Incentive Design for Agile Teams*

Zhuoran Lu[†] Yangbo Song[‡]

February 15, 2022

Abstract

Motivated by the ever-increasing prevalence of agile teams, we study a team incentive design problem where multiple agents are located on a network and work on a joint project. The principal seeks the least costly mechanism to incentivize full efforts, by choosing the work assignment sequence and the rewards to the agents upon success. Whereas the agents' actions are hidden to the principal, they may be observed among the agents given the internal information that is determined by the network and the sequence. We show that under effort complementarity, the transparency of the agents' actions can reduce their incentive costs. Moreover, the effectiveness of transparency decreases as an action becomes more transparent. In the optimal sequence, the agents work sequentially in the order of ascending individual importance to the project. The agents who move later effectively monitor their preceding peers and have higher incentive costs. When multiple agile teams collaborate, more important agents also move later in their respective teams, while larger teams are allocated toward both ends of the optimal sequence. Meanwhile, only a small fraction of the agents, as compared to the overall agent group, will serve as monitors. For several typical networks, simple algorithms are offered to explicitly characterize the optimal mechanism.

Keywords: Sequential task assignment, Peer information, Network, Incentive design

JEL Classification: D82, L14, L23

*We are grateful to Simon Board, Bob Gibbons, Moritz Meyer-ter-Vehn, Ichiro Obara, Marek Pycia, John Riley, Tomasz Sadzik and Birger Wernerfelt, as well as seminar audience at Fudan, MIT Sloan, UCLA, UIBE and XMU for helpful comments. Any remaining errors are our own.

[†]School of Management, Fudan University. Email: luzhuoran@fudan.edu.cn.

[‡]School of Management and Economics, the Chinese University of Hong Kong, Shenzhen. Email: yangbosong@cuhk.edu.cn.

1 Introduction

"Agile teams" refer to a modern business model that has presented both opportunities and challenges to the organization structure for value creation. In a typical agile team, a cross-functional group of people, usually with various individual expertise, are assembled on a temporary basis to collaborate on building or testing a product. Notable examples include Amazon's "two-pizza teams" and Spotify's invention of its "Discover Weekly" function. On a larger scope, collaboration among multiple agile teams is also common, for instance film production that involves independent crews in different locations, or an NSF-funded R&D project requiring joint effort by various laboratories or universities. Such a work process has two outstanding features. First, it still reflects the classical idea of teamwork, namely efforts of members are interrelated for the product's success. Second, it is less constrained in terms of work assignment compared to a conventional waterfall team with formalized development steps and milestones, meaning that it has more flexibility in designing its micro-structure and associated incentives.

In this paper, we study the incentive design problem for an agile team, as well as that for the collaboration among multiple agile teams. We seek to address, in this particular setting, Holmstrom (1982)'s classic remark: "...monitoring technologies were exogenously given. In reality, they are not. The question is what determines the choice of monitors; and how should output be shared so as to provide all members of the organization (including monitors) with the best incentives to perform?" The core economic issue we focus on can be summarized as "endogenous internal information" or "endogenous peer monitoring", i.e., which agents essentially serve as monitors and which are to be monitored. In teamwork, members may observe others' efforts in a possibly asymmetric way; for instance, in an agile team there is typically one member (sometimes called the "scrum master") who is better informed than others of the team members' actions. The project owner, on the other hand, cannot monitor the entire team and thus faces the classical moral hazard problem, and will craft a solution by exploiting this internal observation structure. We propose a concise and tractable model of incentive design in this environment, and explicitly characterize the principal's optimum using a strategically designed sequence of work.

Our model considers a group of multiple agents who jointly work on a risky project, each agent responsible for an individual task. An agent can increase the chance of the project's success by exerting costly effort, while the exact marginal effect of his effort also depends on other team members' inputs. Throughout the paper, we mainly focus on complementary technologies, i.e., the project's success probability is supermodular in agents' efforts.

Agents choose their (binary) effort level sequentially on a discrete time line, and they may be able to observe the efforts chosen by earlier movers. Specifically, there exists an internal observation structure among the agents, which we model by an exogenous network with bilateral links. If two agents i and j are linked, j can observe whether i exerted effort when j moves after i, and vice versa. A denser network suggests more transparent internal information, with the highest level of transparency represented by a complete network and the lowest by an empty network. For studying agile teams, we typically use a star network to represent one such team, with the scrum master as the center and other team members as peripheral agents, but our framework readily applies to other types of networks.

The principal cannot observe the effort choice of any agent but is aware of the network topology, and aims at inducing effort from each agent at the lowest possible costs. Our main departure from the literature on incentive design in teams is the set of available mechanisms that the principal can implement. Specifically, the principal determines (1) the sequence of moves by the agents, and (2) the reward to each agent upon success of the project. To the best of our knowledge, this paper is the first to consider the task assignment sequence in a team incentive design problem, which is a distinctive feature of agile teams.

Drawing on relevant literature such as Winter (2010), we have a ready characterization of the optimal reward scheme given a fixed sequence of moves, and thus our analysis may focus on the novel element of designing an optimal sequence. A carefully designed sequence benefits the principal because it exploits the externality between agents' efforts when internal information among agents is present. For instance, consider again agents i and j who are linked, and suppose that j moves after i. At i's turn, i realizes that under a complementary technology, shirking from his own task makes j's effort less effective for the project's success, leading to the chain reaction that j also tends to shirk from his task. Therefore, i's effort now matters more for the project and the principal only needs to promise i a lower reward. As a whole, the principal's problem is to maximize the leverage of such externality given the internal observation structure, so as to minimize the sum of effort-inducing rewards.

Our results characterize the optimal incentive design in two representative environments. We first focus on a "micro" scenario and investigate the principal's optimum for one single team. We consider agents that are heterogeneous in importance, i.e., a ranking order exists in the contribution of their efforts to the project's success. For the full transparency benchmark represented by a complete network, we find that the uniquely optimal sequence is by reversed importance: the least important agent moves first, then the second least important, etc., and the most important agent moves last. This is because an agent requires less incentive to work when they realize that their shirking will incur a larger loss in the success probability, and postponing a more important agent's turn to move applies this leverage to more predecessors.

For a typical agile team represented by a star network, the optimal sequence is also unique with two properties. First, analogous to the full transparency benchmark, every predecessor of the center is more important than every successor. Second, the center is generically at an interior position in the sequence, with both successors and predecessors. The optimal sequence essentially represents a balance in both the extensive margin and intensive margin. On the one hand, sequential work assignment is always desired as some internal transparency helps the principal in reducing total reward needed for incentivizing full efforts. By making some peripheral agents successors of the center, the predecessors' incentive costs can be significantly reduced. On the other hand, the marginal benefit of increasing transparency is decreasing in the degree of transparency. When more agents become successors so that the observation structure becomes more transparent, the net marginal effect on the total incentive costs varies monotonically from a decreasing reduction to an increasing addition. Hence, the exact set of successors at optimum reflects a trade-off between the principal's two countervailing motives: a larger set of predecessors to increase the number of less-rewarded agents, and a larger set of successors to lower the reward to each predecessor.

Our analysis then turns to optimal incentive design from a "macro" perspective, focusing on the collaboration among several agile teams for a larger project. We model this scenario by connecting the team centers, which essentially represents a hierarchy where members in central position are also better informed, both within and across teams.

We identify and highlight cross-team effects which lead to different implications when a team's importance to the whole project's success has different origins. When teams have the same size, i.e., agent number, but can be ranked in terms of individual agents' importance, the optimal task assignment is asymmetric across teams in a monotonic pattern with respect to agents' importance ranking. In every team, more important agents are successors of their center while less important ones predecessors. Across teams, more important teams are assigned to move later, and have more successors, than less important ones. Reflecting the trade-off in extensive and intensive margins across teams, this property is consistent with the previous optimal design in a single team. However, a new insight arises when heterogeneity across teams is reflected by size instead of agents' individual importance: although larger teams still bear a larger weight for the project's success, the effect of size heterogeneity on incentive design is not monotone. The optimal sequence exhibits a "V-shape" in team size, with larger teams located closer to either end. This is because the principal finds it beneficial to both create more leverage by delaying task assignment of some large teams, and utilize such leverage to a greater extent by bringing forward task assignment of other large teams.

As it turns out, the above distinct findings can be unified under one perspective. The principal essentially tries to identify two sets of agents—the monitored who require less reward, and the monitors to ensure incentive compatibility of the reward reduction. The former should be sufficient in number while the latter in importance. Between teams of the same size, a more important team makes better monitors for the principal since the members require less reward by themselves already and impose greater implicit threat onto the monitored. In contrast, between teams of identical individual importance, a larger team is preferred for both roles, resulting in a V-shaped sequence to maximize reward reduction with the centers being information intermediaries. Hence, in response to Holmstrom (1982)'s question in the realm of agile teams, we point out that an agent who ranks higher in importance for the project's success is more likely to be endogenously assigned a monitor's role, while an agent in a larger team either monitors more peers (indirectly), or is monitored by more peers. Moreover, due to the diminishing benefit from transparency, we show that only a small proportion of agents will serve in the optimal sequence as monitors. In other words, when the overall agent group enlarges, the endogenous selection criterion for monitors becomes increasingly stringent.

Finally, in designing the optimal incentive scheme for both one single team and multiple collaborating teams, we propose simple algorithms to explicitly characterize the optimum for several typical networks. These algorithms guarantee that the optimal sequence is identified within a number of calculations bounded above by the number of agents.

Related literature. The theoretical literature on incentive design for teamwork is extensive and growing. The trade-off an agent faces between working and shirking originated from the classical literature on moral hazard in teams (Alchian and Demsetz, 1972; Holmstrom, 1982; Holmstrom and Milgrom, 1991; Itoh, 1991). Subsequent studies developed this literature to static contracting on teamwork with a number of variations, such as externalities (McAfee and McMillan, 1991; Segal, 1999; Babaioff et al., 2012), specialization versus multitasking (Balmaceda, 2016), loss-averse agents (Balmaceda, 2018), and network-based production spillover (Sun and Zhao, 2021). Our main contribution to this literature is to consider the endogenization of internal information among the agents in the presence of moral hazard.

Several recent papers have investigated how including or altering the scheme of information sharing among agents affects incentive design. Zhou (2016) shows that the welfare-optimal organization of team members is a chain when the first mover observes the state of nature and the later movers observe their immediate predecessor's effort. Our analysis produces a similar result when the exogenous network of internal information is complete. Gershkov et al. (2016) study the efficient contract design given that some team member may share information about a payoff-relevant state, and they show that efficiency can be achieved if contracts take into account a contest ranking across agents. Au and Chen (2021) characterize the optimal long-term contract in teams of two members, with efforts observable between the paired agents. In Camboni and Porcellacchia (2021), the principal observes noisy signals about efforts, and may condition the contract offered to each team member on both her individual signal and the whole project's outcome. The optimal incentive scheme features a partition between insulated and non-insulated agents ranked by signal precision.

A comprehensive study on the role of internal information in effort-based teamwork, with an exogenous sequence of task assignment, has been provided by Winter (2004, 2006, 2010). This is the main strand of literature we follow on building the theoretical framework. Our results indicate that some peer information architectures are more likely to emerge than others, once the sequence becomes the principal's choice. Halac et al. (2021) also investigate the incentive design problem in teamwork in the presence of moral hazard, but the principal leverages uncertainty of ranking among agents instead of internal information. Consequently, they show that discrimination is suboptimal in contrast to Winter (2004). The broad idea of the interplay between early and late movers in teams, and how a designer can exploit such structure for efficiency or cost saving, have also been studied from other perspectives including information-based leadership (Hermalin, 1998; Zhou and Chen, 2015; Zhou, 2016), tournaments in team production (Gershkov et al., 2009, 2016) and various forms of payoff externalities (Che and Yoo, 2001; Segal, 2003; Bernstein and Winter, 2012).

The importance of internal information in incentive design has also been noted by data. Empirical evidence suggests that workers' productivity and willingness to work respond positively to observed efforts of peers (Ichino and Maggi, 2000; Heywood and Jirjahn, 2004; Gould and Winter, 2009; Mas and Moretti, 2009). Experimental studies on behavior in team production have also indicated that an agent's contribution in teamwork is highly responsive to internal information (Carpenter et al., 2009; Steiger and Zultan, 2014) and that unequal rewards tend to facilitate coordination and improve efficiency (Goerg et al., 2010).

Organization. The rest of the paper is organized as follows. Section 2 lays out the model. Section 3 analyzes the full transparency benchmark and presents some preliminary results. Section 4 and 5 characterizes, respectively, the optimal incentive design in a single team and in collaborating teams. Section 6 concludes. All proofs are in the Appendix.

