Submitted to *Theoretical Economics*

Robust Performance Evaluation of Independent Agents

ASHWIN KAMBHAMPATI
Department of Economics, United States Naval Academy

A principal provides incentives for independent agents. The principal cannot observe the agents' actions, nor does she know the entire set of actions available to them. It is shown that an anti-informativeness principle holds: very generally, robustly optimal contracts must link the incentive pay of the agents. In symmetric and binary environments, they must exhibit *joint performance evaluation* — each agent's pay is increasing in the performance of the other, with the degree of this increase depending on individual performance.

KEYWORDS. moral hazard, robustness, teams.
JEL CLASSIFICATION. D81, D82, D86.

1. INTRODUCTION

Should members of a group be compensated on the basis of individual performance, relative to the performance of others, or jointly? Conventional economic wisdom builds upon the "Informativeness Principle", which states that

Ashwin Kambhampati: kambhamp@usna.edu

Earlier versions have circulated under the titles "Robust Performance Evaluation" and "Robust Performance Evaluation of Independent and Identical Agents". An earlier draft appears as the first chapter of my dissertation at the University of Pennsylvania and an abstract appears in *EC '21: Proceedings of the 22nd ACM Conference on Economics and Computation.* I thank Nageeb Ali, Aislinn Bohren, Gabriel Carroll, George Mailath, Steven Matthews, Juuso Toikka, and Rakesh Vohra for their time, guidance, and encouragement. I thank two anonymous referees and several seminar and conference participants for helpful comments and suggestions. The views expressed here are solely my own and do not in any way represent the views of the U.S. Naval Academy, U.S. Navy, or the Department of Defense.

only signals that are statistically informative about an agent's action are valuable for incentive provision (Holmström (1979), Shavell (1979)). Hence, encouraging competition or cooperation among multiple agents through *relative performance evaluation* or *joint performance evaluation* is more profitable than *independent performance evaluation* only if these compensation schemes better extract information (Holmström (1982)).

In practice, however, joint performance evaluation often appears in settings in which statistical considerations suggest it is suboptimal. For instance, salespeople may make sales calls alone. Yet, they are sometimes compensated according to increasing and nonlinear functions of individual and group revenue (Rees et al. (2003)). Members of the same company may perform independent tasks. Yet, companies often use bonus pools to distribute reward pay, with the size of the pool determined by the company's overall performance (Benson and Sajjadiani (2018)). Finally, CEO compensation often exhibits "pay for luck", with remuneration increasing with random positive shocks (Bell et al. (2021)). In these settings, independent performance evaluation might be used if there is insufficient correlation in productivities across agents conditional on their actions. Otherwise, relative performance evaluation better extracts information than joint performance evaluation; success in the face of others' failure is a stronger indicator of effort than success when others succeed.

This paper provides non-Bayesian foundations for joint performance evaluation in such settings. The production environment considered is standard. There is a risk-neutral principal who compensates a finite number of risk-neutral agents. Each agent takes a hidden action to produce observable individual output. The agents are protected by limited liability. Hence, "selling the firm" to the agents is infeasible and there is a trade-off between incentive provision and rent-extraction (see, e.g., Chapter 4 of Laffont and Martimort (2009)).

Agents are commonly known to be independent; there is no correlation in output conditional on agents' actions and no agent can directly affect the output of any of the others. The purpose of the assumption is to rule out all known mechanisms leading to interdependent incentive schemes, which rely on productive or informational linkages across agents (see Section 1.1 for a detailed discussion). It is also a reasonable first-order approximation of the production environment of many economic agents (e.g., salespeople, teachers, and fruit pickers).[1] To derive foundations for interdependent incentive schemes, it is instead assumed that, while the principal knows some actions the agents can take, there may be others she does not know about.[2] Following Carroll (2015), the principal chooses a contract that ensures her the highest worst-case payoff.

The first main result is an “anti"-informativeness principle: robustly optimal contracts *must* link the incentive pay of independent agents under a wide range of assumptions about agents' behavior. Moreover, the robustly optimal independent performance evaluation contract is outperformed by a joint performance evaluation contract that resembles a bonus pool (Theorem 1 and Corollary 1). The intuition is as follows. Suppose there are two agents and each receives a positive wage for individual success. Now, suppose each agent's wage for individual success is made contingent upon the other's success. Specifically, reduce his wage when the other fails and increase it when the other succeeds so that, when the other takes his targeted action, expected wages are held constant. Then, incentives to shirk are also held constant. Nevertheless, when both agents actually shirk — the “bad" state of the world that matters under robustness considerations — the principal pays each strictly less than under independent performance evaluation (see Example 1).

The second main result is that joint incentives are uniquely optimal in a canonical partial-implementation setting in which there are two agents with the same set of known actions, two output levels, and the principal is constrained to use symmetric contracts (Theorem 2). Interestingly, there does not exist a relative performance evaluation contract that yields the principal a strictly larger payoff than the optimal independent performance evaluation contract (Lemma 2). In contrast to joint performance evaluation contracts, these contracts increase

[1] See Hackman (2002) pages 42-43 for additional examples.

[2] It is worth remarking that there are no symmetry assumptions made on the set of uncertainty; actions need not be common across agents. An earlier version of this paper analyzes the case in which the principal knows that the agents possess a common action set.

expected wage payments when agents shirk, eliminating any of their incentive advantages (see Example 2).

The results provide a plausible explanation for several of the joint incentive schemes mentioned in the motivating applications. For instance, in a sales context, group managers may know some tactics of their sales representatives — representatives can always follow the company's script. But, there are a myriad of less costly (but, potentially less productive) ways in which a sales representative might deviate from this script. Thus, a manager might use team-based incentive pay to reduce expected wage payments if her subordinates discover such tactics. Regarding bonus pools, empirical work has found that these schemes are not particularly beneficial for firm productivity (Benson and Sajjadiani (2018)). But the analysis in this paper shows that this is not a requirement for their *profitability*; it is only important that the size of the pool is sufficiently responsive to firm performance. The results also contribute to the debate surrounding executive compensation. Within a firm, it may make sense to compensate all executives on the basis of the firm's overall performance. But, to capture the rent-extraction benefits of joint performance evaluation, it is important that this pay responds as much to declines in performance as it does to increases in performance. This is not always the case in practice, as shown by Bell et al. (2021).[3]

1.1 *Related Literature*

This paper makes two main contributions to the theoretical literature. First, it establishes a fundamentally new justification for team-based incentive pay. In the Bayesian contracting paradigm, the Informativeness Principle prescribes independent performance evaluation whenever one agent's performance is statistically uninformative of another's action. Hence, if the set of actions available to agents in a team is common knowledge, then it is impossible to improve upon independent performance evaluation. To justify incentive schemes commonly

[3]Members of a board of directors may also wish to tie the pay of a CEO to the performance of other firms in the same industry. But, to capture the benefits of joint incentives, it is important for members of the board to also control the incentives of the CEOs of these other firms. This type of ownership structure is typically prohibited by conflict-of-interest regulations.

used in practice, such as relative performance evaluation and joint performance evaluation, the literature has instead introduced productive and informational linkages among agents.[4] Specifically, one agent's action either has a direct effect on another's performance or there is correlation in performances conditional on an action profile. The model studied in this paper explicitly rules out these channels.

Second, it contributes to the literature on robust contracting by considering a multi-agent environment in which the principal's uncertainty set is bounded.[5] The pioneering work of Carroll (2015) considers a principal-single agent model in which the principal has non-quantifiable uncertainty about the actions available to the agent. His main result is that there exists a robustly optimal contract that is linear in individual output. The model and analysis in this paper enrich that of Carroll (2015) by introducing seemingly irrelevant agents and showing that multiple agents lead to the optimality of joint incentive schemes.[6]

Dai and Toikka (2022) extend the analysis of Carroll (2015) to multi-agent settings, but consider a model in which the principal deems *any* game the agents might be playing plausible. In this setting, they find that contracts that are linear in team output are worst-case optimal under partial Nash implementation.

[4] In the absence of productive interaction, joint performance evaluation may be optimal if agents are affected by a common, negatively correlated productivity shock (Fleckinger (2012)). In the absence of a common shock, joint performance evaluation may be optimal if efforts are complements in production (Alchian and Demsetz (1972)), if it induces help between agents (Itoh (1991)) or, alternatively, if it discourages sabotage (Lazear (1989)). Finally, joint performance evaluation may be optimal if agents are engaged in repeated production and it allows for more effective peer sanctioning (Che and Yoo (2001)).

[5] Related work not discussed here include the papers of Hurwicz and Shapiro (1978), Garrett (2014), Frankel (2014), and Rosenthal (2020), who consider contracting with unknown preferences; Marku et al. (2023), who consider a robust common agency problem; and Chassang (2013), who considers a dynamic agency problem.

[6] Building upon Carroll (2015)'s single-agent model, Antic (2015) imposes bounds on the principal's uncertainty over the productivity of unknown actions (see also Section 3.1 of Carroll (2015), which studies lower bounds on costs). In contrast, the model studied here places no restrictions on the technology available to each agent in isolation beyond those of Carroll (2015). Instead, the restrictions concern the relationship between the agents.

