Costless Coordination through Public Contracting

Yating Yuan*

August 17, 2025

Abstract

A principal incentivizes a team of agents to work on a joint project. Building on Winter (2004), this paper proposes a simple, efficient mechanism that implements work as the unique outcome under any procedure of Iterative Elimination of Weakly Dominated Strategies. The mechanism asks agents to choose between two public messages, "collaborate" and "monopolize," and the message profile decides their bonuses upon team success. Unlike Winter (2004) and Halac et al. (2021), the equilibrium bonus allocation is both non-discriminatory and public. Thus, efficiency need not come at the cost of fairness nor transparency.

JEL: D23 D62 D86

Keywords: team incentives, pay discrimination, pay transparency

1 Introduction

A trade-off between efficiency, fairness, and transparency sits at the heart of team incentives. To ensure everyone works hard, a cost-minimizing manager must either accept some inefficiency and treat similar workers unequally (e.g., Winter (2004)) or keep contracts private (e.g., Halac et al. (2021)).

*Department of Economics, University of Warwick, Coventry, CV47AL, United Kingdom. Email: yating.yuan@warwick.ac.uk. I am grateful to my supervisors, Jacob Glazer and Motty Perry, for their guidance, support, and insightful discussions. I would also like to thank Costas Cavounidis, James Fenske, Hyungmin Park, Herakles Polemarchakis and Debraj Ray for their comments.

But private contracts are often far from ideal. Achieving "efficient coordination"[1] typically requires interdependent rewards: each worker's compensation depends on what others are offered. When those offers are private, a worker may not even be able to verify their own pay terms until the work is done. Moreover, norms are shifting. Salary transparency laws are spreading (Cullen, 2024). Organizations want open pay structures. The demand for contracts that are transparent, fair, and still deliver effort is only growing.

This paper shows that such contracts are not only possible — they can be simple and more robust than the alternatives. Building on the model of Winter (2004), I propose a mechanism that keeps pay public and fair, yet achieves efficient coordination. The key is to use signaling to coordinate effort: a worker's public contract choice signals his future effort and aligns beliefs through forward-induction reasoning (e.g., Ben-Porath and Dekel (1992); Cavounidis and Park (2025)).[2]

The mechanism offers each worker a binary menu of contracts. After observing each other's choices, workers decide whether to shirk. In the unique outcome that survives Iterative Elimination of Weakly Dominated Strategies (IEWDS), all workers choose the contract that signals "I will work," and the manager pays approximately the second-best cost upon team success. The solution concept matters: IEWDS admits forward induction. It also ensures the mechanism works even without common beliefs. No one is paid unfairly. Effort is coordinated out in the "open." And the mechanism breaks the trilemma.

To fix ideas, consider a project with two tasks: data collection and marketing design. A manager hires Alice to collect data and Bob to design the marketing strategy. Each of them can privately exert costly effort to increase the chance of task success. The project succeeds only if both tasks succeed. Everyone observes the project outcome, but if it fails, no one can tell which task went wrong. The manager's goal is to induce both workers to exert effort as the unique outcome under IEWDS while minimizing total payments.

[1]To focus on coordination, I take the second-best as the benchmark (See also Ma (1988) and Arya et al. (1997)). A manager achieves efficient coordination if she implements the outcome *all workers exert effort* at their second-best cost, i.e., the total payment that incentivizes team effort when workers believe their teammates will also work. Workers still receive a positive rent due to hidden action.

[2]Cavounidis and Park (2025) is the first paper that brings the signaling effect and forward-induction reasoning to the moral-hazard-in-teams setting.

A natural benchmark is to reward each worker a bonus just above the second-best threshold upon team success, which is just enough to incentivize effort *if* the other worker works hard. This contract seems efficient and fair—it does not discriminate between symmetric workers. But we cannot rule out coordination failure: if either worker expects the other to shirk, the bonus falls short, and shirking becomes a best response.

My solution, which I call a *collaborating* mechanism, works as follows. The manager picks an arbitrary worker, say Alice, to send a public message: *collaborate* or *monopolize*. If Alice chooses *collaborate*, both she and Bob are promised a "collaborating" bonus that is ε-close to the second-best benchmark. If Alice chooses *monopolize*, she is promised a larger "monopoly" bonus, carefully chosen by the manager, while Bob receives less. After observing the message, both privately decide whether to exert effort.

In the unique outcome under IEWDS, Alice chooses *collaborate* and both workers exert effort. Why? If the monopoly bonus is large enough, choosing *collaborate* while planning to shirk becomes weakly dominated by the monopoly option for Alice ("signaling constraint"). So if she chooses *collaborate*, Bob infers that she intends to exert effort. This belief makes it worthwhile for him to do the same. Knowing this, Alice will indeed choose *collaborate* rather than *monopolize*, provided that the monopoly bonus is not too attractive ("incentive constraint"). Since promising the team a collaborating bonus above the second-best makes *collaborate and then both work* the best outcome for everyone, the manager can always find a monopoly bonus that satisfies both constraints and coordinate efforts for free. Section 2 explains the two-agent example in detail.

Section 3 and 4 characterize a general setting and describe the N-agent mechanism. When more than two workers signal simultaneously, coordination may break down (Ben-Porath and Dekel, 1992). To avoid this failure, the mechanism introduces a voice hierarchy: the manager ranks the workers arbitrarily, and a worker gets the chance to monopolize only if all the higher-ranked workers have chosen *collaborate*. This voice hierarchy replaces the payment hierarchy in Winter (2004) and serves a similar coordination role. Yet, the order of voices—who speaks or speaks louder—does not matter even when agents are asymmetric in effort costs or the contributions to team success.

Section 5 briefly demonstrates how the main result can be applied to improve incentive provision in real-world settings, including fee-splitting in legal work, credit allocation in academia, and subcontracting.