2 Model

Players and actions. A principal (she) owns a project that is collectively managed by a team, denoted as a set I of n agents. Each agent (he) is responsible for a single task and decides whether to exert effort. Formally, each agent i chooses $a_i \in A_i \equiv \{0, 1\}$, with $a_i = 1$ if he chooses to exert effort and $a_i = 0$ if he shirks. The cost of effort is $c > 0$ and constant across all the agents, whereas shirking is costless. To save on notation, we normalize c to 1 without loss of generality. Hereafter, we use the terms *work* and *exert effort* interchangeably.

Technology. The organization's technology is a mapping from a profile of effort levels to a probability of the project's success. For a subset $S \subseteq I$ of working agents, the probability of the project's success is $p(S)$. Throughout the paper, we assume that p is increasing in the sense that if $T \subset S$, then $p(T) < p(S)$. Moreover, we distinguish between the technology's properties of complementarity and substitutability. A technology p satisfies *complementarity* among agents if for every two sets of agents S and T with $T \subset S$ and every agent $i \notin S$, we have $p(S \cup \{i\}) - p(S) > p(T \cup \{i\}) - p(T)$; that is, i's effort is more effective if the set of other agents who exert effort enlarges. Conversely, we say that p satisfies *substitutability* among agents if $p(S \cup \{i\}) - p(S) \leq p(T \cup \{i\}) - p(T)$. In addition, we distinguish between different agents' *importances* to the project. We say that agent i is (weakly) more important than j if for every coalition $S \subseteq I$ with $i, j \in S$, we have $p(S \backslash \{i\}) \leq p(S \backslash \{j\})$; that is, i's shirking is more detrimental than j's to the chance of success.[1] We assume that the set I is totally ordered in terms of agent importance. Analogously, we say that a set of agents T is (weakly) more important than another set T' if for every coalition $S \subseteq I$ with $T, T' \subseteq S$, we have $p(S \backslash T) \leq p(S \backslash T')$; in particular, if $T' \subset T$, then this holds automatically due to the monotonicity of p.

Network. The organizational structure of the team, also referred to as the *network* of the agents, is represented by an exogenous and undirected graph g. We write $ij \in g$ to indicate that agents i and j are directly linked, and say that i and j are *neighbors*; in particular, we assume that $ii \notin g$ for any agent i.

In applications of our framework, network g could stem from the workplace architecture, the authority structure, geographical locations, informal social networks and so forth. While over time g may change, we assume in the context of agile team that g is fixed within the relatively short duration of the project, and that the structure of g is common knowledge.

[1] Alternatively, one may find it more intuitive to say that if i is (weakly) more important than j, then for any S with $i, j \notin S$, $p(S \cup \{i\}) \geq p(S \cup \{j\})$, i.e., i's effort is more effective than j's, *ceteris paribus*. Whereas the two definitions are clearly equivalent, we find the former more convenient for our subsequent analysis.

Mechanism. Before the agents perform the tasks, the principal designs a work sequence, or simply a *sequence*, π, such that agent i is the π_i-th player to move, with $\pi_i \in \{1, \ldots, n\}$. The principal cannot monitor the agents' efforts, but simply knows whether the project is a success after all the tasks have been performed. In addition, the principal designs a reward scheme $v = (v_1, \ldots, v_n)$, such that agent i receives $v_i \geq 0$ if the project turns out to be a success, and receives zero payoff otherwise.

A *mechanism* $\{\pi, v\}$ consists of a sequence π and a reward scheme v. Throughout, we assume that the principal can commit to the mechanism.

Internal information. The agents' *internal information* about their peers' effort levels is jointly determined by the graph g and the sequence π. Specifically, agent i observes agent j's action, or simply i *sees* j, before i moves if and only if i and j are neighbors and i moves after j.[2] That is, $ij \in g$ means that i *can* see j based on the network, and i *will* see j when he moves after j. As g is exogenous, we drop it in the subsequent notations. For every π, we define $N_i(\pi) := \{j | ij \in g, \pi_i > \pi_j\}$ to be the set of agents whom agent i sees given the internal information. To save on notation, we write N_i for the set $N_i(\pi)$ henceforth.

Principal's problem. Consider the game that is defined by the set of agents I, the agents' action space $\{A_i\}_{i \in I}$, the network g and a mechanism $\{\pi, v\}$. A (pure) strategy of agent i is a mapping $s_i : 2^{N_i} \to \{0, 1\}$, which specifies the agent's action as a function of his information about the effort levels of the agents in N_i. Given a strategy profile $s = (s_1, \ldots, s_n)$, agent i's expected utility equals

$$U_i(s) := p(W(s))v_i - \mathbb{1}(s_i = 1),$$

where $W(s)$ is the set of agents who work given s, and $\mathbb{1}(\cdot)$ is the indicator function.

A mechanism $\{\pi, v\}$ is *effort-inducing (EFI)* with respect to the network if there exists a perfect Bayesian equilibrium (PBE) in the resultant game, so that all the agents exert effort. The principal's problem is to design an EFI mechanism that yields minimal total rewards to the agents among all EFI mechanisms, which is called an optimal mechanism. In particular, given a sequence π, a reward scheme $v^*(\pi)$ is optimal if $\{\pi, v^*(\pi)\}$ is an optimal mechanism. The principal's objective is meaningful when the project's value is relatively high and the agents' efforts are efficient to raise the probability of success. Alternatively, one can consider the mechanisms that maximizes the principal's expected net profit, but we refrain from this approach as it does not provide new insights while complicates the analysis remarkably.

[2]If i and j move simultaneously, then neither of them can see the other.

2.1 Optimal Reward Scheme

As a helpful preliminary result, we first characterize the optimal reward scheme for a fixed sequence π, applying Winter (2010)'s main results.[3]

Define $M_i(\pi)$, M_i for short, to be the set of agents such that for each $j \in M_i$, there exists a sequence $\{k_r\}$ in which j sees k_1 sees k_2 sees $\ldots k_r$ sees i. For expositional convenience, we call the agents in M_i those who can *ultimately learn* i's action based on the notion that everyone in M_i would be informed of i's action if an agent could share his information with those who see him. The proposition below characterizes the optimal reward scheme $v^*(\pi)$ with respect to an arbitrary sequence π.

Proposition 1. *For any fixed sequence π: (i) if p is complementary, then the optimal reward scheme $v^*(\pi)$ pays agent i $v_i^* = [p(I) - p(I \backslash (\{i\} \cup M_i))]^{-1}$; (ii) if p is substitutable, then $v^*(\pi)$ is invariant with respect to π and pays agent i $v_i^* = [p(I) - p(I \backslash \{i\})]^{-1}$.*

When the agents perform the tasks sequentially, they face an *implicit threat of shirking*. Specifically, the exposure of a low effort might induce an agent who observes this action to shirk and consequently triggers a *domino effect* of shirking, making success less likely. This implicit threat thus reduces the agent's incentive cost. Under a complementary technology and the optimal reward scheme, it is indeed sequentially rational for an agent to shirk once he sees someone shirking, making the implicit threat credible. Moreover, Proposition 1 implies that if agent i's action becomes more transparent in the sense that the set M_i enlarges, then i should be rewarded less because he is more willing to work due to a greater implicit threat. That is, under complementarity action transparency can reduce an agent's incentive cost.

In contrast, under a substitutable technology, the internal information has no impact on incentives as if all the agents moved simultaneously. To implement full efforts, the principal must provide the agents with sufficient incentives when they believe that all their teammates are working. However, under substitutability, such a reward scheme gives each agent an even stronger incentive to work when he sees someone shirking. This eliminates the implicit threat of shirking that is critical in the complementarity case, thereby preventing the principal from saving the incentive costs through the design of internal information.

In what follows, we shall focus on the more interesting case of complementarity. Because I is finite, Proposition 1 ensures that an optimal mechanism exists; thus, it remains to find such a mechanism by characterizing the optimal sequence.

[3] Winter (2010) characterizes the optimal reward scheme for a fixed internal information structure under substitutability (Proposition 2) and complementarity (Proposition 4), respectively.

3 Full Transparency Benchmark: Complete Networks

As a reference point, we study a complete network where all the agents are interconnected. This network topology leads to the richest internal observation structure in the sense that in any sequence, each agent can observe all preceding actions.[4] Before analyzing complete networks, we lay out two general results which hold for any network g.

First, we show that if two agents are neighbors, they cannot move simultaneously in the optimal sequence. This is summarized by the lemma below:

Lemma 1. *For any two agents i and j, if $ij \in g$, then $\pi_i^* \neq \pi_j^*$.*

Under complementarity, it is suboptimal to make two neighbors perform simultaneously. This is because doing so reduces the transparency of actions, thereby mitigating the implicit threat of shirking and thus increasing incentive costs.

Second, we show that if two agents share the same set of neighbors other than themselves and either one can learn the other's action in the optimal sequence, then the more important agent moves later. Formally, we have the following lemma:

Lemma 2. *For any two agents i and j such that $\{k|ik \in g, k \neq j\} = \{k|jk \in g, k \neq i\}$ and i is more important than j, if in π^* either $i \in M_j^*$ or $j \in M_i^*$, then $\pi_i^* > \pi_j^*$.*

Intuitively, if the more important agent i moves before the less important agent j, then the principal can profit by switching their orders. Because i and j share the same set of neighbors other than themselves, swapping i and j would only affect the rewards of i, j and every agent k such that $j \in M_k$ and $i \notin M_k$, i.e., the agents whose actions can be ultimately learned by j but not i. In particular, the principal can reduce k's reward by replacing j's position with i, because if k shirks he will then trigger a more important agent to shirk, leading to a greater implicit threat. Analogously, i's new reward will be lower than j's old reward, while j's new reward equals i's old reward. Thus, the principal can reduce the total incentive costs by assigning the more important agent to the later stage.

Now consider a complete network. Lemma 1 implies that in the optimal sequence, the agents move sequentially in the order $1, 2, \ldots, n$; thus, agents in later stages effectively serve as the monitors of the team, and will punish their peers' shirking actions by shirking as well. In addition, Lemma 2 implies that the agents move in ascending order of importance; thus, the monitors are relatively more important. For ease of exposition, relabel the agents such

[4] Winter (2006) considers a similar setting in which the tasks must be performed sequentially in the order $1, 2, \ldots, n$ and the agents can observe all preceding actions, and identifies the optimal sequence with respect to the importance of agent, as well as the associated reward scheme.

that agent i is (weakly) less important than $i+1$, $i \leq n-1$, with n being the most important. The next proposition characterizes the optimal mechanism in a complete network.

Proposition 2. *If g is a complete network and the agents are increasingly important, then the optimal mechanism $\{\pi^*, v^*\}$ satisfies that (i) π^* is the identity permutation; (ii) agent i has a reward $v_i^* = [p(I) - p(\{j|j < i\})]^{-1}$. In particular, if two agents are equally important, then exchanging their orders only still leads to an optimal sequence.*

A prominent implication of Proposition 2 is that agents in later stages (the monitors) should be rewarded more generously even if all the agents are equally important. Intuitively, an agent moving later has a less transparent action, and thus, is more costly to incentivize. Moreover, Proposition 2 implies that the gap between two adjacent agents' rewards increases along the optimal sequence. This is because under complementarity a low effort is less detrimental to the project when fewer agents work. This fact implies that while the transparency of an agent's action can reduce his incentive cost, the effectiveness of transparency on providing incentives is diminishing as the action becomes more transparent. To illustrate, suppose a subset S of agents who could not initially learn agent i's action are now able to learn the action, then by Proposition 1 the change in the optimal reward to i is given by

$$\frac{p(I \backslash (\{i\} \cup M_i)) - p(I \backslash (\{i\} \cup M_i \cup S))}{[p(I) - p(I \backslash (\{i\} \cup M_i))][p(I) - p(I \backslash (\{i\} \cup M_i \cup S))]}. \tag{1}$$

Note that as the set M_i enlarges, the numerator decreases due to complementarity, while the denominator increases due to monotonicity; thus, the ratio decreases. That is, the marginal reduction in the reward is decreasing as i's action becomes more transparent. This implies that in a complete network, an agent's reward is convex in his order. To summarize:

Corollary 1. *If g is a complete network, then under the optimal mechanism $\{\pi^*, v^*\}$, v_i^* is increasing and strictly convex in i.*

Corollary 1 means that ex ante identical agents may obtain increasingly different rewards when the technology is complementary and the principal wishes to implement the least-cost full-effort mechanism by leveraging peer-monitoring. This is because the agents may be ex post different with respect to the optimal internal information, and under complementarity action transparency can reduce agents' incentive costs, with a diminishing marginal effect.

To conclude this section, note that a complete network yields the richest transparency, and thus imposes the greatest implicit threats of shirking on the agents. The corollary below states that the total rewards to the agents are the least in complete networks among all network topologies. This also means that Proposition 2 provides a sharp lower bound for the total rewards of optimal mechanisms.