This result is driven by the finding that any contract that induces competition between agents is non-robust to games in which one agent's action can directly influence the productivity of another. In contrast to Dai and Toikka (2022), this paper considers a setting in which the principal *knows* that output is independently distributed across agents. This has the immediate effect of ruling out such games and ensuring that linear contracts are suboptimal in performance indicators. Despite these differences, the results and management implications of this paper complement Dai and Toikka (2022). Agents in Dai and Toikka (2022)'s model are a "real team" in the sense that they work together to produce value for the principal, while agents in the model of this paper are best thought of as "co-actors" given the assumption of technological independence (Hackman (2002)). Yet, in either case, joint performance evaluation is optimal. What changes is the particular form of the optimal joint performance evaluation contract — in the case of a real team, optimal compensation is always linear in the value the team generates for the principal, while in the case of co-acting agents it is always non-linear and may involve bonus payments that reward each agent for others' success in a manner proportional to their individual contribution.

2. Model

2.1 *Environment*

A risk-neutral principal writes a contract for risk-neutral agents, indexed by $i = 1, 2, ..., n$. Agent i's output, y_i, is observable and belongs to a compact set $Y \subset \mathbb{R}_+$, where $\max(Y) > \min(Y) = 0$. To produce output, agent i chooses an unobservable action, a_i, from a finite set $A_i \subset \mathbb{R}_+ \times \Delta(Y)$, where $\Delta(Y)$ is the set of Borel distributions on Y. Each action a_i is thus identified by an effort cost, $c(a_i) \in \mathbb{R}_+$, and a distribution over output, $F(a_i) \in \Delta(Y)$. Agents are assumed to be independent — there are no informational or productive linkages across agents. Formally, the joint distribution over output vectors induced by any action profile is the product of the marginal distributions over individual outputs,

$$F(a) := F(a_1) \times \cdots \times F(a_n) \in \Delta(Y^n) \quad \text{for all } a \in A := A_1 \times \cdots \times A_n.$$

2.2 *Contracts*

A **contract** is a function for each agent i,

$$w_i : Y^n \to \mathbb{R}_+,$$

where the non-negativity restriction in the co-domain reflects agent limited liability (no agent can receive negative wages). Direct (revelation) mechanisms[7] and random mechanisms[8] are thus ruled out by assumption. It will be useful to classify contracts according to an extension of the typology of Che and Yoo (2001), who consider binary performance evaluations.

DEFINITION 1 (Performance Evaluations). A contract $w = (w_i)_i$ is an

- **independent performance evaluation (IPE)** if, for all i and y_i, $w_i(y_i, y_{-i})$ is constant in y_{-i};
- a **relative performance evaluation (RPE)** if it does not exhibit IPE and, for all i and y_i, $w_i(y_i, y_{-i})$ is decreasing in y_{-i};
- and a **joint performance evaluation (JPE)** if it does not exhibit IPE and, for all i and y_i, $w_i(y_i, y_{-i})$ is increasing in y_{-i} .[9]

[7] It is well known that the principal can partially implement the Bayesian optimal contract technology-by-technology using a revelation mechanism: she can ask agents to report the action set and, if reports disagree, punish them with a contract that always pays zero. The interpretation taken in this paper, however, and in the rest of the literature on robust contracting, is that such a mechanism violates the spirit of the robustness exercise. The principal would like to avoid changing the contract she offers as the agents' environment varies. The performance of alternative indirect mechanisms, such as offering a menu of contracts, awaits further study.

[8] Randomizing over contracts is not helpful if the principal believes that Nature selects the agents' action set after her contract is realized. However, if Nature moves simultaneously, then there is scope for randomization to improve the principal's payoff. See Kambhampati (2023) for an analysis of the single-agent case.

[9] Output vectors are equipped with the usual partial order: $y' \geq y$ if y' is weakly larger than y in all components. So, a function of output vectors, f, is increasing if $y' \geq y$ implies $f(y') \geq f(y)$.

2.3 *Payoffs*

Agent i's ex post payoff given a contract w, action profile a, and output vector y is

$$w_i(y) - c(a_i),$$

while his expected payoff is

$$U_i(a;w) := \mathbb{E}_{F(a)}[w_i(y)] - c(a_i).$$

The principal's ex post payoff given a contract w and output vector y is

$$\sum_{i=1}^{n} (y_i - w_i(y)) .$$

Given a contract w and action set A, the first part of the analysis assumes only that the principal believes that the agents' behavior is consistent with common knowledge of rationality. That is, she assumes play of a (correlated) rationalizabile action profile.[10] Let $\mathcal{R}(w,A) \subseteq \Delta(A)$ be the set of Borel distributions over rationalizable action profiles in the game $\Gamma(w,A)$. Then, the non-empty set of expected payoffs obtainable under some distribution over rationalizable action profiles is

$$V(w,A) := \{\mathbb{E}_\sigma[\sum_{i=1}^{n} (y_i - w_i(y))] : \sigma \in \mathcal{R}(w,A)\}.$$

2.4 *Uncertainty*

When the principal writes a contract for the agents, she has limited knowledge about the action set available to each agent. In particular, she knows only a non-empty subset of actions available to each agent $A_i^0 \subseteq A_i$. To rule out uninteresting cases, it is assumed that, for each agent i, there exists an action $a_i^0 \in A_i^0$ such that $\mathbb{E}_{F(a_i^0)}[y] - c(a_i^0) > 0$. This ensures that the principal obtains a strictly positive payoff from contracting with agent i. In addition, it is assumed that if $a_i^0 \in A_i^0$, then $c(a_i^0) > 0$. This ensures that the principal's optimal contract for

[10] Correlated rationalizable action profiles are those obtained by iterated elimination of strictly dominated actions. See Brandenburger and Dekel (1987).

agent i is different from one that always pays zero. In the face of her uncertainty, the principal evaluates each contract on the basis of its performance across all finite supersets of her knowledge, collected in the feasible set of uncertainty $\mathcal{A} := \{A \subset (\mathbb{R}_+ \times \Delta(Y))^n : |A| < \infty \text{ and } A_i \supseteq A_i^0\}$.

3. Anti-Informativeness Principle

Let

$$v_{IPE} := \max_{w\colon w \text{ is an IPE}} \inf_{A \in \mathcal{A}} \max(V(w, A)) > 0$$

be the principal's highest payoff obtainable under rationalizable behavior given any IPE contract.[11] The first main result is that there exists a JPE contract whose rationalizable payoffs robustly dominate those obtained under any IPE.

THEOREM 1 (Anti-Informativeness Principle). *For each agent i, there exists a base share, $\phi_i > 0$, and bonus factor, $\beta_i > 0$, such that the JPE contract, w, with*

$$w_i(y_i, y_{-i}) = (\phi_i + \frac{\beta_i}{n-1} \sum_{j \neq i}^{n} y_j) y_i \quad \textit{for } i = 1, \dots, n$$

has rationalizable payoffs that robustly dominate those obtained under any IPE:

$$\inf_{A \in \mathcal{A}} \max(V(w, A)) > \inf_{A \in \mathcal{A}} \min(V(w, A)) = v_{IPE}.$$

PROOF. See Appendix A.1. □

The JPE contract, w, considered in the statement of Theorem 1 induces a supermodular game among the agents under an appropriately defined partial order on actions, $\succeq$: $a_i \succeq a_i'$ if either $\mathbb{E}_{F(a_i)}[y_i] > \mathbb{E}_{F(a_i')}[y_i]$, or $\mathbb{E}_{F(a_i)}[y_i] = \mathbb{E}_{F(a_i')}[y_i]$ and $c(a_i) \leq c(a_i')$. Hence, by standard results in the literature on supermodular

[11] Carroll (2015) establishes the existence of an optimal IPE under principal-preferred action selection. Under the assumption that known actions are costly, there continues to exist an optimal contract even under principal least-preferred action selection. Moreover,

$$v_{IPE} = \max_{w\colon w \text{ is an IPE}} \inf_{A \in \mathcal{A}} \min(V(w, A)).$$

games (see, e.g., Vives (1990) and Milgrom and Roberts (1990)), there is a maximal and minimal rationalizable action profile and each profile is a Nash equilibrium. When the parameters $(\phi_i, \beta_i)_i$ are chosen so that the principal's payoff is increasing in the agents' action profile, it follows that, given any contract w and action set A, the principal's rationalizable payoffs are an interval[12]

$$I(w, A) = [\min(V(w, A)), \max(V(w, A))].$$

Theorem 1 establishes that, for any action set A, all payoffs in $I(w, A)$ are weakly larger than v_{IPE} and there are a continuum of payoffs strictly larger than v_{IPE}.[13] This dominance is also retained when taking limits; under any sequence of action sets in which the principal's smallest rationalizable (Nash) payoff is either equal to or approaches v_{IPE}, there is a sequence of rationalizable (Nash) payoffs bounded away from v_{IPE}.