Related Literature

This paper contributes to the literature on contracting against strategic uncertainty in team production. In the setting of public contracting, in which all bonus offers are publicly known, Winter (2004) provides a benchmark analysis. He studies the optimal independent contract where agents cannot influence their own or others' contracts. The key insight there is that discriminatory bonuses are necessary to ensure full effort in any Nash equilibrium at a minimum cost: a large bonus makes exerting effort a dominant strategy for one agent, so incentivizing the other requires only a small bonus.

Cavounidis and Park (2025) extends this framework by allowing for interdependent contracting, but focuses on a natural, constrained class of mechanisms—subcontracting. The principal sets a bonus budget and delegates bonus allocation to a sequence of agents. The authors incorporate extensive-form rationalizability (Pearce, 1984) in Nash equilibrium to capture forward induction. Subcontracting can sometimes outperform Winter (2004)'s contract, but not always. In both papers, the constraints on the mechanism space or the equilibrium concept push the cost above the second-best level and lead to discrimination. Building on the subcontracting mechanism, my paper designs a public mechanism that achieves the second-best and eliminates discrimination under IEWDS.[3]

When it comes to private contracting, the main finding in Halac et al. (2021) reveals that the optimal private independent contract outperforms Winter (2004) and eliminates discrimination by creating rank uncertainty and mutual assurance among agents. Beyond independent contracting, the principal can approximate the second-best payoff with a mechanism in which agents send private messages before/after effort decisions (Ma, 1988; Arya et al., 1997; Cavounidis and Ghosh, 2021) or a private incentive scheme where an agent's

[3]Both IEWDS and extensive-form rationalizability capture forward-induction reasoning. Their relationship has been discussed in specific classes of games (Battigalli, 1997; Shimoji, 2002, 2004), but remains unclear in general. Under the optimal mechanism of this paper, the unique outcome under IEWDS is also extensive-form rationalizable. A similar observation appears in money-burning games (Shimoji, 2002, 2004).

bonus offer depends on the private offers of others (Halac, Lipnowski, and Rappoport, 2021). In these papers, agents are not informed of others' payments and messages, nor can they verify their own contractual terms before work begins.[4] This lack of transparency in one's own contract is less common in practice. My paper provides a simple mechanism that achieves the second-best payoff without requiring agents to work under missing contract terms.

The signaling effect, that public contract choices signal future action, is closely related to the literature on money-burning (Kohlberg and Mertens, 1986; Van Damme, 1989; Ben-Porath and Dekel, 1992; Hurkens, 1996; Shimoji, 2002) and cheap talk (Antić and Persico, 2023). In particular, Ben-Porath and Dekel (1992) give a pre-game money-burning option to one player to signal their future action in a two-player game. The signaling player can implement their preferred outcome, without actually burning money, as a unique outcome under the maximal IEWDS procedure that deletes all weakly dominated strategies at each stage. In the moral-hazard-in-teams setting, subcontracting works through a similar effect (Cavounidis and Park, 2025), but grants the principal more leeway in designing the game form and the option. My paper expands the mechanism space and optimally applies this signaling logic to achieve efficient coordination.

Finally, this paper shows how coordination concerns can generate a new form of hierarchy in organizations. Organizational hierarchies are typically tied to differential pay since monetary incentives play a key role in shaping production and allocation decisions (see e.g., Mookherjee (2006); Garicano et al. (2013); Winter (2004); Halac et al. (2021)). In contrast, the mechanism here features a voice hierarchy: agents choose between contracts in a public, hierarchical (or sequential) way. This structure facilitates multi-agent signaling and supports coordinated effort. While differential pay exists off-path, symmetric agents receive the same pay on-path. The hierarchy lies not really in who earns more, but in who speaks first.

[4]Various (sequential) mechanisms have also been explored to eliminate undesired equilibria in principal-agent problems with hidden types (Demski and Sappington, 1984; Ma et al., 1988; Mookherjee and Reichelstein, 1990; Glover, 1994) and, more broadly, in contract design problems with externalities (Segal, 2003; Genicot and Ray, 2006; Kapon et al., 2024).

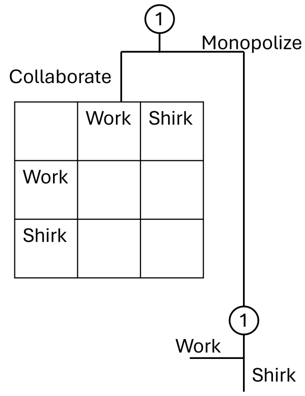

Figure 1: The game tree

	s_2		
		w	s
s_1	$collaborate, w$	$\mathbf{b_1^c - c, b_2^c - c}$	$pb_1^c - c, pb_2^c$
	$collaborate, s$	$pb_1^c, pb_2^c - c$	$p^2b_1^c, p^2b_2^c$
	$monopolize, w$	$pb_1^m - c, 0$	$\mathbf{pb_1^m - c, 0}$
	$monopolize, s$	$p^2b_1^m, 0$	$p^2b_1^m, 0$

Table 1: The reduced normal-form game

Note: In the reduced normal-form game, agent 1's strategy $s_1 \in S_1$ consists of a message and an effort decision following the message. Agent 2's (simplified) strategy $s_2 \in S_2$ is an effort decision after a *collaborate* message. The unique outcome under IEWDS is $((collaborate, w), w)$, and subgame-perfect Nash equilibria are highlighted in bold.

2 A Two-Agent Example

A principal hires two agents for a team project. Each agent performs a task and privately decides whether to exert effort $a_i \in \{w(ork), s(hirk)\}$ at cost $c \in (0, 1)$. Effort raises the chance of task success from $p \in (0, 1)$ to 1. The project succeeds if and only if both tasks succeed. While the project outcome is publicly observable, task outcomes are not. The principal aims to implement team effort as a unique outcome that survives any IEWDS procedure. Agents maximize expected payments net of the effort cost.