Corollary 2. *A complete network generates minimal total payoffs to the agents, and thus maximal payoff to the principal.*

A complete network is the most favorable for the principal as it yields the richest internal information. Such a network can represent the emerging workplace architecture "war room" that is adopted by different organizations. The movement to such open-space environment allows workers to better monitor their peers' efforts, thereby enhancing the implicit incentive of working, which is generated by this mutual observability. In our model, the network g is exogenously given; in other situations, the principal may be able to improve the connection between agents (i.e., by adding links to g); hence, she may find it profitable to transform g into a complete network by for example upgrading IT infrastructure, adopting open-space workplace architecture, organizing more recreational activities for employees, and so forth.

4 Incentive Design for a Single Team

In this section, we study the optimal incentive design for a single team that is represented by a star network. A star network features a particular node i such that every link in the network involves node i; thus, agent i is termed as the *center* of the star, and the rest of the agents are termed as the *peripheral agents*. In what follows, we assume that $n \geq 3$.

Star network structures are common in organizations, and often serve as basic units for more complex networks. For example, in most scientific labs, when a project leader works with his/her fellow researchers, the leader often serves as the center of the team, while each fellow researcher works on an individual task and communicates the progress only to the leader. Such a team thus has a star network structure. Usually, the principal investigator (PI) of the lab can only observe the outcome of the entire project and chooses how to reward the team based on the final result. Other examples of star networks may include a general contractor and subcontractors, a book editor and chapter contributors, and so forth.

To find the optimal sequence for a star network, note that it suffices to characterize the set of the center's successor(s), with the possibility of an empty set. For ease of exposition, we relabel the peripheral agents by importance from 1 to $n-1$, with a higher index referring to a more important agent. Provided there is no confusion, let the center be the n-th agent who is not necessarily the most important agent. Note that every peripheral agent has the same unique neighbor, i.e., the center. Then by Lemma 2, we have the following lemma:

Lemma 3. *If in π^* the center has both a nonempty set of predecessors and a nonempty set of successors, then the successors are uniformly more important than the predecessors.*

The intuition of Lemma 3 has been suggested already; that is, if more important agents move in later stages, then a low effort will induce agents with higher importance to shirk and is thus more detrimental to success, thereby allowing the principal to reduce incentive costs. The relative importance between the center's predecessors and successors implies that the optimal sequence for a star network can be summarized by a sufficient statistic, that is, the number of the center's successor(s).[5] Let m be the number of the center's successor(s), with $0 \leq m \leq n-1$. Thus, the center has $n-1-m$ predecessors; if each of them shirks, then the center and all his successors shirk accordingly under the optimal reward scheme. Similarly, if the center shirks, then all his successors shirk as well. In contrast, the center's successors cannot trigger anyone to shirk because their actions are unobservable. Define $V^*(m)$ as the total rewards to the agents under the optimal reward scheme when the m most important peripheral agents move after the center. Thus, by Proposition 1, $V^*(m)$ is given by

$$V^*(m) = \underbrace{\sum_{i=1}^{n-1-m} \frac{1}{p(I) - p(\{j|j < n-m\}\backslash\{i\})}}_{\text{rewards to the predecessors}} + \underbrace{\frac{1}{p(I) - p(\{j|j < n-m\})}}_{\text{reward to the center}} + \underbrace{\sum_{i=n-m}^{n-1} \frac{1}{p(I) - p(I\backslash\{i\})}}_{\text{rewards to the successors}}.$$

To find the optimizer m^*, we compare $V^*(m)$ with $V^*(m+1)$; the difference between the two items is the marginal effect of an additional successor on the total rewards. By direct calculation, for any m with $0 \leq m \leq n-2$, we have

$$\begin{aligned} V^*(m+1) - V^*(m) = & \sum_{i=1}^{n-2-m} \frac{1}{p(I) - p(\{j|j < n-m-1\}\backslash\{i\})} \\ & - \sum_{i=1}^{n-2-m} \frac{1}{p(I) - p(\{j|j < n-m\}\backslash\{i\})} \\ & - \frac{1}{p(I) - p(\{j|j < n-m\})} + \frac{1}{p(I) - p(I\backslash\{n-m-1\})}. \end{aligned} \tag{2}$$

The sum of the first three terms on the RHS of (2) is the net change in the rewards to the

[5] This is because the relative orders between the center's predecessors or successors does not affect their incentive costs, as each individual's action is equally transparent for predecessors and successors, respectively.

center and his predecessors. Since p is increasing, this value is negative, i.e., by increasing the number of successors, the total rewards to the center and his predecessors decrease. The reason is two-fold: first, creating more successors reduces the number of the rest of the agents; more importantly, doing so makes the actions of the center and his predecessors more transparent, thereby enhancing the implicit threat of shirking for these agents and reducing the incentive costs. In this regard, we call these terms together the marginal benefit (MB) of more successors. Formally, we define

$$MB(m) := \sum_{i=1}^{n-2-m} \left[\frac{1}{p(I) - p(\{j|j < n-m\} \backslash \{i\})} - \frac{1}{p(I) - p(\{j|j < n-m-1\} \backslash \{i\})} \right] + \frac{1}{p(I) - p(\{j|j < n-m\})}.$$

In contrast, the last term on the RHS of (2) is positive, which is the extra reward to the new successor. Analogously, we call this term the marginal cost (MC) of more successors. Formally, we define,

$$MC(m) := \frac{1}{p(I) - p(I \backslash \{n-m-1\})}.$$

Note that $MC(m)$ is nondecreasing in m. This is because each new successor of the center is (weakly) less important than the current ones, meaning that his incentive cost is higher. It follows that if $MB(m)$ is decreasing in m, then there exists a unique optimizer m^* (either an interior solution or a corner solution). Lemma 4 below shows that under complementarity, the marginal benefit of more successors is indeed decreasing.

Lemma 4. *$MB(m)$ is decreasing in m, whereas $MC(m)$ is nondecreasing in m.*

Intuitively, as the center obtains more successors, his action becomes more transparent, so do his predecessors' actions. As argued previously, the effectiveness of action transparency on providing incentives is diminishing. Consequently, the reduction in payments decreases as the center obtains more successors. This effect is amplified by the decreasing number of predecessors. In addition, because each successor is more important than all the predecessors, a low effort of the center or his predecessor will trigger on average less important agents to shirk when there are more successors, meaning that the average implicit threat of shirking is weaker. In summary, the marginal benefit of more successors is diminishing.

Lemma 4 ensures that the optimal sequence is essentially unique and can be succinctly characterized by an integer m^* which is given by

$$m^* := \min\{m | MB(m) \leq MC(m)\}. \tag{3}$$

The next proposition shows that in the optimal sequence, the center never moves the first; if the center is sufficiently more important than any peripheral agent, then he moves the last.

Proposition 3. *If g is a star network, then the optimal mechanism $\{\pi^*, v^*\}$ satisfies that (i) the center has m^* successor(s), each of them is more important than the center's predecessors, where m^* is given by (3) with $0 \leq m^* \leq n-2$; (ii) v^* is given by Proposition 1 accordingly. Moreover, if $[p(I) - p(I \backslash \{n-1\})] < \delta\, [p(I) - p(I \backslash \{n\})]$ for some small $\delta > 0$, then $m^* = 0$, where agent $n-1$ is the most important peripheral agent and agent n is the center.*

The intuition of why the center never acts the first is straightforward. Suppose the center moves the first, then it is profitable to swap the center's position with that of a successor. This is because in the new sequence, the agent's action is as transparent as the center's old action, which will be observed by all the other agents, while the center's action is more transparent than the agent's old action which is unobservable. Because the rest of the agents obtain the same rewards, the total rewards are lower in the new sequence.

Proposition 3 states that if the center is sufficiently more important than the peripheral agents, in the sense that the center's shirking is much more detrimental to success, then he should act the last. Intuitively, if the center is significantly important, then his predecessors' incentive costs are relatively low due to the strong implicit threat of shirking. By contrast, the center's successors are free from the implicit threat and thus have relative high incentive costs.[6] This means that creating more successors will never be profitable. Thus, in this case the center will serve as the unique monitor of the team.

In particular, if all the agents are equally important, then $1 \leq m^* \leq n-2$; that is, the center always acts in an interior stage. This is because with identical importance, each agent i's payoff depends only on the cardinality of M_i, irrespective of his identity. By assigning the center to some interior stage, the mechanism allows the peripheral agents to learn their peers' actions through the center, as if the center served as an internal information intermediary through which the successors could monitor the predecessors. Such a layout maximizes the layers of informational hierarchy, thereby improving the transparency of actions.

[6]Indeed, for any $1 \leq m \leq n-2$, the incentive costs of the center and his predecessors' are bounded above by $[p(I) - p(I \backslash \{n\})]^{-1}$, whereas that of a center's successor is bounded below by $[p(I) - p(I \backslash \{n-1\})]^{-1}$. Given the assumption, the latter incentive cost is more than $1/\delta$ times of the former, for some small $\delta > 0$.

An algorithm to identify π^*. The previous analysis indicates that there exists a simple algorithm to find the optimal sequence for star networks. Specifically, starting from $m = 0$, allocate the peripheral agents to the set of the center's successors one by one from the most important to the least, until the first time when $MB(m) \leq MC(m)$ and therefore $m = m^*$. When the team size is relatively large, this algorithm can remarkably simplify the search of the optimal sequence. In addition, the relative orders among the predecessors or successors do not affect the total rewards, and thus, the principal has substantial flexibility in choosing the optimal mechanism.

As a comparative-statics analysis, we study the impacts of the importance of individual task on the optimal sequence for a star network. Specifically, we examine how the number of the center's successors in the optimal sequence varies with the importance of individual task. For ease of exposition, in the following, we consider a numerical example and assume that the agents are equally important to the project.

Example 1. Suppose that g is a star network with $n \geq 3$ agents, and that the project is a success if and only if all the tasks are successful. Each task is successful with probability 1 if the agent works, and is successful with probability $\alpha \in (0, 1)$ if the agent shirks. A lower probability α means that an agent's shirking is more detrimental in expectation to the project's success, and hence (effort on) his task is more important. Let w be the number of agents who work, then $p(w) = \alpha^{n-w}$, because all the tasks are independent. Clearly, p is increasing and satisfies complementarity. Applying the previous results, we express $MB(m)$ and $MC(m)$ explicitly in the following:

$$MB(m; \alpha) = \frac{n-m-2}{1-\alpha^{m+2}} - \frac{n-m-2}{1-\alpha^{m+3}} + \frac{1}{1-\alpha^{m+1}} \quad \text{and} \quad MC(m; \alpha) = \frac{1}{1-\alpha}.$$

It follows that for a fixed $\alpha \in (0, 1)$, $MB(m)$ is decreasing in m, and that $MB(0) > MC(0)$ and $MB(n-2) < MC(n-2)$. Thus, the optimizer m^* exists and is an interior solution for any $\alpha \in (0, 1)$. In addition, from basic mathematical analysis, we have that for a fixed m, both $MB(\alpha)$ and $MC(\alpha)$ are increasing and strictly convex in α, and that $MB(\alpha)$ is single-crossing $MC(\alpha)$ from below in the domain $\alpha \in (0, 1)$. This is illustrated in Figure 1. It thus follows that the optimizer $m^*(\alpha)$ is nondecreasing in α.[7] To summarize,

Corollary 3. *In the optimal sequence of Example 1, the number of the center's successors $m^*(\alpha)$ is nondecreasing in α for $\alpha \in (0, 1)$.*

[7] Since m is an integer, $m^*(\alpha)$ is not necessarily increasing in α.

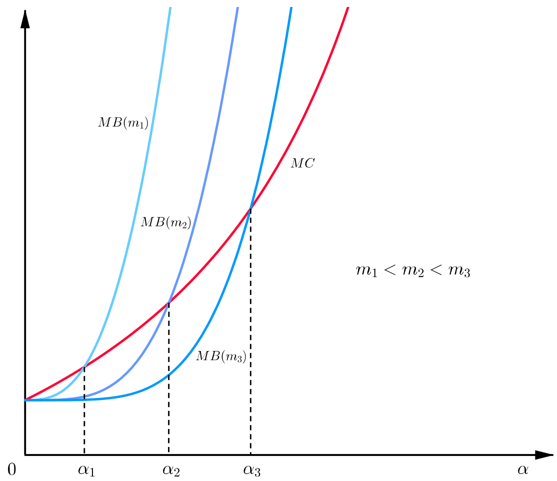

Figure 1: Marginal Benefit and Marginal Cost as a Function of Importance.

Corollary 3 states that the more important each task is, the fewer successors the center obtains. Intuitively, if each task is important to the project, then each agent has a relatively strong incentive to work, and thus, the implicit threat of shirking is not critical in providing incentives. This means that improving the transparency of actions by creating more successors is not efficient in reducing incentive costs. In contrast, if each task has little effects on the project, then it may be profitable to increase the center's successors.

5 Incentive Design for Collaborating Teams

In this section, we take our analysis further to consider multiple agile teams collaborating to finish a joint project. Alternatively, one may interpret the collaboration as organizing a large agile team within which multiple sub-teams are each responsible for a different sub-project.