Theorem 1 has a number of important implications for the selection of robustly optimal contracts under single-valued solution concepts. For instance, under principal-preferred Nash equilibrium selection (the assumption in the literature discussed in Section 1.1), JPE strictly outperforms IPE. In addition, because the maximal rationalizable (Nash) action profile in any supermodular game with positive spillovers Pareto dominates all other rationalizable (Nash) action profiles, under any selection of rationalizable action profiles satisfying weak Pareto efficiency for the agents, e.g., selection of a Pareto undominated Nash equilibrium, JPE strictly outperforms IPE. Put differently, there is only a “tie" between JPE and IPE if agents coordinate against their own interest. These observations are collected in the following Corollary.

COROLLARY 1. *The following results are immediate from Theorem 1:*

[12] To see that all payoffs in the interval are attainable, it suffices to randomize over the maximal and minimal action profiles.

[13] Formally, the set of worst-case rationalizable payoffs dominate $\{v_{IPE}\}$ in the interval order $\succeq_I$: for closed intervals $X \subseteq \mathbb{R}$ and $Y \subseteq \mathbb{R}$, $X \succeq_I Y$ if $x \in X$ and $y \in Y$ implies $x \geq y$.

1. *(Ambiguity Aversion over Nash Equilibria) Suppose that, given a contract w and action set A, the principal has ambiguity aversion over the Nash equilibrium the agents will play as captured by the λ-Maxmin Expected Utility representation of Ghirardato et al. (2004):*

$$V_\lambda(w,A) := \lambda \min_{\sigma\in\mathcal{E}(w,A)} \mathbb{E}_\sigma[\sum_{i=1}^n (y_i - w_i(y))] + (1-\lambda) \max_{\sigma\in\mathcal{E}(w,A)} \mathbb{E}_\sigma[\sum_{i=1}^n (y_i - w_i(y))],$$

   where $\lambda \in [0,1]$ and $\mathcal{E}(w,A)$ is the set of Nash equilibria in $\Gamma(w,A)$. Then, under the JPE contract in Theorem 1, denoted by w_{JPE},

$$\inf_{A\in\mathcal{A}} V_\lambda(w_{JPE},A) \quad \geq \quad \sup_{w:\, w \text{ is an IPE}} \quad \inf_{A\in\mathcal{A}} \quad V_\lambda(w,A),$$

   where the inequality is strict for any $\lambda < 1$. Notice that $\lambda = 0$ corresponds to partial Nash implementation.

2. *(Cooperative Solutions) Let $f : \mathcal{W} \times \mathcal{A} \to \Delta((\mathbb{R}_+ \times \Delta(Y))^n)$ be any selection of a distribution over rationalizable action profiles, i.e., $f(w,A) \in \mathcal{R}(w,A)$, that is weakly Pareto efficient for the agents, i.e., $f(w,A) \neq \sigma \in \mathcal{R}(w,A)$ if there exists a distribution $\sigma' \in \mathcal{R}(w,A)$ such that, for all agents i, $\mathbb{E}_\sigma[U_i(\cdot;w)] < \mathbb{E}_{\sigma'}[U_i(\cdot;w)]$. Given a contract w and action set A, let*

$$V_f(w,A) := \mathbb{E}_{f(w,A)}[\sum_{i=1}^n (y_i - w_i(y))].$$

   Then, under the JPE contract in Theorem 1, denoted by w_{JPE},

$$\inf_{A\in\mathcal{A}} V_f(w_{JPE},A) \quad > \quad \sup_{w:\, w \text{ is an IPE}} \quad \inf_{A\in\mathcal{A}} \quad V_f(w,A).$$

A simple example illustrates the key intuition behind the result.

EXAMPLE 1 (JPE versus IPE). There are two agents ($n = 2$) and output is binary ($Y = \{0,1\}$). There is a single, common known action ($A_1^0 = A_2^0 = \{a_0\}$). The known action results in success, $y = 1$, with probability $p(a_0) > 0$ and failure, $y = 0$, with probability $1 - p(a_0)$. Its effort cost is $c(a_0) \in (0, p(a_0))$. The principal is concerned about unknown action sets of the form $A = \{a_0, a^*\} \times \{a_0, a^*\}$, where a^* is a “shirking" action available to both agents. She knows that shirking entails zero effort cost, $c(a^*) = 0$, and is less productive than the known action, i.e., the probability of success is $p(a^*) < p(a_0)$. Moreover, she assumes that she can select her most-preferred Nash equilibrium in case of multiplicity.

Consider first the principal's payoff guarantee from an optimal IPE, which pays each agent a share of output $\alpha \in (c(a_0), 1)$. A naïve intuition is that principal's worst-case payoff is obtained when $p(a^*) = 0$; if agents take a shirking action with this success probability, then the principal obtains an expected payoff of zero. But, this logic ignores incentives, as pointed out by Carroll (2015). In particular, each agent has a strict incentive to shirk only if she obtains a higher expected utility from doing so. Hence, (a_0, a_0) is a Nash equilibrium whenever

$$p(a^*)\alpha \le p(a_0)\alpha - c(a_0) \iff p(a^*) \le p(a_0) - \frac{c(a_0)}{\alpha},$$

yielding the principal a payoff per agent of

$$p(a_0)(1-\alpha).$$

The principal's worst-case payoff is instead obtained as p^* approaches $p(a_0) - c(a_0)/\alpha$ from above. Along this sequence, (a^*, a^*) is the unique Nash equilibrium and the principal's payoff per agent becomes arbitrarily close to

$$(p(a_0) - \frac{c(a_0)}{\alpha})(1-\alpha).$$

Now, consider a contract of the form described in the statement of Theorem 1, parameterized by $\phi := \phi_1 = \phi_2 > 0$ and $\beta := \beta_1 = \beta_2 > 0$. Specifically, choose ϕ so that it is strictly smaller than the benchmark IPE share, α, and choose

$$\beta = \frac{\alpha - \phi}{p(a_0)}.$$

This contract is calibrated to the optimal IPE in the following sense: If an agent succeeds at her task, then her expected wage payment remains the same conditional on the other agent working. That is,

$$\phi + \beta p(a_0) = \alpha.$$

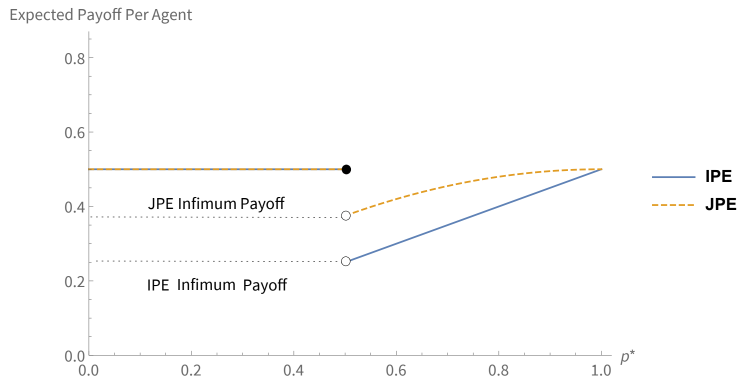

FIGURE 1. Principal's expected payoff per agent as a function of $p^* := p(a^*)$. Parameters: $p(a_0) = 1$, $c(a_0) = 1/4$, $\alpha = 1/2$, $\phi = 0$, and $\beta = 1/2$.

Hence, (a_0, a_0) is, again, a Nash equilibrium whenever

$$p(a^*) \leq p(a_0) - \frac{c(a_0)}{\alpha}.$$

Moreover, the principal's worst-case payoff is again obtained as $p(a^*)$ approaches $p(a_0) - c(a_0)/\alpha$ from above. (Along this sequence, (a^*, a^*) is the unique Nash equilibrium.) However, a simple calculation shows that the principal obtains a strictly higher payoff per agent in worst-case scenarios:

$$(p(a_0) - \frac{c(a_0)}{\alpha})(1 - (\phi + \beta p(a^*))) > (p(a_0) - \frac{c(a_0)}{\alpha})(1 - \alpha),$$

where the inequality follows from $\phi + \beta p(a^*) < \phi + \beta p(a_0) = \alpha$. See Figure 1 for an illustration. $\diamondsuit$

The intuition behind Example 1 is simple. By constructing a mean-preserving spread of an agent's wage with respect to the targeted action of the other, worst-case productivity is held constant. But, under joint performance evaluation, the principal pays agents less in expectation in worst-case scenarios. Each is punished for the shirking of the other.

While Example 1 identifies the key rent-extraction advantage of JPE over IPE, both the focus on partial Nash implementation and the class of games considered is with loss of generality. A previous version of this paper established that, under partial Nash implementation, the worst-case symmetric game for the principal is the limit of an n-sequence of dominance solvable games with n unknown actions. In each game in this sequence, agents "undercut" each other as dominated strategies are eliminated, taking progressively less costly and less productive actions. In asymmetric games, the productivity of a single agent can be driven even lower. When one agent takes a costless and less-productive action, others are willing to take even less-productive, costless actions. Both incentive problems arise because the share of individual output each non-shirking agent receives, on average, is reduced when others shirk. So, "all at once" reductions in expected output are less damaging than incremental or sequential reductions.

