10.48550/arXiv.2509.10147.

Unlocking the AI Agent Economy. https://www.masumi.network/whitepaper. Whitepaper, Masumi Network, accessed December 3, 2025. (n.d.).

M. Veale and R. Binns. 2017. "Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data." *Big Data & Society*, 4, 2, 2053951717743530. doi:10.1177/2053951717743530.

O. Vinyals, I. Babuschkin, W. M. Czarnecki, and et al.. 2019. "Grandmaster level in StarCraft II using multi-agent reinforcement learning." *Nature*, 575, 350–354. doi:https://doi.org/10.1038/s41586-019-1724-z.

M. M. de Weerdt, Y. Zhang, and T. Klos. 2011. "Multiagent task allocation in social networks." *Autonomous Agents and Multi-Agent Systems*, 25, 1, 46–86. doi:https://doi.org/10.1007/s10458-011-9168-3.

M. Wooldridge and N. R. Jennings. 1995. "Intelligent agents: Theory and practice." *The Knowledge Engineering Review*, 10, 2, 115–152. doi:https://doi.org/10.1017/S0269888900008122.

M. Yokoo and K. Hirayama. 2000. "Algorithms for distributed constraint satisfaction: A review." *Autonomous Agents and Multi-Agent Systems*, 3, 2, 185–207. doi:https://doi.org/10.1023/A:1010078712316.

I. de Zarzà, J. de Curtò, G. Roig, P. Manzoni, and C. T. Calafate. 2023. "Emergent cooperation and strategy adaptation in multi-agent systems: An extended coevolutionary theory with LLMs." *Electronics*, 12, 12, 2722. doi:https://doi.org/10.3390/electronics12122722.

B. Zhang, M. Anderljung, L. Kahn, N. Dreksler, M. C. Horowitz, and A. Dafoe. 2021. "Ethics and governance of artificial intelligence: Evidence from a survey of machine learning researchers." *Journal of Artificial Intelligence Research*, 71, 591–677. doi:https://doi.org/10.1613/jair.1.12895.

A. Ziv. 1993. "Performance measures and optimal organization." *The Journal of Law, Economics, and Organization*, 9, 1, 30–50. doi:https://doi.org/10.1093/oxfordjournals.jleo.a037054.

A   Proofs

Proof of Proposition 1: Let ν denote any of the symbols r, c, σ^2. Let $\bar{\nu}$ denote the product of the other two (i.e., if $\nu = r$, then $\bar{\nu} = c\sigma^2$). Implicitly differentiating (7) in ν yields

$$\frac{\partial \hat{n}}{\partial \nu} = -\frac{1}{\bar{\nu}\hat{n}G''(\hat{n})\nu^2} < 0. \tag{35}$$

Proof of Proposition 2: As before, let ν denote any of the symbols r, c, σ^2 and let $\bar{\nu}$ denote the product of the other two. Now, differentiating (6) and applying (35), we obtain

$$\delta_{\text{ex}}^{\nu}(\hat{n}) = \frac{\partial \hat{b}}{\partial \nu}\bigg|_{n=\hat{n}} = -\frac{q\bar{\nu}G(\hat{n})}{(1 + rc\sigma^2 G(\hat{n}))^2} < 0. \tag{36}$$

$$\delta_{\text{end}}^{\nu}(\hat{n}) = \frac{\partial \hat{b}}{\partial \nu}\bigg|_{n=\hat{n}} + \frac{\partial \hat{b}}{\partial n}\bigg|_{n=\hat{n}} \cdot \frac{\partial \hat{n}}{\partial \nu} = -\frac{q\bar{\nu}}{(1 + rc\sigma^2 G(\hat{n}))^2}\left[G(\hat{n}) - \frac{G'(\hat{n})}{rc\sigma^2 \hat{n} G''(\hat{n})}\right]. \tag{37}$$

From (7), we can substitute for $\frac{1}{rc\sigma^2}$ into the latter expression to obtain:

$$\begin{aligned}
\delta_{\text{end}}^{\nu}(\hat{n}) &= -\frac{q\bar{\nu}}{(1 + rc\sigma^2 G(\hat{n}))^2}\left[G(\hat{n}) - \frac{G'(\hat{n})}{\hat{n}G''(\hat{n})} \cdot [\hat{n}G'(\hat{n}) - G(\hat{n})]\right] \\
&= -\frac{q\bar{\nu}G(\hat{n})}{(1 + rc\sigma^2 G(\hat{n}))^2}\left[\frac{G'(\hat{n})}{G''(\hat{n})} \cdot \left(\frac{G''(\hat{n})}{G'(\hat{n})} - \frac{G'(\hat{n})}{G(\hat{n})} + \frac{1}{\hat{n}}\right)\right] \leq 0,
\end{aligned} \tag{38}$$

where the last weak inequality follows from the convexity of $G(\cdot)$ and condition (2).