The *collaborating* mechanism works as follows. The principal arbitrarily selects one agent, say agent 1, to choose between two public messages: *collaborate* or *monopolize*. If agent 1 chooses *collaborate*, both agents receive a bonus offer ε-close to the *second-best* level upon project success: $b_i^c = \frac{c}{1-p} + \varepsilon$ for an arbitrarily small $\varepsilon > 0$. If agent 1 chooses *monopolize*, the principal promises her a large monopoly bonus upon project success, $b_1^m \geq 0$, but nothing for agent 2. The whole team will observe the message and then simultaneously decide whether to work.

The mechanism yields a unique outcome under any IEWDS procedure: agent 1 chooses *collaborate* and both agents subsequently choose *work*. Since *work* gives a negative payoff for agent 2 following a *monopolize* message, for illustration purposes we omit his action in that branch from the game tree (Figure 1) and its purely reduced normal-form representation (Table 1).[5] Two constraints underpin the main result:

1. Signaling constraint: $\max\{pb_1^m - c, p^2 b_1^m\} > pb_1^c$. This ensures that the monopoly option strictly dominates $(collaborate, s)$ for agent 1.

2. Incentive constraint: $\max\{pb_1^m - c, p^2 b_1^m\} < b_1^c - c$. This ensures that the monopoly option is strictly worse than $(collaborate, w)$ for agent 1 when agent 2 chooses *work* after a *collaborate* message.

The signaling constraint guarantees that $s_1 = (collaborate, s)$ never survives any IEWDS procedure. As a result, $s_2 = s$ will not survive either. The incentive constraint then rules out both $(monopolize, w)$ and $(monopolize, s)$ for agent 1. The only surviving outcome is $((collaborate, w), w)$. Since the collaborating bonus b_1^c exceeds the second-best level, i.e., $b_1^c - c > pb_1^c$, there always exists a monopoly bonus b_1^m that satisfies both constraints.

3 A General Model of Teamwork

A principal aims to incentivize a group of N agents, $i \in \mathbb{N} := \{1, 2, ..., N\}$, to complete a team project. Each agent privately decides whether to work or shirk, denoted by $a_i \in \{1(work), 0(shirk)\}$. *Work* incurs a personal cost $c_i > 0$. The team outcome — success or failure — depends on the set of agents who choose to work. For any set of agents, $J \subset \mathbb{N}$, let $P(J)$ denote the probability of team success if the agents in J work and others shirk. $P(\cdot)$ satisfies two assumptions: for any $J, J' \in \mathbb{N}$,

1. (monotonicity) if $J \subsetneq J', P(J) < P(J')$;

2. (complementarity) if J, J' are not nested, $P(J \cup J') - P(J) > P(J') - P(J \cap J')$.

The first assumption states that team success becomes more likely as more agents exert effort. The second introduces a core source of coordination friction: an agent's marginal contribution is higher when more teammates also choose to work.

[5] The proof for Theorem 1 shows that working after being monopolized will not survive IEWDS.

Only the team outcome is contractible. Agents maximize their expected payment net of any effort cost. If the principal promises a bonus $b_i \in [0, \infty)$ upon team success, agent i's payoff will be $P(\{i \in \mathbb{N} : a_i = 1\})b_i - c_i a_i$.

The principal's goal is to induce all N agents to work while minimizing total payment. To avoid distraction, I do not formally define a general mechanism space or the cost-minimization problem because the paper focuses on how a particular mechanism achieves the second-best. Instead, I introduce the concept of unique implementation under IEWDS and characterize the second-best benchmark.

Definition 1 (Unique Implementation under IEWDS)**.** *A mechanism uniquely implements full effort and bonus allocation* $\mathbf{b} := (b_1, b_2, ..., b_N) \in [0, \infty)^N$ *under IEWDS, if in the reduced normal-form representation of the induced game:*

1. *Every IEWDS procedure leads to the same unique outcome; and*
2. *On the equilibrium path of this outcome, each agent* i *is promised the bonus* b_i *conditional on team success, and all agents choose to work.*

Definition 2 (Second-Best Benchmark)**.** *The principal's second-best bonus for each agent* $i \in \mathbb{N}$ *is* $\underline{b}_i = \frac{c_i}{P(\mathbb{N}) - P(\{\mathbb{N}/\{i\}\})}$. *The second-best total bonus is* $\sum_{i \in \mathbb{N}} \underline{b}_i$.

No mechanism can uniquely implement full effort and a bonus allocation with $\sum_i b_i \leq \sum_i \underline{b}_i$. Appendix A provides the formal proof. The intuition is straightforward: If "all agents work" is not a strict Nash equilibrium given the bonus allocation, then some IEWDS procedure will leave a shirking outcome in play.

This second-best benchmark highlights our focus on the coordination issue. If the principal signs a contract with each agent independently and promises each agent (slightly more than) the second-best bonus $\underline{b}_i$ upon team success, miscoordination may happen.[6] No strategy will be weakly dominated because the unique best response is *shirk* when all others shirk. Neither IEWDS nor Nash equilibrium can eliminate the shirking outcome.

[6] Winter (2004) (Proposition 4) shows that offering each agent slightly more than the second-best bonus can induce all agents to work in the unique coalition-proof equilibrium. This equilibrium concept assumes a strong cooperative culture in the workplace.

4 Collaborating Mechanism and Efficiency

This section starts by describing the *collaborating* mechanism. It can uniquely implement full effort and any bonus allocation that does no better than the *second-best* benchmark ($b_i > \underline{b}_i$ for all $i \in \mathbb{N}$). By carefully designing how agents "speak" through contract choices, the principal can eliminate coordination rents and implement a fair, transparent incentive scheme.

4.1 Collaborating Mechanism

The principal selects a target bonus vector and ranks the agents arbitrarily. A collaborating mechanism grants the first $N-1$ agents a binary message space $\{collaborate, monopolize\}$; we refer to them as speakers. If all speakers choose *collaborate*, the mechanism offers each agent the target bonus conditional on team success. If any speaker chooses *monopolize*, the highest-ranked one will receive a large monopoly bonus, while all other agents receive nothing.