Theorem 1 asserts that there nevertheless exists a JPE contract that outperforms any IPE contract. To mitigate the free-riding problem and increase the entire set of rationalizable payoffs, the proof utilizes a more conservative calibration argument than illustrated in Example 1. Specifically, it identifies an improved JPE contract in which each agent's contract is calibrated to an IPE yielding him a larger share of individual output than optimal. Despite encouraging greater productivity, these larger IPE contracts are suboptimal on their own because they leave too much rent to each agent. But, under the calibrated JPE contract, the principal reduces expected wage payments in worst-case scenarios. This reduction is shown to be large enough that the calibrated JPE contract strictly outperforms both the larger IPE contract and the smaller, optimal IPE contract. Hence, the superiority of JPE over IPE is retained even when agents may be playing more complicated games than those considered in Example 1 and under alternative equilibrium selection criterion.

To conclude this section, observe that the dominating JPE contracts in Theorem 1 resemble the "bonus pool" contracts used by many corporations. In such an incentive scheme, the size of the pool depends on the overall performance of the company, i.e., $\sum_i^n y_i$, and each worker's share of the pool depends on their individual contribution, i.e., y_i. The team-based component of w_i,

$$\left(\frac{\beta_i}{n-1}\sum_{j\neq i}^{n} y_j\right) y_i,$$

corresponds to pay from the bonus pool, while the purely individual component,

$$\phi_i y_i,$$

controls worst-case shirking incentives. Notice that y_i is removed from agent i's bonus pool payment to avoid double-counting payments for individual contributions.

4. Optimality of Joint Performance Evaluation

The second main result is that, in a canonical setting, any optimal contract is of the form of the dominating contracts in Theorem 1.[14] Specifically, it is assumed that there are two agents possessing the same known action set, two output levels, and the principal is restricted to use symmetric incentive contracts. Moreover, the principal can select her preferred Nash equilibrium in case of multiplicity. Let $\mathcal{E}(w, A)$ be the set of Nash equilibria in the game $\Gamma(w, A)$. Then, from contract w, the principal obtains a payoff of

$$V(w) := \inf_{A \in \mathcal{A}} \max_{\sigma \in \mathcal{E}(w,A)} \mathbb{E}_\sigma[\sum_{i=1}^{n} (y_i - w_i(y))].$$

Before the result is stated, the contract space and notion of optimality are formally defined. If there are two agents and two output levels, $Y := \{0, 1\}$, a contract for agent i is a quadruple of non-negative wages,

$$w^i := (w^i_{11}, w^i_{10}, w^i_{01}, w^i_{00}) \in \mathbb{R}^4_+,$$

where the first index of each wage indicates agent i's own success ($y_i = 1$) or failure ($y_i = 0$) and the second indicates the success or failure of the other agent. If, in addition, contracts are assumed to be symmetric, then superscripts can be dropped and the set of all contracts is simply the set of all non-negative quadruples. The typology of performance evaluations thus simplifies considerably.

[14] The Bayesian analog of this setting is thoroughly analyzed by Fleckinger (2012) and Fleckinger et al. (2023), who show that symmetric IPE contracts are optimal for independent and identical agents.

DEFINITION 2 (Binary Performance Evaluations). A contract $w = (w_{11}, w_{10}, w_{01}, w_{00}) \in \mathbb{R}^4_+$ is

- an **independent performance evaluation (IPE)** if $(w_{11}, w_{01}) = (w_{10}, w_{00})$;
- a **relative performance evaluation (RPE)** if $(w_{11}, w_{01}) < (w_{10}, w_{00})$;
- and a **joint performance evaluation (JPE)** if $(w_{11}, w_{01}) > (w_{10}, w_{00})$,

where $>$ and $<$ indicate strict inequality in at least one component and weak in both.

Notice that a symmetric bonus pool contract of the form described in Theorem 1 sets $w_{11} = \phi + \beta$, $w_{10} = \phi$, and $w_{01} = w_{00} = 0$ for $\phi > 0$ and $\beta > 0$. A contract w^* is said to be **s-optimal** if it maximizes $V(\cdot)$ over the set of symmetric contracts.

The main result follows below.

THEOREM 2 (Optimality of JPE). *Suppose there are two agents, $n = 2$, with a common set of known actions, $A^0 := A^0_1 = A^0_2$, and the set of output levels is binary, $Y = \{0, 1\}$. Then, any s-optimal contract is a JPE contract with $w_{01} = w_{00} = 0$ and there exists an s-optimal contract.*

PROOF. See Appendix A.2. □

The result is a consequence of two important Lemmas. The first Lemma establishes that it is without loss of generality to consider contracts that do not reward failure.[15]

LEMMA 1 (Suboptimality of Positive Wages for Failure). *For any contract w with $w_{00} > 0$ or $w_{01} > 0$, there either exists an IPE, RPE, or JPE contract w' with $w'_{01} = w'_{00} = 0$ that yields the principal a higher payoff.*

[15] Though the result is familiar, the proof is surprisingly nontrivial. Specifically, while reducing payoffs by a constant rules out many contracts, there are two cases that require different arguments. First, when $w_{11} > 0$ and $w_{00} > 0$ (with $w_{01} = w_{00} = 0$), it must be argued that the probability of success under any equilibrium action decreases in w_{00}. Second, when $w_{10} > 0$ and $w_{01} > 0$ (with $w_{11} = w_{00} = 0$), symmetric and mixed equilibria that might be beneficial for the principal must be ruled out.

PROOF. See Appendix A.3. □

The second Lemma establishes that no RPE contract can outperform the best IPE contract.

LEMMA 2 (IPE Outperforms RPE). *No RPE contract with* $w_{01} = w_{00} = 0$ *can yield the principal a higher payoff than the optimal IPE contract.*

PROOF. See Appendix A.4. □

From Corollary 1 part 1 with $\lambda = 0$ and the proof of Theorem 1, there exists a symmetric JPE contract with $w_{01} = w_{00} = 0$ that strictly outperforms any IPE contract. Hence, if there exists an s-optimal contract, then there exists an s-optimal JPE contract with $w_{01} = w_{00} = 0$. Existence and uniqueness follow from simple, technical verifications.

Symmetric RPE contracts, e.g., salesperson-of-the-year awards, are commonly utilized in practice. So, it is worthwhile to describe the economic intuition behind their suboptimality under robustness considerations. Under RPE, agents are discouraged to take more productive actions when others are more productive. When one agent is productive, the other agent has less of an incentive to take a productive action because his chance of outperforming the other decreases. Given that one agent is willing to shirk, it is then possible to provide incentives for the other agent to shirk. In the resulting equilibrium, expected wage payments actually *increase*; weight is shifted from w_{11} to w_{10} and $w_{10} > w_{11}$, in contrast to the case of JPE. The corresponding increase in expected wage payments offsets the advantage of encouraging productivity by one of the two agents. The mechanics of the argument are illustrated in an elaboration of Example 1.

EXAMPLE 2 (RPE versus IPE). Suppose the environment and space of uncertainty are the same as in Example 1. Consider the performance guarantee of an RPE contract with $w_{11} > w_{10} > 0$. Observe that a^* is a strict best response to a^* if and only if

$$\underbrace{p(a^*)\left(p(a^*)w_{11} + (1-p(a^*))w_{10}\right)}_{\text{Payoff } a^* \text{ against } a^*} > \underbrace{p(a_0)\left(p(a^*)w_{11} + (1-p(a^*))w_{10}\right) - c(a_0)}_{\text{Payoff } a_0 \text{ against } a^*}$$

$$\iff p(a^*) > p(a_0) - \frac{c(a_0)}{p(a^*)w_{11} + (1-p(a^*))w_{10}}.$$

That is, if one agent shirks, the other has a strict incentive to shirk. If this inequality is satisfied, then a^* is also a strict best-response to a_0; the incentive to shirk is larger when the other works because $w_{11} < w_{10}$. So, a^* is a strictly dominant strategy and (a^*, a^*) is the unique Nash equilibrium.

Now, suppose the productivity of the shirking action approaches from above the value, $p(a^*)$, at which

$$p(a^*) = p(a_0) - \frac{c(a_0)}{p(a^*)w_{11} + (1-p(a^*))w_{10}}.$$

Then, (a^*, a^*) is the unique Nash equilibrium along the sequence and the principal's payoff per agent approaches

$$\left(p(a_0) - \frac{c(a_0)}{p(a^*)w_{11} + (1-p(a^*))w_{10}}\right)(1 - (p(a^*)w_{11} + (1-p(a^*))w_{10})).$$

This payoff can be no higher than what is obtained from an IPE contract with share $\alpha := p(a^*)w_{11} + (1-p(a^*))(1-w_{10})$, whose payoff is derived in Example 1. $\diamondsuit$

5. Final Remarks

This paper identifies non-statistical foundations for team-based incentive schemes commonly used in practice. Very generally, it is shown that linking the pay of independent agents is robustly optimal. Moreover, in a canonical environment, joint performance evaluation contracts, e.g., bonus pool incentive programs, are optimal. Such contracts approximate the incentive properties of benchmark independent performance evaluation contracts, while flexibly reducing expected wage payments when agents are less productive than the principal anticipates. The worst-case analysis draws attention to these scenarios, uncovering an economic intuition that had previously gone unnoticed.