Typically, such collaboration represents a hierarchy in both connection and information: The scrum masters (leaders) of individual teams are in charge of cross-team coordination, and thus have an advantage over their team members in observation and inference of other teams' performances. For example, when multiple research units collaborate on a large scale research project, the directors of the research units may observe not only the performances of their colleagues in own groups, but also those of other teams through directors meetings. Therefore, we model the organization structure in this environment by a network composed of multiple stars, with all the star centers connected by a complete sub-network, i.e., there is a link between each pair of centers. Figure 2 illustrates such a network. Let $t \geq 2$ be the number of stars in g, and $\hat{n}_i \geq 3$ be the number of agents in star $i \in \{1, \dots, t\}$. Also, let c_i be the center of star i, and $C \coloneqq \{c_i\} \subset I$ be the set of the centers.

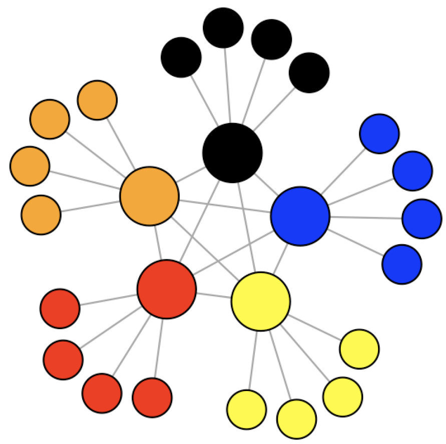

Figure 2: Network among Collaborating Teams.

To solve for the optimal sequence under such networks, note that it suffices to determine the order between the centers, and the order of each peripheral agent relative to the center of his own star. Given a sequence π, denote by π^C the sub-sequence confined to the set C. Because all the centers are interconnected, by Lemma 1 the centers must move sequentially within C, i.e., $\pi_{c_i} \neq \pi_{c_j}$ for any two stars $i \neq j$. We say that star i moves before star j if c_i moves before c_j, and that an agent moves before (after) star i if he belongs to a star that moves before (after) star i. In terms of the sequence within each star, provided there is no confusion, we call a peripheral agent a "predecessor" if he moves before his own center, and a "successor" if he moves after the center. Clearly, in the optimal sequence, every peripheral agent is either a predecessor or a successor. Analogous to the single team, let $m_i \leq \hat{n}_i - 1$ be the number of c_i's successors within star i.

Since the network now consists of multiple layers of connection, the internal information of such an organization can be hierarchical as well. For example, a change in the number of successors in star i will affect not only the action transparency of the center c_i and all the predecessors in star i, but also that of each agent in the preceding stars, whose action c_i can ultimately learn. Consequently, sequencing in such networks will exhibit a similar cross-star effect on the agents' incentives.

Moreover, since the teams may be different in both size and importance, it is remarkably difficult to pin down the optimal sequence when both kinds of heterogeneity are present. To distinguish between the effects of team size and that of team importance on the optimal mechanism, in the following we study each effect respectively by fixing the other. Since the optimal reward scheme follows immediately from Proposition 1 once the optimal sequence is given, in the subsequent results we omit the optimal reward scheme unless necessary.

5.1 Optimal Mechanisms with Homogeneous Agent Importance

We first study the optimal mechanism under heterogeneous team sizes while assuming that all the agents are equally important to the project. In this particular case, the probability of success can be simply written as $p(w)$, where w is the number of working agents.

To characterize the optimal sequence, it is instructive to define analogously the marginal benefit and the marginal cost of more successors m_i in each star i, for a fixed sub-sequence of the centers π^C. As argued previously, the marginal benefit stems from the improvement in transparency of all the actions that cannot be learned by the new successor until now, whereas the marginal cost is simply the reward to the new successor. Since the agents are equally important, the optimal reward to a successor is $[p(n) - p(n-1)]^{-1}$ by Proposition 1. That is, the marginal cost is constant across stars. In contrast, the marginal benefit exhibits the aforementioned cross-star effect, i.e., increasing m_i reduces the rewards of not only c_i and all the predecessors in star i, but also all the centers and predecessors in the preceding stars. Clearly, *ceteris paribus*, the marginal benefit is higher for a star i in a later stage, as increasing m_i can improve the transparency of more actions. Based on these observations, we show a key property of the relative orders of stars in the lemma below:

Lemma 5. *Assume homogeneous agent importance. Suppose in π^* c_i moves before c_j, then (i) if $m_i > 0$, then $m_j = \hat{n}_j - 1$; (ii) if $m_j < \hat{n}_j - 1$, then $m_i = 0$.*

Lemma 5 states that if a star has a nonempty set of successors, then in every star that moves later all the peripheral agents are successors. In contrast, if a star has a nonempty set of predecessors, then in each star that moves earlier all the peripheral agents are predecessors. Intuitively, if it is profitable to make a peripheral agent a successor, then it is more profitable to make every peripheral agent a successor in each star that moves later, since the marginal benefit is greater in the latter. On the other hand, if it is unprofitable to make a peripheral agent a successor, then it is never profitable to have successors in a star that moves earlier. Lemma 5 implies that in the optimal sequence, there exists at most one star that has both predecessors and successors. Thus, the teams can be allocated basically into two groups in the optimal sequence: toward the head of the sequence, in each team every peripheral agent is a predecessor; toward the end of the sequence, every peripheral agent is a successor.

In addition, we can show that in the first group teams are sequenced in descending order of size, whereas in the second group teams are sequenced in ascending order. Therefore, the optimal sequence exhibits a "V-shape" with respect to team size. Specifically,

Proposition 4. *Assume homogeneous agent importance. The optimal sequence satisfies that for any two stars i and j such that i moves before j, (i) if $m_i = m_j = 0$, then $\hat{n}_i \geq \hat{n}_j$; (ii) if $m_i = \hat{n}_i - 1$, $m_j = \hat{n}_j - 1$, then $\hat{n}_i \leq \hat{n}_j$.*

The optimal sequence exhibiting "V-shape" results from the dual role of a center in the internal information. According to Lemma 5, the action of each center and each predecessor will be eventually learned, or simply *monitored*, by all the subsequent centers and successors. Thus, by allocating larger size teams toward both ends of the sequence, the principal allows the centers to monitor more predecessors, while in the meantime to be monitored by more successors, leading to more transparent actions overall. As in a single team, the centers serve as information intermediaries of the organization; as such, the "V-shape" layout improves peer monitoring by facilitating the flow of internal information through the centers.

Whereas Proposition 4 provides only a partial characterization of the optimal sequence, it rules out most sub-optimal sequences. Specifically, given the number of stars t, there are totally $t!$ possible permutations of the stars. However, the number of the permutations that satisfy Proposition 4 is around $\sum_{k=0}^{t} C(t,k) = 2^t$,[8] which is of lower order of $t!$ for large t. Moreover, one could apply the following algorithm to search the optimal sequence based on Proposition 4. With a bit abuse of notation, we write MB and MC for the marginal benefit and the marginal cost of increasing m_i, respectively. Recall that MB equals the reduction in rewards to the agents whose action could not be learned by the new successor before, which is a function of the total number of current successors, $\sum m_i$, and that MC is the reward to the new successor, which is fixed at $[p(n) - p(n-1)]^{-1}$. Then, we have:

Corollary 4. *Assume homogeneous agent importance. Suppose the optimal sub-sequence π^C has been determined, then π^* can be fully characterized by the following algorithm:*

1. *Set each star i such that $m_i = 0$, $i \in \{1, \dots, t\}$;*
2. *From the last star to the first star in π^C, increase m_i one by one until*

$$MB\left(\sum m_i\right) \leq MC.$$

Similar to the single team case, the algorithm in Corollary 4 is derived from the idea that the marginal benefit is decreasing in the total number of successors $\sum m_i$. This is again due

[8] By Lemma 5, we partition the stars into two groups, in the first each peripheral agent is a predecessor, whereas in the second each peripheral agent is a successor. Let k be the number of stars in the first group. Then, the sub-sequence π^C is given by Proposition 4. The number of possible permutations is thus 2^t.

to the diminishing marginal effect of action transparency on incentives. In particular, for t which is not large, the optimal sequence can be easily searched by a two-step algorithm: for each sub-sequence π^C that satisfies Proposition 4, apply the algorithm in Corollary 4; then, select the sub-sequence that yields the least total rewards in step one.

We have demonstrated the efficiency of Proposition 4 in searching the optimal sequence. The proposition below can take this one step further by offering upper and lower bounds for the total number of successors in the optimal sequence. Formally, we have:

Proposition 5. *Assume homogeneous agent importance. Suppose π^* contains totally $\sum m_i^*$ successors. Then, apply the following algorithm:*

1. Align the stars in ascending order of $\hat{n}_i$, $i \in \{1, \ldots, t\}$;

2. Apply Corollary 4 to this π^C, and let $\bar{m}$ be the associated total number of successors;

1' Align the stars in descending order of $\hat{n}_i$, $i \in \{1, \ldots, t\}$;

2' Apply Corollary 4 to this π^C, and let $\underline{m}$ be the associated total number of successors.

It follows that $\underline{m} \leq \sum m_i^* \leq \bar{m}$.

Proposition 5 offers an additional restriction for the set of stars that contain successors. Because there is at most one star that contains both predecessors and successors, the total number of peripheral agents within such a set is approximately bounded by $\underline{m}$ and $\bar{m}$. This narrows down the possible sets of stars that contain successors. In particular, if $\underline{m}$ and $\bar{m}$ are close to each other, then there are relatively few possible permutations.

The idea of Proposition 5 is again attributed to the fact that the effectiveness of action transparency on incentives is diminishing. Intuitively, by sequencing the stars in ascending order of size, the principal assigns as many peripheral agents as possible to later stages, and thus, the agents' actions are less transparent than they would be in any other sub-sequence π^C given a fixed number of successors. This means that the marginal benefit of increasing m_i is always higher in this sub-sequence than in any other π^C, leading to a larger total number of successors.[9] The intuition for the lower bound is analogous.

Presumably, for a large-scale organization, i.e., when both n and t are large, the number of successors $\sum m_i^*$ might also be a large number. This could lead to an enormous amount of possible sets of stars that contain successors, making the previous algorithm impractical.

[9] This does not necessarily mean that the sub-sequence with ascending order of size leads to the optimum. Note that the notion of marginal benefit is applicable to a given sub-sequence. However, one could improve the sequence by altering the sub-sequence.

However, the next result shows that under mild conditions—the following Assumption 1 for example— $\sum m_i^*$ is of lower order of n; that is, the set of successors amounts for only a small fraction of the whole population of the organization.

Assumption 1. *For any positive integer $m \le n-1$, we have*

$$\frac{1}{m}[p(n-1) - p(n-m-1)] \ge K[p(n) - p(n-1)]$$

for some constant $K > 0$.

Assumption 1 states that the marginal productivity of the last piece of effort cannot be infinitely higher than the average productivity of any preceding efforts, however large n is; otherwise, the marginal cost of increasing m_i may be so low that almost all the peripheral agents are successors. Given Assumption 1, we have the following proposition:

Proposition 6. *Assume homogeneous agent importance. If Assumption 1 holds, then the optimal sequence satisfies that $\sum m_i^*$ is bounded above by some number of order $\sqrt{n}$.*

The proof of Proposition 6 can be sketched in two steps. First, we shall show that given any number of successors, the marginal benefit in multiple stars is lower than that in a single star with a same n, if in the former case the number of predecessors plus the number of stars that have no successors is equal to the number of predecessors in the latter case. Intuitively, because the multi-star structure has more tiers in connection, the agents' actions are more transparent than in a single star. This implies that the marginal benefit is lower, and thus, there are fewer successors in the multi-star case. Second, we shall prove that the number of successors in the optimal sequence of a single star is of order $\sqrt{n}$ given Assumption 1. This is because on the one hand a successor has a relatively high reward due to his unobservable action; on the other hand, the effectiveness of action transparency is diminishing. Thus, the marginal benefit of increasing successors will be soon outweighed by the marginal cost.

We have shown that when the agents are equally important, the internal information of the organization has a hierarchical structure. The peripheral agents of teams in later stages serve as the monitors of the organization, whereas those of teams in earlier stages are simply monitored by their peers, and the centers serve as information intermediaries through which the internal information flow. Moreover, Proposition 6 states that the monitors will account for only a small fraction of the population. The idea is that while peer-monitoring can help reduce the agents' incentive costs, the monitors themselves are costly to incentivize because then no one will monitor the monitor.