Let $R : \mathbb{N} \to \mathbb{N}$ be an arbitrary permutation of the agents. The following discussion refers to the agents by their identity $R(i)$, unless specified otherwise.

Definition 3 (Collaborating Mechanism)**.** *A collaborating mechanism, denoted by* $\mathcal{C}(\mathbf{b}^c, \mathbf{b}^m)$*, proceeds as follows:*

1. ***Contracting stage**: Each speaker* $i \in \{1, 2, ..., N-1\}$ *simultaneously sends a public message* $m_i \in M := \{collaborate, monopolize\}$.

   - *If all speakers choose collaborate, each agent receives a success-contingent bonus* b_i^c.
   - *If any speaker chooses monopolize, the highest-ranked speaker among them receives a personalized monopoly bonus* b_i^m *satisfying:*

$$\max\{P(\{i\})b_i^m - c_i,\, P(\phi)b_i^m\} > P(\mathbb{N} \setminus \{i\})b_i^c \tag{SC1}$$

$$\max\{P(\{i,j\})b_i^m - c_i,\, P(\{j\})b_i^m\} > P(\mathbb{N})b_i^c - c_i \quad \forall j \neq i \tag{SC2}$$

$$\max\{P(\{i\})b_i^m - c_i,\, P(\phi)b_i^m\} < P(\mathbb{N})b_i^c - c_i \tag{IC}$$

All other agents receive zero bonus.

2. ***Working stage**: After observing the full message profile, all agents simultaneously choose whether to work or shirk.*

The constraints mirror those in the two-agent example. Constraint (SC1) ensures that choosing *monopolize* weakly dominates *collaborate and then shirk*. Constraint (IC) ensures that *monopolize* is strictly worse than *collaborate* if an agent expects all others to collaborate and then work. The additional constraint (SC2) rules out bad behaviors off-path: it guarantees that, if another agent works after being monopolized, a speaker strictly prefers to seize the monopoly opportunity rather than collaborate and work.

Such a monopoly bonus vector $\mathbf{b}^m$ exists if $b_i^c > \underline{b}_i$ for all i. When each collaborating bonus exceeds the second-best level, constraints (SC1) and (IC) leave a nonempty set of b_i^m for us to choose. By strict monotonicity of $P(\cdot)$, we can always raise b_i^m within this set to satisfy (SC2).

Theorem 1. *Given any target bonus vector* $\mathbf{b}$ *with* $b_i > \underline{b}_i$ *for all* $i \in \mathbb{N}$*, the collaborating mechanism* $\mathcal{C}(\mathbf{b}^c, \mathbf{b}^m)$ *with* $\mathbf{b}^c = \mathbf{b}$ *and any* $\mathbf{b}^m$ *satisfying the constraints* (SC1)-(IC) *uniquely implements full effort and bonus allocation* $\mathbf{b}$ *under IEWDS.*

We now formally define strategies in the purely reduced normal-form game induced by the collaborating mechanism. Let $a_i : M^{N-1} \to \{0, 1\}$ be the effort function, mapping the full message profile to the effort decision of agent i . A speaker's strategy $s_i = (m_i, a_i(m_i, \mathbf{m}_{-i}))$ specifies his own message $m_i \in M$ and, given this message, whether to work or shirk after observing all other messages $\mathbf{m}_{-i} \in M^{N-2}$. A nonspeaker's strategy $s_N = a_N(\mathbf{m})$ simply specifies whether to work or shirk after observing all messages.

Let $\mathbf{s}^* := (s_i^*)_{i\in\mathbb{N}}$ denote the outcome that, as will be demonstrated, uniquely survives IEWDS. It consists of the following strategies. All speakers choose *collaborate*. Upon observing a unanimous collaborating message profile, all agents choose *work*. Off the equilibrium path, any message profile that results in agent i being monopolized (i.e., receiving a zero bonus) triggers a *shirk* response from agent i. Conversely, any message profile that allows agent i to monopolize triggers a *work* response from agent i if $P(\{i\})b_i^m - c_i > P(\phi)b_i^m$, and *shirk* otherwise. Compared to Winter (2004)'s contract, we move discriminatory bonuses off the equilibrium path while maintaining a "flat" structure of payments on path.

The proof has three parts. Part 1 (Lemma 1) shows that $\mathbf{s}^*$ survives any IEWDS procedure because it is a strict Nash equilibrium. Part 2 (Lemma 2) proves that no agent works after being monopolized. Part 3 finalizes the proof by showing that no other outcome survives under IEWDS.

Lemma 1. $\mathbf{s}^*$ *survives any IEWDS procedure.*

Proof of Lemma 1. Note first that a strict Nash equilibrium, in which each agent's strategy is the unique best response to the others, cannot be eliminated under any IEWDS procedure. We therefore aim to show that $\mathbf{s}^*$ is a strict Nash equilibrium.

First, suppose agent i chooses *collaborate* but deviates in the working stage by choosing *shirk*. Since all other agents play *work* under $\mathbf{s}^*_{-i}$, and since $b_i^c > \underline{b}_i$, the deviation yields strictly lower payoff. Second, suppose agent i deviates to choose *monopolize*. In this case, all other agents respond by shirking. Whether agent i works or shirks, his expected payoff is strictly below that from playing s_i^* by constraint (IC).

In both cases, s_i^* delivers a strictly higher payoff than any deviation against $\mathbf{s}^*_{-i}$. Hence, it is the unique best response. □

The next part shows that working after being monopolized cannot survive IEWDS. If such a strategy, call it s_i, were to survive, someone else would find it profitable to exploit this *work* response, making s_i strictly worse than a *shirk* alternative. Then s_i must have been eliminated under IEWDS, leading to a contradiction.