Appendix A: Proofs

A.1 *Proof of Theorem 1*

From Carroll (2015), there exists an optimal IPE contract such that, for each i,

$$w_i(y_i, y_{-i}) = \alpha_i y_i,$$

where $\alpha_i = \sqrt{c(a_i^0)}/\sqrt{\mathbb{E}_{F(a_i^0)}[y]}$ for some $a_i^0 \in A_i^0$. The infimum payoff from agent i is attained in the limit of a sequence of action sets $(A_i(k))_k$ in which the agent's unique rationalizable action has expected output converging to

$$\bar{p}_i := \mathbb{E}_{F(a_i^0)}[y_i] - \frac{c(a_i^0)}{\alpha_i}, \tag{1}$$

yielding the principal a (worst-case) payoff of

$$\bar{p}_i\,(1-\alpha_i)\,.$$

When compensating n agents using only optimal IPE contracts, the principal's payoff is thus

$$v_{IPE} = \sum_{i=1}^{n} \bar{p}_i\,(1-\alpha_i)\,.$$

Fix a collection of optimal IPE shares $(\alpha_i)_i$, where $\alpha_i = \sqrt{c(a_i^0)}/\sqrt{\mathbb{E}_{F(a_i^0)}[y]}$ for some $a_i^0 \in A_i^0$. Consider a nonaffine JPE contract such that, for each agent i,

$$w_i(y_i, y_{-i}) = (\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} y_j)y_i, \tag{2}$$

where $\alpha_i > \phi_i > c(a_i^0)/\mathbb{E}_{F(a_i^0)}[y_i]$ and $\frac{c(a_i^0)}{(\mathbb{E}_{F(a_i^0)}[y_i])^2} > \beta_i > 0$ are chosen so that

$$\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \bar{p}_j = \alpha_i, \tag{3}$$

with $\bar{p}_i$ defined in (1). Notice that, because $\mathbb{E}_{F(a_j^0)}[y] > \bar{p}_j$ for all $j \neq i$, the constructed JPE is calibrated to a contract strictly larger than the optimal IPE for agent i. In addition, if all agents take an action with expected output equal to $\bar{p}_i$, then the principal's payoff under the JPE contract is exactly v_{IPE}, a consequence of the reduction in expected wage payments under JPE.

Now, equip any $A_i \supseteq A_i^0$ with the partial order $\succeq$: $a_i \succeq a_i'$ if either $\mathbb{E}_{F(a_i)}[y_i] > \mathbb{E}_{F(a_i')}[y_i]$, or $\mathbb{E}_{F(a_i)}[y_i] = \mathbb{E}_{F(a_i')}[y_i]$ and $c(a_i) \leq c(a_i')$. Then, $\Gamma(w, A)$ is a supermodular game under the corresponding product order on action profiles: $a' \succeq a$ implies $\mathbb{E}_{F(a_i')}[y_i] \geq \mathbb{E}_{F(a_i)}[y_i]$ for all i and hence

$$\begin{aligned} U_i(a_i', a_{-i}'; w) - U_i(a_i, a_{-i}'; w) &= \left(\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \mathbb{E}_{F(a_j')}[y_j]\right)\left(\mathbb{E}_{F_i(a_i')}[y_i] - \mathbb{E}_{F_i(a_i)}[y_i]\right) \\ &\geq \left(\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \mathbb{E}_{F(a_j)}[y_j]\right)\left(\mathbb{E}_{F_i(a_i')}[y_i] - \mathbb{E}_{F_i(a_i)}[y_i]\right) \\ &= U_i(a_i', a_{-i}; w) - U_i(a_i, a_{-i}; w). \end{aligned}$$

The following Lemma establishes that, in any game, every agent i plays an action with expected output weakly larger than $\bar{p}_i$ in any rationalizable strategy profile. Moreover, there exists a game with a rationalizable action profile in which each agent i produces expected output equal to $\bar{p}_i$.

LEMMA 3. *Suppose, for each i, w_i satisfies (2)-(3). Then, given any action set A satisfying $A_i \supseteq A_i^0$, any rationalizable action for agent i in $\Gamma(w, A)$ has expected output weakly larger than $\bar{p}_i$. However, there exists an action set A satisfying $A_i \supseteq A_i^0$ such that $\Gamma(w, A)$ has a rationalizable action profile in which each agent i produces expected output exactly equal to $\bar{p}_i$.*

PROOF. Given the presence of a_i^0, under any conjecture about other agents' actions, agent i is unwilling to play an action with expected output smaller than

$$p_i^1 := \mathbb{E}_{F(a_i^0)}[y] - \frac{c(a_i^0)}{\phi_i} > 0$$

because $\min(Y) = 0$. If agent i knows each agent $j \neq i$ is unwilling to play an action with expected output smaller than p_j^1, then he is unwilling to play an action with expected output smaller than

$$p_i^2 := \mathbb{E}_{F(a_i^0)}[y] - \frac{c(a_i^0)}{\phi_i + \dfrac{\beta_i}{n-1}\displaystyle\sum_{j\neq i}^{n} p_j^1} > p_i^1.$$

Iterating yields a strictly increasing and bounded sequence, $(p_1^k, ..., p_n^k)_k$. Hence, its limit, $(p_1^\infty, ..., p_2^\infty) \in [0, \max(Y)]^n$, exists by the monotone convergence theorem and must satisfy

$$p_i^\infty = \mathbb{E}_{F(a_i^0)}[y] - \frac{c(a_i^0)}{\phi_i + \dfrac{\beta_i}{n-1}\displaystyle\sum_{j\neq i}^{n} p_j^\infty} \quad \text{for all } i = 1, ..., n.$$

By (3), $p_i^\infty = \bar{p}_i$ for all i is a solution to the system of equations. It is the unique solution because the map $T : \mathbb{R}_+^n \to \mathbb{R}_+^n$ with i-th component

$$T_i(p) = \mathbb{E}_{F(a_i^0)}[y] - \frac{c(a_i^0)}{\phi_i + \dfrac{\beta_i}{n-1}\displaystyle\sum_{j\neq i}^{n} p_j}$$

is a contraction on $(\mathbb{R}_+^n, d)$, where $d(x, y) := \max_i |x_i - y_i|$ is the supremum (Chebyshev) distance. To prove this, observe that for any vectors $p, p' \in \mathbb{R}_+^n$,

$$\begin{aligned}
|T_i(p) - T_i(p')| &= \left|\frac{c(a_i^0)}{\phi_i + \dfrac{\beta_i}{n-1}\displaystyle\sum_{j\neq i}^{n} p_j} - \frac{c(a_i^0)}{\phi_i + \dfrac{\beta_i}{n-1}\displaystyle\sum_{j\neq i}^{n} p'_j}\right| \\
&= \left|\frac{c(a_i^0)\beta_i(\dfrac{1}{n-1}(\displaystyle\sum_{j\neq i} p_j - \sum_{j\neq i} p'_j))}{(\phi_i + \dfrac{\beta_i}{n-1}\displaystyle\sum_{j\neq i}^{n} p_j)(\phi_i + \dfrac{\beta_i}{n-1}\displaystyle\sum_{j\neq i}^{n} p'_j)}\right| \\
&\leq \left|\frac{c(a_i^0)\beta_i}{\phi_i^2}\right| d(p, p'),
\end{aligned}$$

with $|\frac{c(a_i^0)\beta_i}{\phi_i^2}| \leq |\beta_i \left(\frac{(\mathbb{E}_{F(a_i^0)}[y_i])^2}{c(a_i^0)}\right)| < 1$ by $\phi_i > c(a_i^0)/\mathbb{E}_{F(a_i^0)}[y_i]$ and $\beta_i < \frac{c(a_i^0)}{(\mathbb{E}_{F(a_i^0)}[y_i])^2}$. So

$$d(T(p), T(p')) = \max_i |T_i(p) - T_i(p')| \leq \kappa d(p, p'),$$

where $\kappa := \min_i |\frac{c(a_i^0)\beta_i}{\phi_i^2}| < 1$.

Now, consider the action set $A := \times_{i=1}^n A_i^0 \cup \{a_i^*\}$, where $c(a_i^*) = 0$ and $\mathbb{E}_{F(a_i^*)}[y_i] = \bar{p}_i$. In $\Gamma(w, A)$, $(a_1^*, \ldots, a_n^*)$ is rationalizable because a_i^* is a best-response to a_{-i}^* (it yields the same payoff as i's targeted known action, which is a best-response to a_{-i}^* in A_i^0). So, there exists an action set with a rationalizable action profile in which each agent i produces expected output $\bar{p}_i$. □

In addition, there exists some $\epsilon > 0$ such that in any game played by the agents there exists a rationalizable action profile in which there is some agent i with expected output weakly larger than $\bar{p}_i + \epsilon$.