From (36) and (38), it follows that $\delta_{\text{end}}^{\nu}(\hat{n}) > \delta_{\text{ex}}^{\nu}(\hat{n})$ if and only if

$$\frac{G'(\hat{n})}{G''(\hat{n})} \cdot \left(\frac{G''(\hat{n})}{G'(\hat{n})} - \frac{G'(\hat{n})}{G(\hat{n})} + \frac{1}{\hat{n}}\right) < 1, \tag{39}$$

which reduces to: $\hat{n}G'(\hat{n}) - G(\hat{n}) > 0$, which holds for all $\hat{n} > 0$ since $G(0) = 0$ and $G(\cdot)$ is convex.

Proof of Proposition 3: For the "if" part of the result, suppose that $G(n) = An^{\gamma}$, $A > 0$, $\gamma > 1$. Then, for arbitrary n, the expression in square brackets in (38) simplifies to:

$$\begin{aligned}
1 - \frac{G'(n)^2}{G(n)G''(n)} + \frac{G'(n)}{nG''(n)} &= 1 - \frac{A^2\gamma^2 n^{2\gamma-2}}{(An^{\gamma}) \cdot (A\gamma(\gamma-1)n^{\gamma-2})} + \frac{A\gamma n^{\gamma-1}}{nA\gamma(\gamma-1)n^{\gamma-2}} \\
&= 1 - \frac{\gamma}{\gamma-1} + \frac{1}{\gamma-1} \\
&= 0.
\end{aligned} \tag{40}$$

Therefore, $\delta_{\text{end}}^{\nu}(n) = 0$ for all values of n, i.e., reward-for-performance sensitivity is unaffected by changes to the parameter values (σ^2, r, c) once team size is adjusted optimally.

For necessity, note from Proposition 1 that changes to σ^2, r, or c lead to variations in the optimal n via equation (7). Therefore, if b is invariant to the parameters of the problem, it must be the case that $\delta_{\text{end}}^{\nu}(n) = 0$ for an open

set of values of n. From (38), this implies that for all n in some interval:

$$
\begin{aligned}
\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} &= 0, \\
\Leftrightarrow \frac{d}{dn}\left(\ln\left[\frac{nG'(n)}{G(n)}\right]\right) &= 0, \\
\implies \frac{nG'(n)}{G(n)} &= \theta, \\
\implies G'(n) - \frac{\theta}{n} \cdot G(n) &= 0
\end{aligned}
\tag{41}
$$

for some constant $\theta > 0$. Solving the first-order linear differential equation in 41 reveals that $G(\cdot)$ must belong to the class of power functions, thus completing the proof.

Proof of Lemma 1: From equations (21) and (22), the owner's profit function is

$$
\Pi(n) = \frac{n}{2c}\left(q - 2c\sigma\sqrt{rkG(n)/n}\right)^2,
\tag{42}
$$

and this is defined over the relevant range where $b^*(n) = q - 2c\sigma\sqrt{rkG(n)/n} \geq 0$. Note that $h(n) = \sqrt{\frac{G(n)}{n}}$ is an increasing function of n. This implies $n \in [0, \overline{n}]$, where $\overline{n}$ is such that $q - 2c\sigma\sqrt{rk}h(\overline{n}) = 0$. Evaluating $\Pi(n)$ at the two bounds, it is clear that $\Pi(\overline{n}) = 0$. Moreover,

$$
\Pi(0) = \frac{1}{2c} \cdot \lim_{n\to 0}\left(nq^2 + 4c^2\sigma^2rkG(n) - 4qc\sigma\sqrt{rknG(n)}\right) = 0.
\tag{43}
$$

By Rolle's theorem, we are therefore assured that an interior n^* exists with $\Pi'(n^*) = 0$. Now, the derivative of $\Pi(n)$ with respect to team size is given by:

$$
\Pi'(n) = \frac{b^*(n)}{2c}\left(q - 2c\sigma\sqrt{rk}h(n) - 4nc\sigma\sqrt{rk}h'(n)\right).
\tag{44}
$$