Lemma 2. *Any strategy that plays work after being monopolized will not survive IEWDS.*

Proof of Lemma 2. We begin with a useful observation, which follows directly from the definition of IEWDS and will be used repeatedly:

Claim 1. *A strategy s_i will not survive if there exists a surviving strategy profile $\mathbf{s}'$ and an alternative (possibly mixed) strategy $\sigma_i \neq s_i$ such that (1) s_i yields weakly lower payoffs than σ_i against any possibly surviving strategy profiles; and (2) $u_i(s_i, \mathbf{s}'_{-i}) < u_i(\sigma_i, \mathbf{s}'_{-i})$.*[7]

[7]Even if σ_i, or any strategy in the support of σ_i, is deleted, we can find a surviving strategy for agent

Suppose by contradiction that there exists a surviving strategy s_i which plays *work* after a message profile that allows agent $j \neq i$ to monopolize, i.e., a message profile in which agent j is the highest-ranked monopolizing speaker. When facing the surviving profile $(s_i, \mathbf{s}^*_{-ij})$, *monopolize* gives agent j a continuation payoff strictly higher than any collaborating payoff because $\max\{P(\{i,j\})b^m_j - c_j, P(\{i\})b^m_j\} > P(\mathbb{N})b^c_j - c_j$ by (SC2). Hence, at least one strategy with $m_j = monopolize$, say s_j, cannot be eliminated and survives to the end.

However, when $m_j = monopolize$ survives, the *work* response from agent i should not survive. Formally, consider an alternative strategy that differs from s_i only in that it plays *shirk* after being monopolized. This strategy weakly dominates s_i and gives a strictly higher payoff when facing this surviving strategy profile $(s_j, \mathbf{s}^*_{-ij})$, in which agent i is monopolized. By Claim 1, s_i will not survive. □

The final step establishes the uniqueness. We begin by showing that the costly signal *collaborate* is credible. The signaling constraint (SC1) ensures that *collaborate and then shirk* is weakly dominated by *monopolize.* In particular, it is strictly worse when facing a surviving strategy profile in which all other agents collaborate and then work. Hence, it cannot survive regardless of the elimination order.

Once we rule out *collaborate and then shirk* for all speakers, the nonspeaker will work when all speakers choose *collaborate.* Then, starting from the lowest-ranked speaker, each of them anticipates that (1) all lower-ranked agents will collaborate and then work, and (2) his own message only matters if all higher-ranked speakers choose *collaborate.* As such, each speaker faces a similar trade-off as in the two-agent case and strictly prefers *collaborate and then work* over *monopolize.*

Proof of Theorem 1. Following Lemma 2, we restrict attention to strategies in which agents shirk after being monopolized.

Step 1 (Collaborate and then shirk will not survive). Fix a speaker i. If any higher-ranked speaker chooses *monopolize*, all strategies yield identical payoffs. Now suppose all

i that weakly dominates (each of) them accordingly among the set of possibly surviving strategy profiles. Given $\mathbf{s}'$ survives and $u_i(s_i, \mathbf{s}'_{-i}) < u_i(\sigma_i, \mathbf{s}'_{-i})$, there exists a surviving strategy that weakly dominates s_i. So s_i cannot survive.

higher-ranked speakers choose *collaborate*, but at least one lower-ranked speaker chooses *monopolize*. By constraint (SC1), choosing *monopolize* guarantees agent i a strictly positive payoff, while *collaborate* results in zero.

Next, consider the case where all other speakers choose *collaborate* in $\mathbf{s}_{-i}$, but only a subset $K \subseteq \mathbb{N} \setminus \{i\}$ of agents (including the nonspeaker) choose *work* in the second stage. In this case, the expected payoff from playing *collaborate and then shirk* is $P(K)b_i^c$, while *monopolize* yields strictly more than $P(\mathbb{N}\setminus\{i\})b_i^c$ by (SC1). By monotonicity of $P(\cdot)$, we have $P(\mathbb{N} \setminus \{i\}) \geq P(K)$, so at least one strategy with a *monopolize* message weakly dominates *collaborate and then shirk*, and is strictly better against the surviving profile $\mathbf{s}^*_{-i}$, in which all other agents choose *collaborate and then work*.

Therefore, by Claim 1, the strategy *collaborate and then shirk* cannot survive.

Step 2 (Uniqueness of $\mathbf{s}^*$). Once we rule out all strategies that play *collaborate and then shirk*, the strategy s^*_N, in which the nonspeaker chooses *work* following a unanimous *collaborate* message profile, is the unique surviving strategy for the nonspeaker.

We now show that s^*_i is the unique surviving strategy for each speaker $i < N$, using a "backward" induction argument. Assume the induction hypothesis: for every agent $j \in \{i+1, \ldots, N\}$, s^*_j is the unique surviving strategy. We will show that s^*_i will be the unique surviving strategy for agent i.

For each speaker $i < N$, a possibly surviving strategy different from s^*_i must have $m_i =$ *monopolize*. Denote one such strategy by s_i. First, when facing the surviving profile $\mathbf{s}^*_{-i}$, s_i gives strictly lower payoff than s^*_i by constraint (SC1). Second, s_i deliver lower payoffs than s^*_i against any possibly surviving opponent strategies:

- If any higher-ranked speaker chooses *monopolize*, then both s_i and s^*_i yield zero payoff;
- Suppose all higher-ranked speakers choose *collaborate*. By the induction hypothesis, all lower-ranked agents must choose *collaborate*. Then, by Step 1, all opponents work after this unanimous *collaborate* message. Hence, we face the same situation as in $\mathbf{s}^*_{-i}$ and s_i gives strictly lower payoff than s^*_i.

By Claim 1, s_i will not survive. It follows that s^*_i is the unique surviving strategy for agent i. We finish the induction argument. □

4.2 Efficient Coordination and Discussion

By Theorem 1, for all $\varepsilon > 0$, the collaborating mechanism can uniquely implement full effort and bonus allocation $\mathbf{b} = (\underline{b}_i + \frac{c_i}{\sum_i c_i}\varepsilon)_{i \in \mathbb{N}}$ under IEWDS. The total bonus $(\sum_i \underline{b}_i) + \varepsilon$ approaches arbitrarily close to the second-best level. The bonus allocation is non-discriminatory: the on-path markup for agent i, $\lim_{\varepsilon \to 0} \frac{b_i - c_i}{c_i} = \frac{1}{P(\mathbb{N}) - P(\mathbb{N} \setminus \{i\})} - 1$, varies nontrivially only with the agent's marginal contribution to team success. We discuss two key properties of this second-best mechanism below.