LEMMA 4. *There exists an $\epsilon > 0$ such that, in any game $\Gamma(w, A)$ in which, for each i, w_i satisfies* (2)-(3) *and $A_i \supseteq A_i^0$, there is at least one rationalizable action profile in which some agent i plays an action with expected output weakly larger than $\bar{p}_i + \epsilon$.*

PROOF. For each i, define

$$\hat{p}_i := \mathbb{E}_{F(a_i^0)}[y] - \frac{(c(a_i^0)/2)}{\phi_i + \dfrac{\beta_i}{(n-1)} \displaystyle\sum_{j \neq i}^n \bar{p}_j} > \bar{p}_i,$$

a lower bound on the expected output of any rationalizable action with cost weakly greater than $c(a_i^0)/2$ by Lemma 3. Let $\delta \in (0, \frac{1}{3}\min_i(\mathbb{E}_{F(a_i^0)}[y_i] - \bar{p}_i))$ satisfy both

$$\frac{1}{3} \max_i \left(\frac{\phi_i + \dfrac{\beta_i}{n-1} \displaystyle\sum_{j \neq i}^n (\bar{p}_j + 3\delta)}{\phi_i + \dfrac{\beta_i}{n-1} \displaystyle\sum_{j \neq i}^n \bar{p}_j} \right) < \frac{1}{2} \tag{4}$$

and

$$\delta < \frac{1}{3}\min_i \left(\frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^n \bar{p}_j} - \frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^n \hat{p}_j} \right). \tag{5}$$

Define

$$\epsilon := \min_i \left(\frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^n \bar{p}_j} - \frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^n (\bar{p}_j + \delta)} \right) > 0. \tag{6}$$

Towards contradiction, suppose that there exists an action set with $A_i \supseteq A_i^0$ in which all rationalizable action profiles involve each agent producing expected output strictly smaller than $\bar{p}_i + \epsilon$. It suffices to consider action sets with maximal action profile equal to the targeted known action profile, $(a_1^0, ..., a_n^0)$. Let $(a_i^k)_{k=0}^\infty$ be the best-response path for agent i obtained from infinite iteration of maximal best-response functions. Let $a_i^\infty := \lim_{k\to\infty} a_i^k$. Then, $(a_1^\infty, ..., a_n^\infty)$ is a rationalizable action profile (see, e.g., Vives (1990) and Milgrom and Roberts (1990)). Hence, it must satisfy $\mathbb{E}_{F(a_i^\infty)}[y_i] < \bar{p}_i + \epsilon$ for all i, or

$$\mathbb{E}_{F(a_i^\infty)}[y_i] < \mathbb{E}_{F(a_i^0)}[y_i] - \frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^n \bar{p}_j} + \epsilon \tag{7}$$

for all i using the definition of $\bar{p}_i$. For any $k \geq 1$, a_i^k is in agent i's maximal best-response path only if

$$\mathbb{E}_{F(a_i^k)}[y_i] \geq \mathbb{E}_{F(a_i^{k-1})}[y_i] - \frac{c(a_i^{k-1}) - c(a_i^k)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^n \mathbb{E}_{F(a_j^{k-1})}[y_j]}.$$

So,

$$\mathbb{E}_{F(a_i^\infty)}[y_i] \geq \mathbb{E}_{F(a_i^0)}[y_i] - \sum_{k=1}^{\infty} \frac{c(a_i^{k-1}) - c(a_i^k)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \mathbb{E}_{F(a_j^{k-1})}[y_j]}.$$

Hence, from (7),

$$\mathbb{E}_{F(a_i^0)}[y_i] - \frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \bar{p}_j} + \epsilon > \mathbb{E}_{F(a_i^0)}[y_i] - \sum_{k=1}^{\infty} \frac{c(a_i^{k-1}) - c(a_i^k)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \mathbb{E}_{F(a_j^{k-1})}[y_j]},$$

which holds if and only if

$$\epsilon > \frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \bar{p}_j} - \sum_{k=1}^{\infty} \frac{c(a_i^{k-1}) - c(a_i^k)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \mathbb{E}_{F(a_j^{k-1})}[y_j]}.$$

Re-arranging and using the definition of $\epsilon > 0$ yields

$$\sum_{k=1}^{\infty} \frac{c(a_i^{k-1}) - c(a_i^k)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \mathbb{E}_{F(a_j^{k-1})}[y_j]} > \frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} (\bar{p}_j + \delta)} \tag{8}$$

for all i. Let k_i be the largest iteration at which $\sum_{j\neq i}^{n} \mathbb{E}_{F(a_j^{k_i-1})}[y_j] \geq \sum_{j\neq i}^{n}(\bar{p}_j + 3\delta)$. Observe that $k_i \geq 1$ because $\delta < \frac{1}{3}\min_i(\mathbb{E}_{F(a_i^0)}[y_i] - \bar{p}_i)$ and $\mathbb{E}_{F(a_i^{k-1})}[y_i] \geq \mathbb{E}_{F(a_i^k)}[y_i] \geq \bar{p}_i$ for all $k \geq 1$. Moreover, $k_i < \infty$. If not, then (8) would be violated because $c(a_i^{k-1}) \geq c(a_i^k) \geq 0$ for all $k \geq 1$. In addition, $c(a_i^{k_i}) \geq \frac{c(a_i^0)}{2}$. If not, then (8) would be violated because $\mathbb{E}_{F(a_i^{k-1})}[y_i] \geq \mathbb{E}_{F(a_i^k)}[y_i] \geq \bar{p}_i$ for all $k \geq 1$, and the value of x that solves

$$\frac{x}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n}(\bar{p}_j + 3\delta)} + \frac{c(a_i^0) - x}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n} \bar{p}_j} = \frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n}(\bar{p}_j + \delta)} \iff$$

$$x = \frac{c(a_i^0)}{3}\left(\frac{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n}(\bar{p}_j + 3\delta)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n}\bar{p}_j}\right)$$

is smaller than $\frac{1}{2}c(a_i^0)$ by (4). Choose $K := \min_i k_i$. Then, in iteration K, there must exist some agent i and action a_i^K in i's best-response path satisfying $\mathbb{E}_{F(a_i^K)}[y_i] < \bar{p}_i + 3\delta$ and $c(a_i^K) \geq 0$ when $\mathbb{E}_{F(a_i^{K-1})}[y_j] \geq \hat{p}_j$ for all j. That is, it must be that

$$\left(\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n}\hat{p}_j\right)(\bar{p}_i + 3\delta) > \left(\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n}\hat{p}_j\right)\mathbb{E}_{F(a_i^0)}[y] - c(a_i^0).$$

But, re-arranging and using the definition of $\bar{p}_i$ yields

$$\delta > \frac{1}{3}\left(\frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n}\bar{p}_j} - \frac{c(a_i^0)}{\phi_i + \frac{\beta_i}{n-1}\sum_{j\neq i}^{n}\hat{p}_j}\right),$$

which contradicts (5). □

Suppose, for each i, ϕ_i is sufficiently close to α_i, so that the principal's payoff is strictly increasing in the expected output of each agent. Then, from Lemma 3,

$$\inf_{A\in\mathcal{A}} \min(V(w,A)) = v_{IPE}$$

and, from Lemma 4,

$$\inf_{A\in\mathcal{A}} \max(V(w,A)) > v_{IPE}.$$

A.2 *Proof of Theorem 2*

Carroll (2015)'s analysis shows that there is a symmetric IPE contract that is optimal within the class of all IPE contracts when $A_1^0 = A_2^0$: $w_{10} = w_{11} = \alpha$, where $\alpha = \sqrt{c(a^0)}/\sqrt{\mathbb{E}_{F(a^0)}[y]} > 0$ for some $a^0 \in A^0$, and $w_{00} = w_{01} = 0$. From (2)-(3), and the rest of the proof of Theorem 1, there thus exists a symmetric JPE contract parameterized by $\phi \in (0, \alpha)$ and $\beta > 0$ with $w_{11} = \phi + \beta$, $w_{10} = \phi$, and $w_{01} = w_{00} = 0$ that strictly outperforms the optimal IPE contract.

Lemma 1, proved in Appendix A.3, establishes that it suffices to compare RPE contracts satisfying $w_{00} = w_{01} = 0$ to JPE contracts. Lemma 2, proved in Appendix A.4, establishes that there is no such RPE contract that outperforms the best symmetric IPE contract. Hence, from the preceding paragraph, if an s-optimal contract exists, then there exists one that exhibits JPE.

Existence of an s-optimal JPE contract with $w_{01} = w_{00} = 0$ follows because the minimization problem over action sets is jointly continuous in (w_{11}, w_{10}), success probabilities, and cost parameters, and the constraint set is compact when (tight) strict best-response inequalities are made weak (recall, under JPE, it suffices to inspect best-response dynamics leading to a maximal equilibrium in the order defined in Section A.1). By the Maximum Theorem, the principal's payoff is thus continuous in (w_{11}, w_{10}). Because (w_{11}, w_{10}) can be taken to lie in a compact set (both wages are bounded below by zero and $w_{11} \geq w_{10}$ cannot exceed 1 for the principal to make positive profits), existence of an s-optimal contract is then guaranteed by the Weierstrass Theorem.

The proof of Lemma 1 shows that any contract that is not a JPE contract and does not set $w_{11} > 0$, $w_{00} > 0$, and $w_{10} = w_{01} = 0$ is weakly improved upon by an IPE or RPE contract. In addition, it establishes that any contract setting $w_{11} > 0$ and $w_{00} > 0$ (with $w_{10} = w_{01} = 0$) is strictly outperformed by a nonaffine JPE contract with $w_{00} = 0$. The uniqueness result follows.