5.2 Optimal Mechanisms with Homogeneous Team Size

Now we turn to the organization with heterogeneous agent importance, while assuming that every team has the same size $\hat{n} \geq 3$. Note that the optimal sequence depends on both the relative orders between teams and the internal sequence within each team. Accordingly, the optimal sequence will be affected by not only the importance rank between the agents but also that between teams. In this regard, we define the importance of each team as follows. First, for each team i, we denote the k-th important peripheral agent r_i^k, $k \leq \hat{n}-1$. Then, for any two teams i and j, we say that star i is more important than star j if (i) c_i is more important than c_j and (ii) for all $k = 1, \cdots, \hat{n}-1$, r_i^k is more important than r_j^k; i.e., every agent of star i is more important than his counterpart in star j. We assume that the teams are totally ordered with respect to importance. Moreover, we assume that the subsets of I are totally ordered in importance given the definition in Section 2.

Note that the multi-star network is essentially a combination of the two previous types of networks: each team is a star on a micro level, while a node of a complete network from a macro perspective. Thus intuitively, the optimal sequence should exhibit properties of the optimal sequences of star and complete networks. That is, the principal would conceivably benefit from making more important agents successors within each star, and letting more important stars move after less important ones in the entire sequence. It turns out that such a conjecture is valid under certain conditions, for example, the assumption below.

Assumption 2. *For any subsets of agents S, S' and T such that S is more important than S' and $T \cap (S \cup S') = \emptyset$, we have*

$$p(S \cup T) - p(S) \geq p(S' \cup T) - p(S').$$

Assumption 2 states that the organization's technology satisfies the complementarity in importance; that is, the efforts of any set of agents are more productive when the set of the other working agents becomes more important. It generalizes the notion of complementarity defined in Section 2, from comparing the marginal effect of adding a single agent to two sets ordered by inclusion, to comparing that of adding an agent set to two sets ordered by importance. Then, we have the following result:

Proposition 7. *Assume homogeneous team size. The optimal sequence π^* satisfies that (i) in every star, each successor is more important than each predecessor; (ii) each successor in star i is more important than every predecessor in each star that moves after i. Furthermore,* *if Assumption 2 holds, then (iii) the stars move in ascending order of importance; thus,* m_i^* *is nondecreasing along* π^*.

The idea of Proposition 7 is as follows. Property (i) is simply a corollary of Lemma 2. Property (ii) extends the above cross-team effect to heterogeneous team importance: If for some star i it is profitable to make a peripheral agent, say r_i^k, a successor, then for each star that moves later, it is more profitable to make every peripheral agent who is more important than r_i^k a successor. This is because a more important successor in a later team, who incurs a lower incentive cost, can impose greater implicit threats on more agents. These results are extensions of those of complete and star networks, which do not rely on Assumption 2.

However, the property of monotone importance in a single team does not directly apply to the sequence among multiple teams. In both complete and star networks, a complete importance ranking on *individual* agents guarantees that the principal benefits from switching a more important agent with a less important one, should the former move earlier than the latter in a complete network, or the former be a predecessor whereas the latter a successor in a star. In contrast, in the multi-star network, switching the orders of two teams means altering simultaneously the *set* of monitoring agents and the *set* of agents being monitored. For instance, consider two different stars with the same number of successors, one of which moves before the other. Even if the principal simply swap their orders without changing the number or successors, it affects the reward needed for each predecessor of each star, along with the centers, in a way that individual importance ranking alone is not able to capture with regularity. However, once the importance ranking and complementarity of importance extend naturally to all sets of agents (Assumption 2), the marginal effect of agent importance on incentives becomes uniformly diminishing. That is, the reduction in incentive cost for one agent decreases as those who can learn his action become more important. Thus, it is more profitable to let the less important star move first so that its predecessors are monitored by a more important set of agents.

It is worth noting that the diminishing marginal effect of action transparency is indeed a particular case of that of agent importance. Essentially, an action being more transparent means that the action can be learned by a more important (in terms of size) set of agents. In other words, a more transparent action is just a particular type of more influential actions. Because of the diminishing marginal effect of action influence, in both the single-team and multi-team cases, the internal information exhibits an optimal balance between the intensive margin effect (i.e., making one action more influential) and the extensive margin effect (i.e., making more actions influential).

One may wonder if there exists a simple algorithm, similar to Corollary 4, to search for the optimal sequence under heterogeneous agent importance. The difficulty is that we no longer have the monotonicity of the marginal benefit and the marginal cost of increasing a star's successors, even under Assumption 2. For example, it is ambiguous whether making a less important peripheral agent in a later star a successor is more profitable than making a more important peripheral agent in a earlier star a successor. Thus, it is very likely that a simple algorithm to characterize the optimal sequence is only available under certain parametric settings. The corollary below characterizes a typical case and the associated algorithm.

Corollary 5. *Assume homogeneous team size and given Assumption 2. Suppose for any two stars i and j such that i is more important than j, $r_i^{\hat{n}-1}$ is more important than r_j^1, then π^* can be fully characterized by the following algorithm:*

1. *Align the stars in ascending order of importance, and set each star i such that $m_i = 0$;*
2. *In the last star, make the most important peripheral agent a successor, then the second important peripheral agent, etc. When every peripheral agent becomes successor, move backward one star and continue this procedure;*
3. *Continue until the total rewards start to increase for the first time.*

The algorithm is a straightforward variation of that for a single team. The key to it being effective is that the importance rank should be not only local in each single star, but also global across stars, which makes the net marginal benefit of locally rearranging agents—in a one-by-one fashion as specified in the algorithm—strictly monotone. Naturally, this criterion is met in the two extreme cases: when agents are homogeneous so that they can be ranked in any arbitrary order, which has been shown previously, and when agents are sufficiently different so that the importance rank among stars is always consistent with the rank among individual agents from different stars. A resultant feature of the optimal sequence in these cases is that there will be at most one star with both predecessors and successors; all stars moving before it have only predecessors and those after have only successors.

5.3 "Size Effect" vs "Importance Effect"

In the previous part, we established a unified concept of action influence. An action is more influential if it can be learned by a set of more important agents. However, the importance may stem from different sources, having qualitatively different implications on the incentive design for multi-team collaborations. Specifically, a "more important" team may be placed either before or after a "less important" one, depending on what actually underlies the notion of team importance. When a team is more important because of its members' individual capability (Section 5.2), more important teams unanimously move later than less important ones. In contrast, when a team is more important due to larger size (i.e., more members, Section 5.1), more important teams are located toward both ends of the optimal sequence.

The difference here reflects the interplay between two countervailing effects. The first is a "size effect", meaning how large a fraction a team takes up in the total number of agents; the second is an "importance effect", meaning how much a team's (aggregate) effort influences the success probability of the whole project. A larger size effect might induce the principal to make the team move earlier and make its peripheral agents predecessors of the center, so that more incentive costs can be saved by paying a low reward per agent to more agents. A larger importance effect, on the other hand, would lead the principal to make the team move later and make its peripheral agents successors of the center, in order to lower incentive costs for earlier movers. The second proposition here essentially gives rise to our results on single-team incentive design.

In the multi-team environment, a more important team given size homogeneity exhibits a pure importance effect. However, a larger team given agent importance homogeneity has both a larger size effect and a larger importance effect. In addition, we have shown that the importance effect diminishes when more teams move in the fashion of all peripheral agents being successors, which means that the size effect will ultimately dominate given a sufficiently large total number of agents. As a practical implication from our theory, what it means to be a more important team can be two-fold and the principal should examine the notion closely with care before assigning tasks sequentially. The source of the team's importance is a non-negligible factor in designing optimal incentives.

6 Conclusion

In this paper, we proposed a tractable framework to study an incentive design problem in a team where members have access to private internal information about each other's effort level. The feasible information architecture is described by an exogenous network, while the principal may exploit this architecture to minimize total rewards needed by endogenously determining the sequence of task assignment. We find that, both for a single team and for collaborating teams of the same size, the optimal sequence exhibits a uniform feature of delayed assignment for more important agents. However, the effect of heterogeneous team sizes is not monotone: larger teams may be assigned tasks earlier or later depending on their importance ranking.

Internal information within teams remains an intriguing and promising topic in both theoretical and empirical economics, and related fields such as operations management and organization science. Besides hidden action as studied in this paper, the issue of internal information transparency may also arise with agents' private types, knowledge, evolution and updating, etc. We expect richer further studies to be conducted, with our work as part of the groundwork, on revealing relations between the nature of internal information and optimal incentive design in a more general and flexible strategic environment.

A Appendix

A.1 Proofs

Proof of Proposition 1

Proof. Suppose that p satisfies complementarity. We first prove that $\{\pi, v^*(\pi)\}$ is an EFI mechanism. Consider a strategy profile s^* such that $s_i^* = 1$ if and only if $a_j = 1$ for all $j \in N_i$ or N_i is empty; that is, an agent works unless he sees someone shirking. This strategy profile can be sustained by a PBE with the following set of beliefs: if $a_j = 1$ for all $j \in N_i$ or $N_i = \emptyset$, then $a_k = 1$ for all $k \in I \backslash (N_i \cup \{i\} \cup M_i)$; that is, an agent, not seeing anyone shirking, believes that those whom he cannot see and who cannot see him through a sequence of agents will exert effort. To verify this statement, note that if agent i shirks, then by induction every $j \in M_i$ shirks as well. In contrast, if i works then he believes that all the other agents work too unless he sees someone shirking. Suppose i is the first to act, then he believes that if he works then all the other agents also work, and if he shirks then he will induce each agent in M_i to shirk. Thus, i prefers working to shirking if and only if the difference in expected reward exceeds the effort cost, i.e.,

$$[p(I) - p(I \backslash (\{i\} \cup M_i))]v_i \geq 1. \tag{A.1}$$

Clearly, v_i^* satisfies (A.1). It follows by induction that for all $\pi_i \in \{2, \ldots, n\}$, i prefers to work on equilibrium path if and only if (A.1) holds, as he sees no one shirking. Off the path, if i sees a nonempty subset $S_i \subseteq N_i$ of agents shirking, then he knows that each $j \in S_i$ will induce everyone in M_j to shirk. Let $R_i \coloneqq \bigcup_{j \in S_i} M_j \cup S_i$ be the set of agents whom i believes to shirk. Thus, if i works then his expected utility equals $p(I \backslash R_i)v_i^* - 1$. In contrast, if i shirks then his expected utility equals $p((I \backslash R_i) \backslash (\{i\} \cup M_i))v_i^*$. We now provide a useful lemma.

Lemma A.1. *Suppose p satisfies complementarity, then for any two nonempty sets of agents B, $C \subset I$, we have $p(I) - p(I \backslash B) > p(I \backslash C) - p((I \backslash C) \backslash B)$.*

Proof. For two nonempty sets T and S with $T \subset S$ and two agents i, $j \notin S$, we have

$$\begin{aligned} p(S \cup \{i\} \cup \{j\}) - p(S) &= p(S \cup \{i\} \cup \{j\}) - p(S \cup \{i\}) + p(S \cup \{i\}) - p(S) \\ &> p(T \cup \{i\} \cup \{j\}) - p(T \cup \{i\}) + p(T \cup \{i\}) - p(T) \\ &= p(T \cup \{i\} \cup \{j\}) - p(T). \end{aligned}$$

This implies by induction that for any nonempty set $Q \subset I$ with $Q \cap S = \emptyset$ we have

$$p(S \cup Q) - p(S) > p(T \cup Q) - p(T). \tag{A.2}$$

Then, let $T = (I \backslash C) \backslash B$, $S = (I \backslash B)$, and $Q = B$. It is obvious that $T \subset S$ and $Q \cap S = \emptyset$. Then, the lemma follows immediately from (A.2). □

From Lemma A.1, we conclude that $[p(I \backslash R_i) - p((I \backslash R_i) \backslash (\{i\} \cup M_i))]v_i^* < 1$. This means that i prefers to shirk whenever he sees someone shirking. Hence, s^* and the set of beliefs that are constructed above indeed constitute a PBE with full efforts.

It remains to show that any alternative reward scheme v' with $v_i' < v_i^*$ cannot constitute a PBE with full efforts. Suppose not, then the probability of success is $p(I)$ on the equilibrium path. If i shirks unilaterally, then he can at most trigger those in M_i to shirk, irrespective of the strategy profile. In other words, i's effort externality is confined to the coalition M_i. Because p is increasing, the difference in expected reward is less than the effort cost. Thus, i can make a profitable deviation by shirking, leading to a contradiction. Note that all these arguments go through for any fixed π, thus statement (1) is proven.