The Monopoly Option as a Signaling Device

How can speakers credibly promise to work hard in the future? They can do so by making an upfront, public, and costly decision: forgoing the monopoly option. This act of signaling makes it possible to coordinate effort under the approximately second-best bonus allocation.

This intuition echoes earlier work (e.g., Kohlberg and Mertens (1986), Van Damme (1989), Ben-Porath and Dekel (1992)) in how forward induction selects equilibria in coordination games. To see this, we return to the two-agent example and consider an outside option that directly assigns some payoff to the speaker. As long as *all agents work* is a strict Nash equilibrium under our target bonus allocation, we can construct an outside option that offers less than the *all agents work* outcome, but more than what an agent would receive from unilaterally deviating to shirk (see Table 2 for a formal representation). The extended normal-form game will then have a unique surviving outcome where all agents choose to work under the target bonus allocation.

The specific structure of this outside option and the subsequent effort decisions can be quite flexible. For instance, a "who-to-team-up" variant of the mechanism also works: if any speaker chooses *monopolize*, the highest-ranked among them proceeds alone and decides whether to work or shirk, while the principal excludes all other agents from the project, instead of allowing them to stay and receive zero bonus. As before, the principal can fine-tune the monopoly bonus to achieve the second-best result.

	s_2		
s_1		w	s
	collaborate, w	$b_1^c - c_1, b_2^c - c_2$	$pb_1^c - c_1, pb_2^c$
	collaborate, s	$pb_1^c, pb_2^c - c_2$	$p^2 b_1^c, p^2 b_2^c$
	option	$z_1, 0$	$z_1, 0$

Note: The table presents the (simplified) reduced normal form game when agent 1 is offered an option that directly pays her $z_i > 0$. Effort decisions are denoted by $w(ork)$ and $s(hirk)$. $((collaborate, w), w)$ is the unique outcome under IEWDS if $pb_1^c < z_1 < b_1^c - c_1$.

Table 2: An abstract outside option in the two-agent example

Voice Hierarchy

Extending the two-agent result to a multi-agent setting is initially far from straightforward. When multiple speakers choose *monopolize*, who should be granted that right? Worse yet, when many agents try to signal their intentions simultaneously, coordination can break down. As Ben-Porath and Dekel (1992) show, simultaneous signaling in their game may lead to an outcome in which players burn money without achieving cooperation. While Hurkens (1996) tackles this issue with a stronger solution concept, I twist the signaling procedure.

Recall that in our mechanism, agents speak simultaneously in the contracting stage. The trouble with everyone talking at once motivates a hierarchical structure of our speaking rules. Once an agent speaks, they concede the monopoly right to the next in line. Such a voice hierarchy turns out to prevent the "voice clash" that arises in simultaneous signaling. The order in which agents speak is quite flexible: unlike Cavounidis and Park (2025), it does not matter who speaks or whose voice ranks higher when agents are asymmetric.

A sequential procedure yields the same results if we let agents speak in turn, and all messages are publicly observed. Note that in the voice hierarchy, a *monopolize* message takes effect only if no higher-ranked speaker chooses *monopolize*. This structure lends itself naturally to a sequential mechanism. Figure 2 illustrates the game tree induced by the sequential mechanism.

While Ben-Porath and Dekel (1992) also see the potential of a sequential signaling procedure, their approach does not support pay transparency. They require each agent to observe

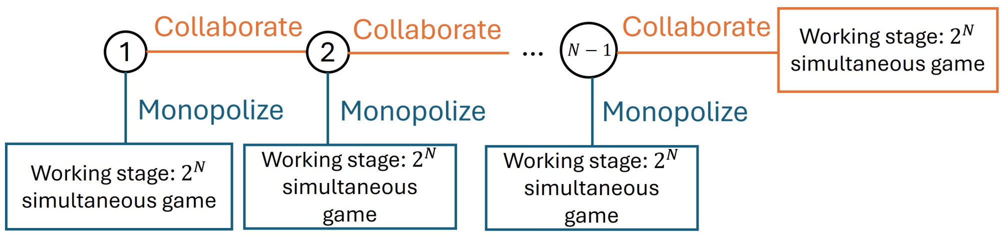

Figure 2: The game tree induced by a sequential *collaborating* mechanism

only prior messages and "leave the scene" after sending their own. My approach differs: everyone sees all messages. In my setting, the principal has more flexibility in designing the off-path subgame when the outside option is chosen. It becomes unnecessary to hide lower-ranked or subsequent messages. As a result, in contrast to other second-best mechanisms (e.g., Halac et al. (2021), Cavounidis and Ghosh (2021)), such transparency lets agents confirm their own contractual terms before deciding whether to work.

5 Examples of Applications

This paper introduces a simple, public mechanism that ensures all agents work under the approximately second-best bonus allocation. The mechanism invites agents to send costly messages to signal their future efforts before the team project starts. Both the messages and the resulting bonus allocation are public information. In this way, contracting with public negotiation procedures serves as a powerful tool to resolve coordination problems in team projects and combat pay discrimination. Several real-life examples and applications are discussed below.

Fee Splitting Among Lawyers

Public bonus-splitting arrangements are common in legal practice, particularly among lawyers from different firms who collaborate on the same case (Grossbaum, 2022). These collaborations often involve contingent fee matters, where payment depends in part on a successful outcome. To avoid disputes, the American Bar Association's Model Rule 1.5(e) requires a written agreement: both the lawyers and the client must consent in advance to the specific share each lawyer will receive. Some jurisdictions go further. For example, Florida Bar rules specify a presumptive 75/25 split between primary and secondary lawyers in contingent fee cases, while allowing equal division by mutual agreement. These legal requirements help institutionalize transparent, pre-work bonus-splitting contracts, which help align expectations and reduce coordination failures in team-based legal work.