A.3 *Proof of Lemma 1*

If $w_{11} \geq w_{01}$ ($w_{10} \geq w_{00}$), setting $w'_{11} = w_{11} - w_{01}$ and $w'_{01} = 0$ ($w'_{10} = w_{10} - w_{00}$ and $w'_{00} = 0$) shifts each agent's payoff by a constant. Similarly, if $w_{11} \leq w_{01}$ ($w_{10} \leq w_{00}$), setting $w'_{01} = w_{01} - w_{11}$ and $w'_{11} = 0$ ($w'_{00} = w_{00} - w_{10}$ and $w'_{10} = 0$) shifts each agent's payoff by a constant. It follows that any Nash equilibrium under w is also a Nash equilibrium under w'. Since the principal's ex post payment decreases, these adjustments must (weakly) increase her payoff.

The argument in the previous paragraph immediately establishes that if $w_{11} \geq w_{01}$ and $w_{10} \geq w_{00}$, then there exists an improved contract w' in which $w'_{00} = w'_{01} = 0$. There are three other cases to consider: (i) $w_{01} \geq w_{11}$ and $w_{00} \geq w_{10}$ (in which case it suffices to set $w_{11} = w_{10} = 0$); (ii) $w_{11} \geq w_{01}$ and $w_{00} > w_{10}$ (in which case it suffices to set $w_{01} = w_{10} = 0$); and (iii) $w_{01} > w_{11}$ and $w_{10} \geq w_{00}$ (in which case it suffices to set $w_{11} = w_{00} = 0$).

If $w_{01} \geq 0$ and $w_{00} \geq 0$, then w cannot yield the principal a positive payoff (and, hence, does not outperform the best IPE). To wit, let $A_i := A^0 \cup \{a_\emptyset\}$ where $p(a_\emptyset) = 0 = c(a_\emptyset)$. Then, $a_\emptyset$ is a strictly dominant strategy and so $(a_\emptyset, a_\emptyset)$ is the unique Nash equilibrium. In this equilibrium, the principal obtains a payoff $-2w_{00} \leq 0$.

If $w_{11} \geq w_{01} = 0$ and $w_{00} > w_{10} = 0$, then it must be that $w_{11} > 0$ or the principal could not attain a positive payoff by the argument in the preceding paragraph. Under such a contract, agent i's payoffs satisfy increasing differences in (a_i, a_j) when A_i is equipped with partial order $\succeq$: $a_i \succeq a'_i$ if either $\mathbb{E}_{F(a_i)}[y_i] > \mathbb{E}_{F(a'_i)}[y_i]$, or $\mathbb{E}_{F(a_i)}[y_i] = \mathbb{E}_{F(a'_i)}[y_i]$ and $c(a_i) \leq c(a'_i)$. Hence, any game this contract induces is supermodular. Moreover, fixing a_j, (a_i, w_{00}) satisfies decreasing differences and (a_i, w_{11}) satisfies increasing differences. Theorem 6 of Milgrom and Roberts (1990) then implies that the maximal and minimal equilibria of any game $\Gamma(w, A)$, $A_i \supseteq A^0$, are decreasing in w_{00} and increasing in w_{11}. Moreover, each agent's expected utility is increasing in expected output if the other has expected output larger than $\frac{w_{00}}{w_{11}+w_{00}}$. Notice that the principal's profit is increasing in expected output if and only if expected output by each agent is larger than $\tilde{p} = \frac{1/2+w_{00}}{w_{11}+w_{00}}$. So, if the principal's worst-case payoff is obtained in a region in which both agents succeed with probability strictly smaller than $\tilde{p}$, then reducing w_{00} by a small amount strictly increases the principal's payoff. On the other hand, if the principal's worst-case payoff is obtained in a region in which both agents succeed with probability strictly larger than $\tilde{p}$, then reducing w_{11} by a small amount constitutes a strict improvement. The knife-edge cases in which either both succeed with probability $\tilde{p}$ or one succeeds with probability less than $\tilde{p}$ and the other larger cannot occur in worst-case scenarios. In the former case, one can always add to each agent's action set a zero-cost action that succeeds with probability one and make the principal strictly worse off. In the latter case, one can construct a symmetric action set that makes the principal strictly worse off.

If $w_{01} > w_{11} = 0$ and $w_{10} \geq w_{00} = 0$, agent i's payoff satisfies decreasing differences. It is shown that the principal's payoff under such a contract cannot exceed the principal's payoff under the best IPE contract, v_{IPE}. Let $a_\emptyset$ be the action satisfying $c(a_\emptyset) = p(a_\emptyset) = 0$. Let a^*_ϵ be an action for which $c(a^*_\epsilon) = 0$ and for which $p(a^*_\epsilon)$ is a fixed point of

$$T_\epsilon(p) := \begin{cases} \max_{a \in A^0 \cup \{a_\emptyset\}} \left[p(a) - \frac{c(a)}{w_{10} - p(w_{10}+w_{01})}\right] + \epsilon & \text{if } w_{10} - p(w_{10}+w_{01}) > 0 \\ 0 & \text{otherwise} \end{cases},$$

where $\epsilon > 0$ is small. To see that T_ϵ has a fixed point, notice that, for any $p \in [0,1]$, $T_\epsilon(p)$ is larger than zero (because $a_\emptyset \in A^0 \cup \{a_\emptyset\}$) and less than one if ϵ is small enough (because A^0 does not contain a zero-cost action that results in success with probability one by the assumption of costly known productive actions). Hence, T_ϵ is a continuous function mapping $[0,1]$ into $[0,1]$.

By construction, $(a^*_\epsilon, a^*_\epsilon)$ is a Nash equilibrium of $\Gamma(w, A_\epsilon)$, where A_ϵ is an action set satisfying $A_i = A^0 \cup \{a^*_\epsilon, a_\emptyset\}$. Now, consider a sequence of strictly positive values $\epsilon_1, \epsilon_2, \ldots$ that converges to zero and for which there is a convergent sequence of fixed points $p(a^*_{\epsilon_1}), p(a^*_{\epsilon_2}), \ldots$ of the mappings $T_{\epsilon_1}, T_{\epsilon_2}, \ldots$. (Because $[0,1]$ is a compact set, such a convergent sequence must exist.) Moreover, if the limit p^* satisfies $w_{10} - p^*(w_{10}+w_{01}) > 0$, then it must equal

$$p^* := \max_{a \in A^0 \cup \{a_\emptyset\}} \left[p(a) - \frac{c(a)}{w_{10} - p^*(w_{10}+w_{01})}\right].$$

It is shown that the principal's worst-case payoff in the limit can be no larger than what she obtains from the optimal IPE contract. If p^* equals zero, then the principal attains less than zero profits and so lower profits than under the optimal IPE contract. Otherwise, let $\hat{a}_0$ denote a maximizer of $p(a) - \frac{c(a)}{w_{10} - p^*(w_{10}+w_{01})}$ over

$A^0 \cup \{a_\emptyset\}$, let $\hat{\alpha} := (1-p^*)w_{10}$, and notice that the principal attains a payoff of

$$\begin{aligned}
&2\left[(p^*)^2 + p^*(1-p^*)(1-w_{01}-w_{10})\right] \\
&= 2\left[p(\hat{a}_0) - \frac{c(\hat{a}_0)}{(1-p^*)(w_{10}+w_{01})}\right]\left[1-(1-p^*)(w_{10}+w_{01})\right] \\
&\leq 2\left[p(\hat{a}_0) - \frac{c(\hat{a}_0)}{(1-p^*)w_{10}}\right]\left[1-(1-p^*)w_{10}\right] \\
&= 2\left[p(\hat{a}_0) - \frac{c(\hat{a}_0)}{\hat{\alpha}}\right]\left[1-\hat{\alpha}\right].
\end{aligned}$$

But,

$$\begin{aligned}
2\left[p(\hat{a}_0) - \frac{c(\hat{a}_0)}{\hat{\alpha}}\right](1-\hat{\alpha}) &\leq 2 \max_{\alpha\in[0,1], a_0 \in A^0 \cup \{a_\emptyset\}} \left[(1-\alpha)(p(a_0) - \frac{c(a_0)}{\alpha})\right] \\
&= 2 \max_{\alpha\in[0,1], a_0 \in A^0} \left[(1-\alpha)(p(a_0) - \frac{c(a_0)}{\alpha})\right] \\
&= v_{IPE},
\end{aligned}$$

where the inequality follows because $p(\hat{a}_0) - \frac{c(\hat{a}_0)}{\hat{\alpha}} \geq 0$ for all $\hat{\alpha} \geq 0$, and the equality follows because setting $\alpha = 1$ yields the principal a payoff of zero given any action in A^0, the payoff attained from choosing $a_\emptyset$ and any $\alpha \in [0,1]$.