Since $b^*(n) > 0$ in the interior, any n^* that satisfies $\Pi'(n^*) = 0$ is characterized by:

$$
h(n^*) + 2n^*h'(n^*) = \frac{q}{2c\sigma\sqrt{rk}}.
\tag{45}
$$

The expression on the left is an increasing function of n if and only if $3h'(n) + 2nh''(n) > 0$. As $h(n) = \sqrt{\frac{G(n)}{n}}$, we have

$$
h'(n) = \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}}\left(\frac{1}{n} - \frac{G(n)}{G'(n)n^2}\right)
\tag{46}
$$

and

$$
h''(n) = \frac{\sqrt{n}G'(n)}{4n\sqrt{G(n)}}\left(\frac{2G''(n)}{G'(n)} + \frac{3G(n)}{n^2G'(n)} - \frac{G'(n)}{G(n)} - \frac{2}{n}\right).
\tag{47}
$$

Together, (46) and (47) imply that

$$
\begin{aligned}
3h'(n) + 2nh''(n) &= \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \cdot \left(2\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n}\right) \\
&> \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \cdot \left(\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n}\right) \\
&\geq 0,
\end{aligned} \tag{48}
$$

where the strict and weak inequalities follow from $G''(\cdot) > 0$ and condition (2), respectively.

Therefore, a unique n^* exists satisfying (45). To verify the second-order conditions at this point, we differentiate $\Pi'(n)$ with respect to n to obtain:

$$
\begin{aligned}
\Pi''(n) &= \frac{b^{*'}(n)}{2c}\left(q - 2c\sigma\sqrt{rk}h(n) - 4nc\sigma\sqrt{rk}h'(n)\right) \\
&\quad + \frac{b^*(n)}{2c}\left(-2c\sigma\sqrt{rk}h'(n) - 4c\sigma\sqrt{rk}h'(n) - 4nc\sigma\sqrt{rk}h''(n)\right).
\end{aligned} \tag{49}
$$

Evaluating this expression at n^*, we have

$$
\Pi''(n^*) = 0 + \frac{b^*(n^*)}{2c}\left(-2c\sigma\sqrt{rk}\right)\left(3h'(n^*) + 2n^*h''(n^*)\right) < 0, \tag{50}
$$

where the last inequality follows from (48) and $b^*(n^*) > 0$. We can therefore conclude that n^* is the globally optimal team size.

Proof of Proposition 4: Let ν denote any of c, σ, $\sqrt{r}$, and $\sqrt{k}$, and let $\bar{\nu}$ denote the product of the other three. Implicitly differentiating the condition for n^* in (45) shows that

$$
\frac{\partial n^*}{\partial \nu} = (3h'(n^*) + 2n^*h''(n^*))^{-1}\left(-\frac{q}{2\bar{\nu}\nu^2}\right) < 0, \tag{51}
$$

as the first term is strictly positive from (48).

Proof of Proposition 5: As before, let ν denote any of c, σ, $\sqrt{r}$, and $\sqrt{k}$, and let $\bar{\nu}$ denote the product of the other three. First, differentiating (21), we immediately obtain

$$
\phi_{\text{ex}}^{\nu}(n^*) = -2\bar{\nu}h(n^*) < 0. \tag{52}
$$

Next, implicitly differentiating the first order-condition (45) in Lemma 1, we have,

$$
\frac{\partial n^*}{\partial \nu} = -\frac{q\bar{\nu}}{2(3h'(n^*) + 2nh''(n^*))(\nu\bar{\nu})^2} < 0 \text{ (from (48))}. \tag{53}
$$

We can then compute $\phi_{\text{end}}^{\nu}(n^*)$ as:

$$
\begin{aligned}
\phi_{\text{end}}^{\nu}(n^*) &= \left.\frac{\partial b^*}{\partial \nu}\right|_{n=n^*} + \left.\frac{\partial b^*}{\partial n}\right|_{n=n^*} \cdot \left.\frac{\partial n^*}{\partial \nu}\right|_{n=n^*} \\
&= -2\bar{\nu}h(n^*) + \frac{2q\bar{\nu}h'(n^*)}{(3h'(n^*) + 2n^*h''(n^*))(2\nu\bar{\nu})}.
\end{aligned} \tag{54}
$$

Substituting for $q/(2\nu\bar{\nu})$ using condition (45), we obtain:

$$
\begin{aligned}
\phi^\nu_{\text{end}}(n^*) &= -2\bar{\nu}\left[h(n^*) - \frac{h'(n^*)(h(n^*) + 2n^*h'(n^*))}{3h'(n^*) + 2n^*h''(n^*)}\right] \\
&= -4\bar{\nu}\left[\frac{h'(n^*)h(n^*) + n^*h(n^*)h''(n^*) - n^*h'(n^*)^2}{3h'(n^*) + 2n^*h''(n^*)}\right].
\end{aligned}
\tag{55}
$$

The denominator of (55) is strictly positive (from (48)). Using equations (46) and (47), the numerator can be simplified to:

$$
\frac{G'(n)}{2} \cdot \left(\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n}\right) \geq 0,
\tag{56}
$$

where the inequality follows from Condition (2).