Credit Allocation in Academia

Discussing authorship order early in a collaboration can also function as a bonus-splitting contract. In economics, many papers adopt alphabetical authorship and, more recently, a certified random authorship to signal equal contributions. Others adopt a non-alphabetical order in which the first author often receives more credit (Laband and Tollison, 2000; Einav and Yariv, 2006; Waltman, 2012).

My paper suggests that the option to choose the authorship rule itself can play a crucial role in incentivizing effort in academic teams. To illustrate this, consider the following simple model of credit allocation. Two agents start a project and expect a total credit of $S \in (0, 1)$ if the project succeeds. Agent 1 brings up the initial idea. She faces a choice: suggest a random order, earning equal credit ($\frac{1}{2}S$ each), or propose a non-alphabetical order that gives her full credit (S) and leaves agent 2 with none. Her goal is to motivate both agents to work. All other aspects of the model follow the setup in Section 2. To ensure that *both agents work* in the unique outcome surviving IEWDS, two constraints must hold:

$$\max\{p \cdot S - c, p^2 \cdot S\} > p \cdot (\frac{1}{2}S)$$
$$\max\{p \cdot S - c, p^2 \cdot S\} < \frac{1}{2}S - c$$

These inequalities hold if and only if $S > \frac{c}{p(1-p)}$ and $p < \frac{1}{2} < 1 - \frac{c}{pS}$.

So, authorship choice can serve as an effective coordination device under two conditions. First, the total credit must be large enough to create room for signaling. Second, the equal allocation of credit resulting from the random order must strike a balance. Given that agent 2 works under equal credit but shirks otherwise, the share $\frac{1}{2}$ must be large enough to discourage agent 1 from taking all the credit (incentive constraint), yet small enough that working under full credit remains more profitable than shirking under half (signaling constraint).

Subcontracting with Partial Centralization

Subcontracting offers another real-life example of public, interdependent contracting. Cavounidis and Park (2025) studies a fully decentralized version in which the subcontractor freely determines how to split a fixed bonus budget. Building on that, my analysis suggests that introducing partial centralization, such as imposing a cap or a floor on how much the subcontractor can claim, might further bring down the cost for the principal.

Consider again the two-agent example from Section 2. In a fully decentralized subcontracting mechanism, the principal sets a fixed bonus budget and delegates agent 1 as the subcontractor, who then decides how to split the bonus budget. Regardless of the subcontractor's choice, the sum of bonuses remains fixed. It implies that, to uniquely implement the approximately second-best bonus allocation, we must have (ignoring the ε's in the contract)

$$b_1^c + b_2^c = b_1^m + 0 \Leftrightarrow \frac{2c}{1-p} = \frac{c}{p(1-p)}.$$

The equality holds only when $p = \frac{1}{2}$. Hence, full decentralization can achieve the second-best benchmark only in this knife-edge case.

What happens if we allow partial centralization, limiting the subcontractor's choices? Suppose first that $p > \frac{1}{2}$. Then, $\frac{2c}{1-p} > \frac{c}{p(1-p)}$: the second-best budget required to motivate both agents is so generous that the subcontractor strictly prefers to claim the entire amount. A simple remedy is to impose a cap on the subcontractor's share, roughly $\frac{c}{p(1-p)}$, to induce sharing.[8] When $p < \frac{1}{2}$, the situation reverses: $\frac{2c}{1-p} < \frac{c}{p(1-p)}$. The second-best budget is insufficient to signal the subcontractor's effort decision. Setting a cap no longer helps. It is unclear whether a natural extension of the subcontracting mechanism can achieve the second-best in this case.

[8] If the subcontractor can choose any bonus below the cap, a technical issue will arise due to the continuous choice set. To uniquely implement full effort and the (approximately) second-best bonus allocation under IEWDS, the principal can also impose a floor, for example, at $\frac{2c}{1-p} + \varepsilon$.

Appendix A Second-Best Benchmark

We prove a slightly more general statement: no mechanism can induce all agents to work with an expected total bonus weakly lower than the *second-best* level in any outcomes that survive IEWDS.

Suppose by contradiction that there exists a mechanism that induces all agents to work with an expected total bonus weakly lower than the *second-best* level in any outcomes that survive IEWDS. For each agent i, let x_i capture the part of her strategy when participating in the mechanism, ϕ_i all the (realized) information from the mechanism, and $a_i(x_i, \phi_i)$ her effort decision rule if she stays. Then we can write the strategy of agent i as $s_i = (x_i, a_i(\cdot))$ in the induced normal-form game.

For each agent i, let S_i^w be the set of strategies that involve working after receiving an expected bonus offer $E(b_i|\phi_i) \leq \underline{b}_i = \frac{c_i}{P(\mathbb{N})-P(\mathbb{N}/\{i\})}$ upon team success on some paths. Formally,

$$S_i^w = \{s_i | a_i(x_i, \phi_i) = 1 \text{ for some } (x_i, \phi_i) \text{ that leads to } E(b_i|\phi_i) \leq \underline{b}_i\}.$$

For each strategy $s_i \in S_i^w$, we can construct an alternative strategy, denoted by $\hat{s}_i$, which replaces *work* by *shirk* on the paths whenever s_i plays *work* after receiving an expected bonus $E(b_i|\phi_i) \leq \underline{b}_i$. More precisely, $\hat{s}_i$ plays the same x_i in the mechanism and exerts the same effort $a_i(x_i, \phi_i)$ whenever $E(b_i|\phi_i) > \underline{b}_i$, but plays *shirk*, $a_i(x_i, \phi_i) = 0$, whenever $E(b_i|\phi_i) \leq \underline{b}_i$. Denote this many-to-one mapping by $h : s_i \mapsto \hat{s}_i$.