Suppose that p satisfies substitutability. As before, we first prove that $\{\pi, v^*(\pi)\}$ is an EFI mechanism. Consider a strategy profile s^* with $s_i^* \equiv 1$, that is, an agent always exerts effort irrespective of his information set. This strategy profile can be sustained by a PBE with the set of beliefs that $a_j = 1$ for all $j \notin N_i$; that is, an agent believes that those whom he cannot see will exert effort. Note that if agent i sees no one shirking then he believes that all the other agents work. Hence, he prefers to work if and only if $[p(I) - p(I \backslash \{i\})]v_i \geq 1$, which holds for v_i^*. In contrast, if i sees a nonempty subset of agents $S_i \subseteq N_i$ who shirk, then his expected utility equals $p(I \backslash S_i)v_i^* - 1$ if he works, and $p((I \backslash S_i) \backslash \{i\})v_i^*$ if he shirks. Then by substitutability, we have $p(I \backslash S_i) - p((I \backslash S_i) \backslash \{i\}) \geq p(I) - p(I \backslash \{i\})$. This implies that i still prefers to exert effort. Thus, s^* and the set of beliefs constitute a PBE. Finally, we argue that there does not exist a reward scheme v' with $v_i' < v_i^*$ that admits a PBE with full efforts. Suppose not, then i must prefer working to shirking if he encounters no shirking. Due to substitutability, if i shirks unilaterally then each $j \in M_i$ prefers to work, as argued above. It follows that the difference in expected reward equals $p(I) - p(I \backslash \{i\})$. Because i is indifferent under v_i^*, he must prefer shirking under v', a contradiction. Thus, $v^*(\pi)$ is indeed optimal. In summary, the proposition is proven. □

Proof of Lemma 1

Proof. Suppose not, then $\pi_i^* = \pi_j^*$. Consider a new sequence π' which differs from π^* only in that j acts in π' immediately after i and before all the agents who act after i in π^*; thus, $\pi_j' > \pi_j^*$ and $\pi_k' = \pi_k^*$ for any agent $k \neq j$. It thus follows that $N_j^* \subset N_j'$ and $M_j' = M_j^*$. Consider an agent $k \neq j$. Clearly, if $\pi_k^* > \pi_j^*$, then $M_k' = M_k^*$. If $\pi_k^* \leq \pi_j^*$, then we partition M_k^* into two groups: $M_{k\backslash j}^*$ and $M_k^* \backslash M_{k\backslash j}^*$, where $M_{k\backslash j}^*$ is the set of agents who will remain in M_k^* if all j's links are eliminated and the agents act in the order of π^*. Pick any agent $l \in M_k^*$. If $l \in M_{k\backslash j}^*$, then clearly he will remain in M_k' under π'. If $l \in M_k^* \backslash M_{k\backslash j}^*$, it must be that $l \in M_j^*$. Since $N_j^* \subset N_j'$ and $M_j' = M_j^*$, l will still remain in M_k', and thus, $M_k^* \subseteq M_k'$. In summary, for any agent $k \in I$, we have $M_k^* \subseteq M_k'$, meaning that $v_k^*(\pi') \leq v_k^*(\pi^*)$ due to Proposition 1. But because $ij \in g$ and $\pi_i^* = \pi_j^*$, we have $M_i^* \subset M_i'$; thus, $v_i^*(\pi') < v_i^*(\pi^*)$. This means that the total payoffs to the agents are strictly lower under π' than under π^*, leading to a contradiction. Thus, the lemma is proven. □

Proof of Lemma 2

Proof. Because we assume that either $i \in M_j^*$ or $j \in M_i^*$, both i's and j's neighbors are nonempty. Thus, there are two cases to consider. First, suppose $ij \notin g$, then there exists some agent $l \neq i, j$ such that $il, jl \in g$ and l acts between i and j. Suppose $\pi_i^* < \pi_l^* < \pi_j^*$, then $j, l \in M_i^*$. Now swap i and j and denote the new sequence π'. Therefore, $i, l \in M_j'$. Because $\{k|ik \in g, k \neq j\} = \{k|jk \in g, k \neq i\}$, we have $N_i' = N_j^*$, $N_j' = N_i^*$, $M_i' = M_j^*$ and $M_j' \cup \{j\} = M_i^* \cup \{i\}$. Thus, for any agent $k \neq i, j$, there are three possibilities. First, $i, j \notin M_k^*$. Because $N_i' = N_j^*$ and $N_j' = N_i^*$, we have $M_k' = M_k^*$. Proposition 1 thus implies that $v_k^*(\pi') = v_k^*(\pi^*)$. Second, $i \in M_k^*$. In this case, $j \in M_k^*$ as $j \in M_i^*$. Since $M_i' = M_j^*$ and $M_j' \cup \{j\} = M_i^* \cup \{i\}$, we have $M_k' = M_k^*$, and thus, $v_k^*(\pi') = v_k^*(\pi^*)$. Third, $j \in M_k^*$ and $i \notin M_k^*$. It follows that there exists an agent $k' \in M_k^*$ with $ik', jk' \in g$. Then by Lemma 1, we have $\pi_i^* < \pi_{k'}^* < \pi_j^*$. It follows that $M_k^* \backslash \{j\} = M_k' \backslash \{i\}$, and thus, we have

$$
\begin{aligned}
p(I\backslash(\{k\} \cup M_k^*)) &= p(I\backslash(\{k\} \cup (M_k^*\backslash\{j\}) \cup \{j\})) \\
&= p((I\backslash(\{k\} \cup (M_k^*\backslash\{j\}))\backslash\{j\}) \\
&= p((I\backslash(\{k\} \cup (M_k'\backslash\{i\}))\backslash\{j\}) \\
&> p((I\backslash(\{k\} \cup (M_k'\backslash\{i\}))\backslash\{i\}) \\
&= p(I\backslash(\{k\} \cup (M_k'\backslash\{i\}) \cup \{i\})) = p(I\backslash(\{k\} \cup M_k')).
\end{aligned}
$$

The inequality above is due to that i is more important than j. Then, from Proposition 1, we have $v_k^*(\pi') < v_k^*(\pi^*)$. Moreover, because $M_i' = M_j^*$, we have

$$p(I\backslash(\{j\} \cup M_j^*)) = p((I\backslash M_j^*)\backslash\{j\}) > p((I\backslash M_j^*)\backslash\{i\}) = p(I\backslash(\{i\} \cup M_i')).$$

It follows from Proposition 1 that $v_i^*(\pi') < v_j^*(\pi^*)$. Finally, since $M_j' \cup \{j\} = M_i^* \cup \{i\}$, we have $v_j^*(\pi') = v_i^*(\pi^*)$. In summary, the total payoff is strictly lower under π' than under π^*, leading to a contradiction. Thus, $\pi_j^* < \pi_l^* < \pi_i^*$.

Second, suppose $ij \in g$, then from Lemma 1, $\pi_i^* \neq \pi_j^*$; as such, either $i \in N_j^*$ or $j \in N_i^*$. Suppose $\pi_i^* < \pi_j^*$. Again, swap i and j and denote the new sequence π'. Note that we now have $N_i' \cup \{i\} = N_j^* \cup \{j\}$, $N_j' = N_i^*$, $M_i' = M_j^*$ and $M_j' \cup \{j\} = M_i^* \cup \{i\}$. Analogously, the above argument goes through in this case; thus, we have $\pi_j^* < \pi_i^*$. To summarize, if in π^* either $i \in M_j^*$ or $j \in M_i^*$, then $\pi_j^* < \pi_i^*$. Thus, the lemma is proven. □

Proof of Proposition 2

Proof. Because g is a complete network, for any two agents i and j, $\{k|ik \in g, k \neq j\} = \{k|jk \in g, k \neq i\}$. Moreover, by Lemma 1, the agents act sequentially in π^*. Thus, we have either $i \in M_j^*$ or $j \in M_i^*$ for any $i \neq j$. It follows from Lemma 2 that if i is more important than j, then $\pi_i^* > \pi_j^*$. By induction, we have that π^* is identity permutation. The optimal reward scheme v^* thus follows immediately from Proposition 1. In particular, if i and j are equally important, then swapping i and j in π^* will not change any agent's incentive cost, thereby maintaining the optimal sequence. Thus, the proposition is proven. □

Proof of Corollary 1

Proof. The corollary follows immediately from Proposition 2 and formula (1). □

Proof of Corollary 2

Proof. Let g_1 be a complete network and g_2 be an arbitrary network with the identical set of vertices as g_1. Suppose $\pi^*(g_2)$ is the optimal sequence for g_2. Consider a sequence $\pi(g_1)$ for g_1 such that each agent has the same order in $\pi(g_1)$ as in $\pi^*(g_2)$. By induction, one can easily show that for each agent i, $M_i(\pi^*(g_2)) \subseteq M_i(\pi(g_1))$. Then by Proposition 1, we have $v_i^*(\pi(g_1)) \leq v_i^*(\pi^*(g_2))$. Because $\pi(g_1)$ is not necessarily optimal, the optimal total payoffs must be (weakly) lower under g_1 than under g_2. Thus, the corollary is proven. □

Proof of Lemma 4

Proof. Define $\Delta MB(m) := MB(m+1) - MB(m)$. By direct calculation, we have

$$\begin{aligned}
\Delta MB(m) = &\sum_{i=1}^{n-3-m}\left[\frac{1}{p(I) - p(\{j|j<n-m-1\}\backslash\{i\})} - \frac{1}{p(I) - p(\{j|j<n-m-2\}\backslash\{i\})}\right] \\
&- \sum_{i=1}^{n-3-m}\left[\frac{1}{p(I) - p(\{j|j<n-m\}\backslash\{i\})} - \frac{1}{p(I) - p(\{j|j<n-m-1\}\backslash\{i\})}\right] \\
&+ \left[\frac{1}{p(I) - p(\{j|j<n-m-1\})} - \frac{1}{p(I) - p(\{j|j<n-m\})}\right] \\
&+ \left[\frac{1}{p(I) - p(\{j|j<n-m-2\})} - \frac{1}{p(I) - p(\{j|j<n-m\}\backslash\{n-m-2\})}\right].
\end{aligned}$$

Note that for any fixed i, the difference

$$\frac{1}{p(I) - p(\{j|j<n-m\}\backslash\{i\})} - \frac{1}{p(I) - p(\{j|j<n-m-1\}\backslash\{i\})}$$

is decreasing in m. This follows from the discussion of formula (1) in the text. Therefore, the difference between the above two summations is negative. Moreover, the values of the third and forth bracket in the expression of $\Delta MB(m)$ are both negative because p is increasing. It follows that $\Delta MB(m)$ is negative, so $MB(m)$ is decreasing. That $MC(m)$ is nondecreasing follows from the argument in the text. Thus, the lemma is proven. $\square$

Proof of Proposition 3

Proof. The optimal sequence follows immediately from Lemmas 3 and 4. Then, the optimal reward scheme is given by Proposition 1. To see that $m^* \leq n-2$, note that

$$MB(n-2) = \frac{1}{p(I) - p(\{1\})} < \frac{1}{p(I) - p(I\backslash\{1\})} = MC(n-2).$$

To prove the last statement of the proposition, note that

$$\begin{aligned}
MB(0) &= \sum_{i=1}^{n-2}\left[\frac{1}{p(I) - p(\{j|j<n\}\backslash\{i\})} - \frac{1}{p(I) - p(\{j|j<n-1\}\backslash\{i\})}\right] + \frac{1}{p(I) - p(I\backslash\{n\})} \\
&< \sum_{i=1}^{n-2}\frac{1}{p(I) - p(\{j|j<n\}\backslash\{i\})} + \frac{1}{p(I) - p(I\backslash\{n\})} \\
&< \frac{n-1}{p(I) - p(I\backslash\{n\})},
\end{aligned}$$

and that

$$MC(0) = \frac{1}{p(I) - p(I \backslash \{n-1\})}.$$

Let $\delta = \frac{1}{n}$; thus, if $[p(I) - p(I\backslash\{n-1\})] < \delta\,[p(I) - p(I\backslash\{n\})]$, then $MB(0) < MC(0)$. This implies that $m^* = 0$. The proposition is thus proven. □

Proof of Lemma 5

Proof. We first prove statement (i). Suppose in π^* we have $m_i > 0$ and $m_j < \hat{n}_j - 1$, then the principal can profit by decreasing m_i by 1 and increasing m_j by 1. Let the new sequence be π', and denote the new predecessor and new successor $\hat{i}$ and $\hat{j}$, respectively. Because the agents are equally important, we have $v^*_{\hat{i}}(\pi^*) = v^*_{\hat{j}}(\pi')$. Moreover, because the centers move sequentially and c_i moves before c_j, we have $|M'_{\hat{i}}| > |M^*_{\hat{j}}|$; thus, $v^*_{\hat{i}}(\pi') < v^*_{\hat{j}}(\pi^*)$. In addition, note that for any agent k other than $\hat{i}$ and $\hat{j}$, $|M'_k| \geq |M^*_k|$. Specifically, if $k = c_j$, or $c_i \notin M_k$ and $c_j \in M_k$, then $|M'_k| > |M^*_k|$; otherwise, $|M'_k| = |M^*_k|$. It follows that $v^*_k(\pi') \leq v^*_k(\pi^*)$; therefore, the total rewards are lower under π' than under π^*, a contradiction. The proof of statement (ii) is analogous. Thus, the lemma is proven. □

Proof of Proposition 4

Proof. Without loss of generality, let star i moves immediately before star j, i.e., $j = i + 1$. We first prove statement (i). Suppose not, then $\hat{n}_i < \hat{n}_j$. Note that for any peripheral agent $\hat{i}$ of star i and any peripheral agent $\hat{j}$ of star j, we have $|M^*_{\hat{i}}| = |M^*_{\hat{j}}| + 1$. Now swap stars i and j with m_i and m_j both remaining 0. Denote the new sequence π'. Thus, $|M'_{c_i}| = |M^*_{c_j}|$, $|M'_{c_j}| = |M^*_{c_i}|$, $|M'_{\hat{i}}| = |M^*_{\hat{j}}|$ and $|M'_{\hat{j}}| = |M^*_{\hat{i}}|$. For any agent k who belongs to neither star i nor j, $|M'_k| = |M^*_k|$. Then by Proposition 1, the difference in total rewards is given by

$$\begin{aligned} V^*(\pi') - V^*(\pi^*) &= \left[\frac{\hat{n}_i - 1}{p(n) - p(n-1-|M'_{\hat{i}}|)} + \frac{\hat{n}_j - 1}{p(n) - p(n-1-|M'_{\hat{j}}|)}\right] \\ &\quad - \left[\frac{\hat{n}_i - 1}{p(n) - p(n-1-|M^*_{\hat{i}}|)} + \frac{\hat{n}_j - 1}{p(n) - p(n-1-|M^*_{\hat{j}}|)}\right] \\ &= \frac{\hat{n}_j - \hat{n}_i}{p(n) - p(n-1-|M^*_{\hat{i}}|)} - \frac{\hat{n}_j - \hat{n}_i}{p(n) - p(n-1-|M^*_{\hat{j}}|)} < 0. \end{aligned}$$

The inequality is due to the monotonicity. Thus, we obtain a contradiction.