We have therefore established that $\phi^\nu_{\text{end}}(n^*) \leq 0$. Moreover, using (48) and the fact that $h'(\cdot) > 0$, a direct comparison of 52 and (55) reveals that $\phi^\nu_{\text{ex}}(n^*) < \phi^\nu_{\text{end}}(n^*) \leq 0$.

Turning to the supervisor's incentives, we know that with identical agents,

$$
\beta^* = b^*\sqrt{\frac{n}{2}}.
\tag{57}
$$

Therefore,

$$
\psi^\nu_{\text{ex}}(n^*) = \sqrt{\frac{n^*}{2}} \cdot \phi^\nu_{\text{ex}}(n^*) = -2\bar{\nu}h(n^*)\sqrt{\frac{n^*}{2}} < 0.
\tag{58}
$$

A comparison of (52) and (58) immediately reveals that $\psi^\nu_{\text{ex}}(n^*) \leq \phi^\nu_{\text{ex}}(n^*)$, with strict inequality for all $n^* > 2$.

To complete the proof, we next demonstrate that

$$
\psi^\nu_{\text{ex}}(n^*) = \psi^\nu_{\text{end}}(n^*).
\tag{59}
$$

Since

$$
\psi^\nu_{\text{end}}(n^*) = \psi^\nu_{\text{ex}}(n^*) + \left.\frac{\partial \beta^*}{\partial n}\right|_{n=n^*} \cdot \left.\frac{\partial n^*}{\partial \nu}\right|_{n=n^*},
\tag{60}
$$

it is sufficient to show that $\left.\frac{\partial \beta^*}{\partial n}\right|_{n=n^*} = 0$. Using (21), we can expand β^* as follows:

$$
\beta^* = b^*\sqrt{\frac{n}{2}} = \sqrt{\frac{n}{2}} \cdot (q - 2c\sigma\sqrt{rkG(n)/n}) = \sqrt{\frac{n}{2}} \cdot (q - 2\nu\bar{\nu}\sqrt{G(n)/n}).
\tag{61}
$$

Differentiating this expression, we obtain:

$$
\begin{aligned}
\left.\frac{\partial \beta^*}{\partial n}\right|_{n=n^*} &= \frac{q}{\sqrt{2}}\frac{1}{2\sqrt{n^*}} - \frac{\nu\bar{\nu}}{\sqrt{2}}\frac{G'(n^*)}{\sqrt{G(n^*)}} \\
&= \frac{\nu\bar{\nu}}{\sqrt{2n^*}} \cdot \left(\frac{q}{2\nu\bar{\nu}} - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}}\right) \\
&= \frac{\nu\bar{\nu}}{\sqrt{2n^*}} \cdot \left(h(n^*) + 2nh'(n^*) - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}}\right) \quad \text{(from (45))} \\
&= \frac{\nu\bar{\nu}}{\sqrt{2n^*}} \cdot \left(\sqrt{\frac{G(n^*)}{n^*}} + \frac{\sqrt{n^*}G'(n^*)}{\sqrt{G(n^*)}} - \frac{\sqrt{G(n^*)}}{\sqrt{n^*}} - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}}\right) \\
&= 0.
\end{aligned}
\tag{62}
$$

Proof of Proposition 6: Suppose $G(n) = An^\gamma$, $A > 0$, $\gamma > 1$. Then, for arbitrary n^*, the numerator of $\phi_{\text{end}}^\nu(n^*)$, given in reduced form in (56), simplifies to:

$$\frac{G'(n)}{2} \cdot \left(\frac{G''(n^*)}{G'(n^*)} - \frac{G'(n^*)}{G(n^*)} + \frac{1}{n^*} \right) = \frac{G'(n)}{2} \cdot ((\gamma - 1)(n^*)^{-1} - \gamma(n^*)^{-1} + (n^*)^{-1})$$

$$= 0. \tag{63}$$

Thus, $\phi_{\text{end}}^\nu(n^*) = 0$ everywhere, i.e., reward-for-performance is totally insensitive to changes in any of the four parameters of interest, (σ^2, r, c, k), provided team size is adjusted optimally.

For the "only if" argument, we have shown in Proposition 5 that changes to σ^2, r, c or k lead to strict variations in the optimal n via equation (45). Therefore, if b is invariant to any of these parameters, it implies that $\phi_{\text{end}}^\nu(n) = 0$ for an open set of values of n. From equations (55) - (56), this implies that for all n in some open interval:

$$\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} = 0. \tag{64}$$

As in the proof of Proposition 3, it can be shown that this implies that $G(\cdot)$ belongs to the class of power functions.