Next, we show that s_i is weakly dominated by $\hat{s}_i = h(s_i)$. Fix any strategy profile from opponents, $\mathbf{s}_{-i}$. $\hat{s}_i$ and s_i lead to the same expected bonus $E(b_i|\phi_i)$ because they differ only in effort decisions. If agent i receives an expected bonus $E(b_i|\phi_i) \leq \underline{b}_i$ on path, *shirk* gives a weakly higher payoff than *work*. If instead $E(b_i|\phi_i) > \underline{b}_i$ on path, $\hat{s}_i$ predicts the same effort and thus gives the same payoff as s_i.

Now consider an IEWDS procedure in which we always delete strategies that play *work* after an agent gets an expected bonus less than the second-best level, before we delete those playing *shirk* instead. That is, whenever we can and want to delete a 'shirking' strategy $\hat{s}_i \in h(S_i^w)$, we replace it with a 'working' strategy s_i in the set $h^{-1}(\hat{s}_i)$. This is feasible because any strategy s_i in $h^{-1}(\hat{s}_i)$ is weakly dominated by $\hat{s}_i$. Since h is many-to-one, for any x_i and ϕ_i that lead to $E(b_i|\phi_i) \leq \underline{b}_i$, we must delete all 'working' strategies s_i with $a_i(x_i, \phi_i) = 1$, before deleting 'shirking' strategies $\hat{s}_i$ with $\hat{a}_i(x_i, \phi_i) = 0$.

If no outcome survives under the procedure, the mechanism fails, leading to a contradiction. If the surviving set is nonempty, all agents must work on path in any surviving outcome by assumption. Then the way we construct the procedure tells us that all agents must receive an expected bonus strictly higher than $\underline{b}_i$. Otherwise, if some IEWDS procedure induces a surviving outcome in which agent i receives $E(b_i|\phi_i) \leq \underline{b}_i$ and plays *work* on path, we must have deleted the strategy $\hat{s}_i = h(s_i)$ that plays *shirk* on the same path. This contradicts how we construct the IEWDS procedure. However, if all agents receive an expected bonus $E(b_i|\phi_i) > \underline{b}_i$ in any surviving outcome, the expected total cost must exceed the *second-best* level. This leads to a contradiction.

References

Antić, N. and N. Persico (2023). Equilibrium selection through forward induction in cheap talk games. *Games and Economic Behavior 138*, 299–310.

Arya, A., J. Glover, and J. S. Hughes (1997). Implementing coordinated team play. *Journal of Economic Theory 74*(1), 218–232.

Battigalli, P. (1997). On Rationalizability in Extensive Games. *Journal of Economic Theory 74*(1), 40–61.

Ben-Porath, E. and E. Dekel (1992). Signaling future actions and the potential for sacrifice. *Journal of Economic Theory 57*(1), 36–51.

Cavounidis, C. and S. Ghosh (2021). Losing by betting on success. *Working Paper*.

Cavounidis, C. and H. Park (2025). Subcontracting Away Strategic Uncertainty - Contracts from the Horse's Mouth. *Working Paper*.

Cullen, Z. (2024). Is pay transparency good? *Journal of Economic Perspectives 38*(1), 153–80.

Demski, J. S. and D. Sappington (1984). Optimal incentive contracts with multiple agents. *Journal of Economic Theory 33*(1), 152–171.

Einav, L. and L. Yariv (2006). What's in a surname? The effects of surname initials on academic success. *Journal of Economic Perspectives 20*(1), 175–188.

Garicano, L., T. Van Zandt, et al. (2013). Hierarchies and the division of labor. *Handbook of Organizational Economics*, 604–654.

Genicot, G. and D. Ray (2006). Contracts and externalities: How things fall apart. *Journal of Economic Theory 131*(1), 71–100.

Glover, J. (1994). A simpler mechanism that stops agents from cheating. *Journal of Economic Theory 62*(1), 221–229.

Grossbaum, D. (2022). Lessons from Lawyer Fee-Sharing Agreements Gone Wrong. Available at `https://www.hinshawlaw.com/newsroom-newsletters-lln-sharing-fees-between-lawyers.html`.

Halac, M., E. Lipnowski, and D. Rappoport (2021). Rank Uncertainty in Organizations. *The American Economic Review 111*(3), 757–786.

Hurkens, S. (1996). Multi-sided Pre-play Communication by Burning Money. *Journal of Economic Theory 69*(1), 186–197.

Kapon, S., L. Del Carpio, and S. Chassang (2024). Using divide-and-conquer to improve tax collection. *The Quarterly Journal of Economics 139*(4), 2475–2523.

Kohlberg, E. and J.-F. Mertens (1986). On the strategic stability of equilibria. *Econometrica 54*(5), 1003–1037.

Laband, D. N. and R. D. Tollison (2000). Intellectual collaboration. *Journal of Political Economy 108*(3), 632–662.

Ma, C.-T. (1988). Unique implementation of incentive contracts with many agents. *The Review of Economic Studies 55*(4), 555–572.

Ma, C.-T., J. Moore, and S. Turnbull (1988). Stopping agents from "cheating". *Journal of Economic Theory 46*(2), 355–372.

Mookherjee, D. (2006). Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective. *Journal of Economic Literature 44*(2), 367–390.

Mookherjee, D. and S. Reichelstein (1990). Implementation via augmented revelation mechanisms. *The Review of Economic Studies 57*(3), 453–475.

Pearce, D. G. (1984). Rationalizable Strategic Behavior and the Problem of Perfection. *Econometrica 52*(4), 1029.

Segal, I. (2003). Coordination and discrimination in contracting with externalities: divide and conquer? *Journal of Economic Theory 113*(2), 147–181.

Shimoji, M. (2002). On forward induction in money-burning games. *Economic Theory 19*(3), 637–648.

Shimoji, M. (2004). On the equivalence of weak dominance and sequential best response. *Games and Economic Behavior 48*(2), 385–402.

Van Damme, E. (1989). Stable equilibria and forward induction. *Journal of Economic Theory 48*(2), 476–496.

Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics 6*(4), 700–711.

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