Second, we prove statement (ii). Suppose not, then $\hat{n}_i > \hat{n}_j$. Now swap stars i and j with

$m_i = \hat{n}_i - 1$ and $m_j = \hat{n}_j - 1$. Denote the new sequence π'. Thus, $|M'_{c_i}| = |M^*_{c_j}| + \hat{n}_i - \hat{n}_j$ and $|M'_{c_j}| = |M^*_{c_i}|$. For any other agent k, $|M'_k| = |M^*_k|$. Then by Proposition 1, we have

$$\begin{aligned} V^*(\pi') - V^*(\pi^*) = & \left[\frac{1}{p(n) - p(n-1-|M'_{c_i}|)} + \frac{1}{p(n) - p(n-1-|M'_{c_j}|)} \right] \\ & - \left[\frac{1}{p(n) - p(n-1-|M^*_{c_i}|)} + \frac{1}{p(n) - p(n-1-|M^*_{c_j}|)} \right] \\ = & \frac{1}{p(n) - p(n-1-|M'_{c_i}|)} - \frac{1}{p(n) - p(n-1-|M^*_{c_j}|)} < 0. \end{aligned}$$

The inequality is due to the monotonicity and that $|M'_{c_i}| = |M^*_{c_j}| + \hat{n}_i - \hat{n}_j$. Thus, we obtain a contradiction. In summary, the proposition is proven. $\square$

Proof of Corollary 4

Proof. Given the sub-sequence π^C, it suffices to show that MB is decreasing in the number of successors $\sum m_i$. Fix some $\sum m_i \geq 0$, add an extra successor following the algorithm in Corollary 4. Suppose for some agent j, M_j enlarges due to the increase in successors. Then, the change in j's reward, $v^*_j(\sum m_i + 1) - v^*_j(\sum m_i)$, is given by

$$\frac{1}{p(n) - p(n-1-\kappa - \sum m_i - 1)} - \frac{1}{p(n) - p(n-1-\kappa - \sum m_i)},$$

where κ is the number of centers in M_j. Then by the discussion of formula (1) in the text, the above formula is decreasing in $\sum m_i$. Moreover, as the number of successors increases, the number of such agent j also decreases. These facts imply that the marginal reduction in rewards decreases as the number of successors increases; that is, MB is decreasing in $\sum m_i$. Because MC is constant, the algorithm is valid. Thus, the corollary is proven. $\square$

Proof of Proposition 5

Proof. We show that the upper bound of $\sum m^*_i$ is given by $\bar{m}$. The proof for the lower bound is analogous. Pick any two stars i and j such that $\hat{n}_i > \hat{n}_j$. Consider two sub-sequences π^C_1 and π^C_2 with the only difference that c_i moves immediately after c_j in π^C_1, whereas c_j moves immediately after c_i in π^C_2. Then, apply the algorithm in Corollary 4 to π^C_1 and π^C_2. The next lemma shows that for any fixed number of successors $\sum m_i \geq 0$, the marginal benefit in π^C_1, $MB_1(\sum m_i)$, is greater than that in π^C_2, $MB_2(\sum m_i)$.

Lemma A.2. *For any integer* $\sum m_i \geq 0$, $MB_1(\sum m_i) \geq MB_2(\sum m_i)$.

Proof. First note that for any $\sum m_i$, if an agent belongs to neither star i nor star j, then he receives the same reward in both sequences. This implies that we can restrict our attention to the impacts of $\sum m_i$ on the agents who belong to either star i or star j. For both sequences, let x be the number of stars after both stars i and j, and y be the total number of peripheral agents of these stars. We consider three cases. First, when $\sum m_i < y + \hat{n}_j - 1$, because MC is constant, $MB_1(\sum m_i) - MB_2(\sum m_i)$ equals

$$\begin{aligned}
&\left[V_2^*(\sum m_i + 1) - V_1^*(\sum m_i + 1)\right] - \left[V_2^*(\sum m_i) - V_1^*(\sum m_i)\right] \\
&= (\hat{n}_i - \hat{n}_j)\left[\frac{1}{p(n) - p(n-2-x-\sum m_i)} - \frac{1}{p(n) - p(n-3-x-\sum m_i)}\right] \\
&\quad - (\hat{n}_i - \hat{n}_j)\left[\frac{1}{p(n) - p(n-3-x-\sum m_i)} - \frac{1}{p(n) - p(n-4-x-\sum m_i)}\right] > 0.
\end{aligned}$$

The inequality is due to the complementarity and that $\hat{n}_i > \hat{n}_j$.

Second, when $y + \hat{n}_j - 1 \leq \sum m_i < y + \hat{n}_i - 1$, $MB_1(\sum m_i) - MB_2(\sum m_i)$ equals

$$\begin{aligned}
&\left[\frac{1}{p(n) - p(n-1-x-\sum m_i)} - \frac{1}{p(n) - p(n-2-x-\sum m_i)}\right] \\
&- (\hat{n}_i - \sum m_i + y - 2)\left[\frac{1}{p(n) - p(n-3-x-\sum m_i)} - \frac{1}{p(n) - p(n-4-x-\sum m_i)}\right] \\
&+ (\hat{n}_i - \sum m_i + y - 1)\left[\frac{1}{p(n) - p(n-2-x-\sum m_i)} - \frac{1}{p(n) - p(n-3-x-\sum m_i)}\right] > 0.
\end{aligned}$$

The inequality is due to the monotonicity and the complementarity.

Finally, when $\sum m_i \geq y + \hat{n}_i - 1$, $MB_1(\sum m_i) - MB_2(\sum m_i)$ equals

$$\begin{aligned}
&\left[\frac{1}{p(n) - p(n-\hat{n}_j-x-y)} - \frac{1}{p(n) - p(n-\hat{n}_i-x-y)}\right] \\
&- \left[\frac{1}{p(n) - p(n-\hat{n}_j-x-y)} - \frac{1}{p(n) - p(n-\hat{n}_i-x-y)}\right] = 0.
\end{aligned}$$

In summary, for any $\sum m_i$, $MB_1(\sum m_i) \geq MB_2(\sum m_i)$. Thus, the lemma is proven. □

Lemma A.2 implies, by induction, that for any integer $\sum m_i$, $MB(\sum m_i)$ is the highest in the sub-sequence in which the stars are aligned in ascending order of size. It follows from Corollary 4 that the optimal number of successors is the highest in this sub-sequence. This means that $\sum m_i^* \leq \bar{m}$. Thus, the proposition is proven. □

Proof of Proposition 6

Proof. Suppose the optimal sub-sequence of centers has been determined, then the optimal sequence π^* can be obtained by applying the algorithm in Corollary 4. Given any number of successors $\sum m_i$, let star j be the last star in the sub-sequence such that $m_j < \hat{n}_j - 1$. It follows that for any star $k \neq j$, $m_k = 0$ if k moves before j or $m_k = \hat{n}_k - 1$ otherwise. Thus, for any agent $l \neq c_j$, either $l \in M_{c_j}$ or $c_j \in M_l$. Then, re-configure g hypothetically into a star network such that c_j is the center and agent l is c_j's successor (predecessor) if $l \in M_{c_j}$ ($c_j \in M_l$). Note that if initially $c_j \in M_l$, then after the re-configuration l's action is (weakly) less transparent, i.e., $|M_l|$ is (weakly) lower. It follows from the discussion of formula (1) in the text that if $\sum m_i$ increases by 1, then the reduction in l's reward will be (weakly) larger after the re-configuration than before. This implies that the marginal benefit of increasing $\sum m_i$ in the multi-star graph is lower than that of the single star with $m = \sum m_i$. Because the marginal cost is the same in both cases, it remains to show that in the optimal sequence of the single star, m^* is bounded above by some number of order $\sqrt{n}$. Following Section 4, we have that $MB(m) - MC(m)$ is given by

$$
\begin{aligned}
&\frac{n-m-2}{p(n)-p(n-m-2)} - \frac{n-m-2}{p(n)-p(n-m-3)} + \frac{1}{p(n)-p(n-m-1)} - \frac{1}{p(n)-p(n-1)} \\
&= \frac{(n-m-2)[p(n-m-2)-p(n-m-3)]}{[p(n)-p(n-m-2)][p(n)-p(n-m-3)]} - \frac{p(n-1)-p(n-m-1)}{[p(n)-p(n-m-1)][p(n)-p(n-1)]} \\
&\leq \frac{(n-m-2)[p(n-m-2)-p(n-m-3)]}{[p(n)-p(n-m-2)][p(n)-p(n-m-3)]} - \frac{Km}{[p(n)-p(n-m-1)]} \\
&< \frac{(n-m-2)[p(n-m-2)-p(n-m-3)]}{[p(n)-p(n-m-1)][p(n)-p(n-m-3)]} - \frac{Km}{[p(n)-p(n-m-1)]} \\
&\propto \frac{(n-m-2)[p(n-m-2)-p(n-m-3)]}{p(n)-p(n-m-3)} - Km \\
&= \frac{(n-m-2)[p(n-m-2)-p(n-m-3)]}{\sum_{i=0}^{m+2}[p(n-i)-p(n-i-1)]} - Km \\
&< \frac{(n-m-2)[p(n-m-2)-p(n-m-3)]}{(m+3)[p(n-m-2)-p(n-m-3)]} - Km = \frac{-Km^2-(3K+1)m+n-2}{m+3}.
\end{aligned}
$$

The first inequality is due to Assumption 1, and the last inequality is due to complementarity. Clearly, the positive root of the RHS of the last equality is of order $\sqrt{n}$. Then by (3), the number of successors in the optimal sequence of this star, m^*, is bounded above by some integer of order $\sqrt{n}$. It follows from the above argument that the number of successors in the optimal sequence of the original multi-star network is also bounded above by some integer of order $\sqrt{n}$. Thus, the proposition is proven. □

Proof of Proposition 7

Proof. First note that property (i) is a corollary of Lemma 2. Suppose property (ii) does not hold, then in the optimal sequence π^* there exist a predecessor i and a successor j such that agent i is more important than agent j and i moves after j's star. Consider a different sequence π' which differs from π^* only in that i is a successor and j is a predecessor. Because i is more important than j, $v_i^*(\pi') < v_j^*(\pi^*)$. Moreover, note that $\{i\} \cup M_i^* \subset \{j\} \cup M_j'$; thus, $v_j^*(\pi') < v_i^*(\pi^*)$. For any agent $k \neq i, j$, if k's a successor, then k's reward is the same in both π^* and π'. Suppose k is a predecessor, then for any agent $l \neq i, j$ and $l \in M_k^*$, we have $l \in M_k'$. Moreover, if $j \in M_k^*$, then j is replaced by i in M_k'; if $j \notin M_k^*$, then we have possibly $i \in M_k'$. It follows that $v_k^*(\pi') \leq v_k^*(\pi^*)$. In summary, the total rewards are lower in π' than in π^*, a contradiction. Thus, a successor is more important than each predecessor in every star that moves later. If further property (iii) holds, then m_i^* is nondecreasing along π^*.

To prove property (iii), suppose in π^* there exist two stars i and j such that i is more important than j, and j is immediately after i. By property (i), star i's (j's) successors are its m_i^* (m_j^*) most important peripheral agents. Let $\bar{s} = \max\{m_i^*, m_j^*\}$ and $\underline{s} = \min\{m_i^*, m_j^*\}$. Swap stars i and j such that i and j has now $\bar{s}$ and $\underline{s}$ successors, respectively, and every successor is more important than any predecessor of his star. Denote the new sequence π'. To compare the total rewards needed in π^* and π', we categorize the agents in four groups:

1. All successors except those in stars i and j, and all the agents that move after j in π^* and i in π'. The rewards to these agents are the same in both π^* and π'.

2. All the centers and predecessors in the stars that move before star i in π^* and j in π'. For each of such agent k, we have $|M_k^*| = |M_k'|$. Because star i is more important than star j, for each agent $l \in M_k^*$, either $l \in M_k'$ or l is replaced by a more important agent in M_k'. Thus, the reward to agent k is lower in π' than in π^*.

3. c_i and c_j. First note that $|M_{c_i}^*| = |M_{c_j}'|$. Because star i is more important than star j, for each agent $k \in M_{c_i}^*$, either $k \in M_{c_j}'$ or k is replaced by a more important agent in M_{c_j}'. Hence, $v_{c_j}^*(\pi') \leq v_{c_i}^*(\pi^*)$. Note too that for each agent $k \in M_{c_j}^*$, either $k \in M_{c_i}'$ or k is replaced by a more important agent in M_{c_i}'. Also, c_i is more important than c_j. Moreover, when $m_i^* > m_j^*$, $|M_{c_i}'| > |M_{c_j}^*|$. These imply that $v_{c_i}^*(\pi') < v_{c_j}^*(\pi^*)$.

4. The peripheral agents of stars i and j. Choose any two peripheral agents k and l from stars i and j, respectively, such that they share the same importance rank in own star. Note that if both k and l are successors in both π^* and π', then their rewards do not

change between π^* and π'. If both k and l are predecessors in both π^* and π', then we consider two cases. First, if $m_i^* \geq m_j^*$, then $v_k^*(\pi^*) + v_l^*(\pi^*) - [v_k^*(\pi') + v_l^*(\pi')]$ equals

$$\begin{aligned}
&\frac{1}{p(I) - p(I \backslash (\{k\} \cup M_k^*))} + \frac{1}{p(I) - p(I \backslash (\{l\} \cup M_l^*))} \\
&- \left[\frac{1}{p(I) - p(I \backslash (\{k\} \cup M_k'))} + \frac{1}{p(I) - p(I \backslash (\{l\} \cup M_l'))} \right] \\
&> \frac{1}{p(I) - p(I \backslash (\{k\} \cup M_k^*))} + \frac{1}{p(I) - p(I \backslash (\{l\} \cup M_k'))} \\
&- \left[\frac{1}{p(I) - p(I \backslash (\{k\} \cup M_k'))} + \frac{1}{p(I) - p(I \backslash (\{l\} \cup M_l'))} \right]. \qquad \text{(A.3)}
\end{aligned}$$

The inequality is because $M_l^* = \{c_j\} \cup M_{c_j}^*$ and $M_k' = \{c_i\} \cup M_{c_i}'$, and by part (3) we have M_l^* is less important than M_k'. Note that $\bar{s} = m_i^* \geq m_j^* = \underline{s}$, thus $M_k^* = M_l'$. It follows that the RHS of (A.3) can be rewritten as

$$\begin{aligned}
&\frac{1}{p(I) - p(I \backslash (\{l\} \cup M_k'))} - \frac{1}{p(I) - p(I \backslash (\{l\} \cup M_k^*))} \\
&- \left[\frac{1}{p(I) - p(I \backslash (\{k\} \cup M_k'))} - \frac{1}{p(I) - p(I \backslash (\{k\} \cup M_k^*))} \right].
\end{aligned}$$

Rearranging, this is equal to

$$\begin{aligned}
&\frac{p(I \backslash (\{l\} \cup M_k')) - p(I \backslash (\{l\} \cup M_k^*))}{[p(I) - p(I \backslash (\{l\} \cup M_k'))][p(I) - p(I \backslash (\{l\} \cup M_k^*))]} \\
&- \frac{p(I \backslash (\{k\} \cup M_k')) - p(I \backslash (\{k\} \cup M_k^*))}{[p(I) - p(I \backslash (\{k\} \cup M_k'))][p(I) - p(I \backslash (\{k\} \cup M_k^*))]}. \qquad \text{(A.4)}
\end{aligned}$$

Let $S = I \backslash (\{l\} \cup M_k^*)$, $S' = I \backslash (\{k\} \cup M_k^*)$ and $T = M_k^* \backslash M_k'$. Note that $M_k' \subset M_k^*$ and k is more important than l; thus, $\{S, S', T\}$ satisfies the conditions of Assumption 2. It follow from Assumption 2 and the monotonicity that (A.4) is positive. This implies that $v_k^*(\pi^*) + v_l^*(\pi^*) - [v_k^*(\pi') + v_l^*(\pi')] > 0$. Second, if $m_i^* < m_j^*$, then (A.3) still holds, but M_k^* is now less important than M_l' because some successor in M_k^* is replaced by a more important successor in M_l'. This means that the RHS of (A.3) is greater than (A.4). Note that now $M_k' \not\subseteq M_k^*$. If M_k' is less important than M_k^*, then by the same argument we have that (A.4) is positive. Otherwise, the LHS of (A.3) is automatically positive because M_l' is more important than M_l^*. To summarize, if both k and l are predecessors in both π^* and π', then $v_k^*(\pi^*) + v_l^*(\pi^*) - [v_k^*(\pi') + v_l^*(\pi')] > 0$.

Finally, if k and l have different positions relative to own centers in either π^* or π' (i.e., in either π^* or π', one of k and l is a predecessor while the other is a successor), then there are two cases. First, if $m_i^* > m_j^*$, then k is a successor while l is a predecessor, in both π^* and π'. Thus, $v_k^*(\pi^*) = v_k^*(\pi')$ and $v_l^*(\pi^*) > v_l^*(\pi')$. Second, if $m_i^* < m_j^*$, then k (l) is a predecessor (successor) in π^* but a successor (predecessor) in π'. Thus, $v_k^*(\pi^*) + v_l^*(\pi^*) - [v_k^*(\pi') + v_l^*(\pi')]$ equals

$$\begin{aligned}
&\frac{1}{p(I) - p(I\backslash(\{k\} \cup M_k^*))} + \frac{1}{p(I) - p(I\backslash\{l\})} \\
&- \left[\frac{1}{p(I) - p(I\backslash(\{k\})} + \frac{1}{p(I) - p(I\backslash(\{l\} \cup M_l'))}\right] \\
&> \frac{1}{p(I) - p(I\backslash(\{k\} \cup M_k^*))} + \frac{1}{p(I) - p(I\backslash\{l\})} \\
&- \left[\frac{1}{p(I) - p(I\backslash(\{k\})} + \frac{1}{p(I) - p(I\backslash(\{l\} \cup M_k^*))}\right] > 0.
\end{aligned}$$

The first inequality is because $M_l' = \{c_j\} \cup M_{c_j}'$, by part (3), it is more important than $M_k^* = \{c_i\} \cup M_{c_i}^*$. The second inequality is due to the complementarity in importance. Therefore, we always have $v_k^*(\pi^*) + v_l^*(\pi^*) - [v_k^*(\pi') + v_l^*(\pi')] > 0$.

In summary, the total rewards are strictly lower in π' than in π^*, a contradiction. Therefore, property (iii) is proven. This completes the proof. □

Proof of Corollary 5

Proof. Consider a sequence that aligns the stars in ascending order of importance. To prove that the proposed algorithm identifies the optimal sequence, it suffices to show that if at the optimum a star i has a predecessor r_i^k, it is never optimal to make any peripheral agent in any previous star j, who is weakly less important than r_i^k, a successor. This follows from the proof of Proposition 7. Thus, the corollary is proven. □

References

Alchian, A.A. and Demsetz, H. "Production, Information Costs, and Economic Organization." *American Economic Review*, Vol. 62(5) (1972), pp. 777–795.

Au, P.H. and Chen, B.R. "Matching with Peer Monitoring." *Journal of Economic Theory*, Vol. 192 (2021), p. 105172.

Babaioff, M., Feldman, M., Nisan, N., and Winter, E. "Combinatorial Agency." *Journal of Economic Theory*, Vol. 147(3) (2012), pp. 999–1034.

Balmaceda, F. "Optimal Task Assignments." *Games and Economic Behavior*, Vol. 98 (2016), pp. 1–18.

Balmaceda, F. "Optimal Task Assignments with Loss-Averse Agents." *European Economic Review*, Vol. 105 (2018), pp. 1–26.

Bernstein, S. and Winter, E. "Contracting with Heterogeneous Externalities." *American Economic Journal: Microeconomics*, Vol. 4(2) (2012), pp. 50–76.

Camboni, M. and Porcellacchia, M. "Monitoring Team Members: Information Waste and the Self-Promotion Trap." working paper (2021).

Carpenter, J., Bowles, S., Gintis, H., and Hwang, S.H. "Strong Reciprocity and Team Production: Theory and Evidence." *Journal of Economic Behavior and Organization*, Vol. 71(2) (2009), pp. 221–232.

Che, Y.K. and Yoo, S.W. "Optimal Incentives for Teams." *American Economic Review*, Vol. 91(3) (2001), pp. 525–541.

Gershkov, A., Li, J., and Schweinzer, P. "Efficient Tournaments within Teams." *The RAND Journal of Economics*, Vol. 40(1) (2009), pp. 103–119.

Gershkov, A., Li, J., and Schweinzer, P. "How to Share It out: The Value of Information in Teams." *Journal of Economic Theory*, Vol. 162 (2016), pp. 261–304.

Goerg, S.J., Kube, S., and Zultan, R. "Treating Equals Unequally: Incentives in Teams, Workers' Motivation, and Production Technology." *Journal of Labor Economics*, Vol. 28(4) (2010), pp. 747–772.

Gould, E.D. and Winter, E. "Interactions between Workers and the Technology of Production: Evidence from Professional Baseball." *Review of Economics and Statistics*, Vol. 91(1) (2009), pp. 188–200.

Halac, M., Lipnowski, E., and Rappoport, D. "Rank Uncertainty in Organizations." *American Economic Review*, Vol. 111(3) (2021), pp. 757–86.

Hermalin, B.E. "Toward an Economic Theory of Leadership: Leading by Example." *American Economic Review*, Vol. 88(5) (1998), pp. 1188–1206.

Heywood, J.S. and Jirjahn, U. "Teams, Teamwork and Absence." *Scandinavian Journal of Economics*, Vol. 106(4) (2004), pp. 765–782.

Holmstrom, B. "Moral Hazard in Teams." *The Bell Journal of Economics*, Vol. 13(2) (1982), pp. 324–340.

Holmstrom, B. and Milgrom, P. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization*, Vol. 7 (1991), pp. 24–52.

Ichino, A. and Maggi, G. "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm." *The Quarterly Journal of Economics*, Vol. 115(3) (2000), pp. 1057–1090.

Itoh, H. "Incentives to Help Multi-Agent Situations." *Econometrica*, Vol. 59(3) (1991), pp. 611–636.

Mas, A. and Moretti, E. "Peers at Work." *American Economic Review*, Vol. 99(1) (2009), pp. 112–145.

McAfee, R.P. and McMillan, J. "Optimal Contracts for Teams." *International Economic Review*, Vol. 32(3) (1991), pp. 561–577.

Segal, I. "Contracting with Externalities." *The Quarterly Journal of Economics*, Vol. 114(2) (1999), pp. 337–388.

Segal, I. "Coordination and Discrimination in Contracting with Externalities: Divide and Conquer?" *Journal of Economic Theory*, Vol. 113(2) (2003), pp. 147–181.

Steiger, E.M. and Zultan, R. "See No Evil: Information Chains and Reciprocity." *Journal of Economic Behavior and Organization*, Vol. 109 (2014), pp. 1–12.

Sun, Y. and Zhao, W. “Moral Hazard on Networks.” working paper (2021).

Winter, E. “Incentives and Discrimination.” *American Economic Review*, Vol. 94(3) (2004), pp. 764–773.

Winter, E. “Optimal Incentives for Sequential Production Processes.” *The RAND Journal of Economics*, Vol. 37(2) (2006), pp. 376–390.

Winter, E. “Transparency and Incentives among Peers.” *The RAND Journal of Economics*, Vol. 41(3) (2010), pp. 504–523.

Zhou, J. “Economics of Leadership and Hierarchy.” *Games and Economic Behavior*, Vol. 95 (2016), pp. 88–106.

Zhou, J. and Chen, Y.J. “Key Leaders in Social Networks.” *Journal of Economic Theory*, Vol. 157 (2015), pp. 212–235.