The previous argument establishes that if there exists a K such that, for all $k \geq K$, $(a^*_{\epsilon_k}, a^*_{\epsilon_k})$ is the unique Nash equilibrium of $\Gamma(w, A_{\epsilon_k})$, then the principal's worst-case payoff is no higher than v_{IPE}. But, other pure and mixed strategy equilibria may exist that benefit the principal, even as k grows large. First, consider the case in which the limit of $(a^*_{\epsilon_k})$ is $a_\emptyset$. If multiplicity arises, then there exists an action $a_0 \in A^0$ that results in success with strictly positive probability and is a weak best response to any action that succeeds with zero probability; if not, then there would exist a K such that for all $k \geq K$, $(a^*_{\epsilon_k}, a^*_{\epsilon_k})$ is the maximal Nash equilibrium of $\Gamma(w, A_{\epsilon_k})$ and hence the unique Nash equilibrium. If $p(a_0) \leq \frac{w_{10}}{w_{10}+w_{01}}$, then the principal's payoff in any equilibrium in which such an action is played with positive probability is less than zero. This follows from

$$p(a_0)(1-w_{10}-w_{01}) \leq \frac{w_{10}}{w_{10}+w_{01}} - w_{10} < 0.$$

If, on the other hand, $p(a_0) > \frac{w_{10}}{w_{10}+w_{01}}$, then add to each A_{ϵ_k} the action a_0' for which $c(a_0') = 0$ and $p(a_0') = p(a_0) - \frac{c(a_0)}{w_{10}}$ if $p(a_0) - \frac{c(a_0)}{w_{10}} > \frac{w_{10}}{w_{10}+w_{01}}$ and $p(a_0') = \frac{w_{10}}{w_{10}+w_{01}} + \epsilon_k$ otherwise. In the first case, the principal attains a payoff of

$$\left[p(a_0) - \frac{c(a_0)}{w_{10}}\right](1 - w_{10} - w_{01}) \leq 2 \max_{\alpha \in [0,1], a_0 \in A^0} \left[(1-\alpha)(p(a_0) - \frac{c(a_0)}{\alpha})\right] = v_{IPE}.$$

In the second case, there exists a K such that for all $k \geq K$, the principal's payoff in the equilibrium $(a_0', a^*_{\epsilon_k})$ is less than zero because the inequality in the previous displayed equation is strict. Finally, no mixed equilibria can exist in any of the cases considered because $a_\emptyset$ is a strict best response to any action larger than $\frac{w_{10}}{w_{10}+w_{01}}$ (the marginal benefit of succeeding with higher probability is less than zero).

Second, consider the case in which the limit of $(a^*_{\epsilon_k})$ is $p^* > 0$. Any other pure or mixed Nash equilibrium of $\Gamma(w, A_{\epsilon_k})$ must involve one agent succeeding with probability $\hat{p} \geq \frac{w_{10}}{w_{10}+w_{01}} > p^*$. If not, then $p(a^*_{\epsilon_k})$ would be a best-response to the distribution $\hat{p}$ and, if $p(a^*_{\epsilon_k})$ is played, then any distribution $\hat{p}$ could not be a best-response.[16] However, any equilibrium in which one agent generates a distribution $\hat{p}$ must have the other play either $a_\emptyset$ (if $\hat{p} > \frac{w_{10}}{w_{10}+w_{01}}$), $a^*_{\epsilon_k}$ (only if $\hat{p} = \frac{w_{10}}{w_{10}+w_{01}}$), or a mixture between the two (again, only if $\hat{p} = \frac{w_{10}}{w_{10}+w_{01}}$); known productive actions are costly and the marginal benefit of succeeding with higher probability is less than zero (strictly so if $\hat{p} > \frac{w_{10}}{w_{10}+w_{01}}$). It suffices to consider the case in which $\hat{p} > \frac{w_{10}}{w_{10}+w_{01}}$. In the other two cases, introducing an action that has the same productivity as the most productive action in the support of the player's strategy that succeeds with probability $\hat{p}$, but an (arbitrarily) smaller cost, reduces the problem to this case, or alternatively, results in the equilibrium $(a^*_{\epsilon_k}, a^*_{\epsilon_k})$. So, consider any action, $a_0 \in A^0$, satisfying $p(a_0) \geq \frac{w_{10}}{w_{10}+w_{01}}$ in the support of the strategy succeeding with probability $\hat{p} > \frac{w_{10}}{w_{10}+w_{01}}$. Mirroring the argument in the previous case, add to each A_{ϵ_k} the action a_0' for which $c(a_0') = 0$ and $p(a_0') = p(a_0) - \frac{c(a_0)}{w_{10}} + \epsilon_k$ if

[16]The first statement follows because $p(a^*_\epsilon)$ has zero cost, profits would still be increasing in the probability with which the agent succeeds, and there are strictly decreasing differences. The second follows because $p(a^*_{\epsilon_k})$ is a strict best-response to $p(a^*_{\epsilon_k})$ by construction.

$p(a_0) - \frac{c(a_0)}{w_{10}} > \frac{w_{10}}{w_{10}+w_{01}}$ and $p(a_0') = \frac{w_{10}}{w_{10}+w_{01}} + \epsilon_k$ otherwise. These adjustments ensure that a_0' is the unique best response to $a_\emptyset$ for every k and so, mirroring the steps in the proof of the previous case, the principal attains a payoff no larger than v_{IPE}.

A.4 *Proof of Lemma 2*

The proof will utlize the following result from the theory of supermodular games. Let $a_{\max}$ and $a_{\min}$ denote the maximal and minimal elements of A, and $\overline{BR} : A \to A$ and $\underline{BR} : A \to A$ denote the maximal and minimal best-response functions for the agents. Define the mapping

$$\widetilde{BR} : A \times A \to A \times A$$
$$(a_i, a_j) \mapsto (\overline{BR}(a_j), \underline{BR}(a_i)).$$

Then, the following Lemma holds.

LEMMA 5 (Vives (1990), Milgrom and Roberts (1990)). *Suppose $(\bar{a}, \underline{a})$ is the limit found by iterating $\widetilde{BR}$ starting from the action profile $(a_{\max}, a_{\min})$. If $\Gamma(w, A)$ is submodular, then both $(\bar{a}, \underline{a})$ and $(\underline{a}, \bar{a})$ are Nash equilibria and any other Nash equilibrium action must be smaller than $\bar{a}$ and larger than $\underline{a}$.*

Now, let $a_\emptyset$ be the action satisfying $c(a_\emptyset) = p(a_\emptyset) = 0$. Let a_ϵ^* be an action for which $c(a_\epsilon^*) = 0$ and for which $p(a_\epsilon^*)$ is a fixed point of

$$T_\epsilon(p) := \max_{a_0 \in A^0 \cup \{a_\emptyset\}} \left[p(a_0) - \frac{c(a_0)}{pw_{11} + (1-p)w_{10}} \right] + \epsilon,$$

where $\epsilon > 0$ is small.[17] To see that T_ϵ has a fixed point, notice that, for any $p \in [0,1]$, $T_\epsilon(p)$ is larger than zero (because $a_\emptyset \in A^0 \cup \{a_\emptyset\}$) and less than one if ϵ is small enough (because A^0 does not contain a zero-cost action that results in success with probability one). Hence, T_ϵ is a continuous function mapping $[0,1]$ into $[0,1]$.

[17] Interpret $-\frac{c(a_0)}{pw_{11}+(1-p)w_{10}}$ as zero if the denominator is zero and $c(a_0) = 0$ and $-\infty$ if the denominator is zero and $c(a_0) > 0$.

Now, define an action space A_ϵ that satisfies $A_i = A^0 \cup \{a^*_\epsilon, a_\emptyset\}$. If A^0 contains an action producing $y_i = 1$ with probability one, consider the least costly among all of them, $\bar{a}_0$, and add to A_ϵ the action $\bar{a}_\epsilon$, where $c(\bar{a}_\epsilon) = c(\bar{a}_0) - \gamma(\epsilon)$ and $p(\bar{a}_n) = 1 - \frac{\gamma(\epsilon)}{2}$ for $\gamma(\epsilon) := \frac{\epsilon(p(a^*_\epsilon)w_{11} + (1-p(a^*_\epsilon))w_{10}}{2}$. Then, $\bar{a}_\epsilon$ strictly dominates $\bar{a}_0$ (and so any other action producing $y_i = 1$ with probability one is as well) and a^*_ϵ is a strictly better reply to a^*_ϵ than $\bar{a}_\epsilon$.

It is shown that $(a^*_\epsilon, a^*_\epsilon)$ is the unique Nash equilibrium of $\Gamma(w, A_\epsilon)$. Notice, by construction, $(a^*_\epsilon, a^*_\epsilon)$ is a strict Nash equilibrium. Now, remove all actions producing $y_i = 1$ with probability one since they are strictly dominated by $\bar{a}_\epsilon$. Upon removing these actions, a^*_ϵ strictly dominates any action smaller than it in the order $\succeq$. So, remove any actions in $\Gamma(w, A_\epsilon)$ below a^*_ϵ and denote the resulting action space by $\hat{A}$. Now, consider the profile $(\bar{a}, a^*_\epsilon)$, where $\bar{a}$ is the largest element of $\hat{A}$. Since a^*_ϵ is the unique best response to a^*_ϵ (because $(a^*_\epsilon, a^*_\epsilon)$ is a strict Nash equilibrium), the maximal best-response to a^*_ϵ is a^*_ϵ. This also implies that a^*_ϵ is the minimal best-response to $\bar{a}$; if not, there exists some $\hat{a}_0 \in \hat{A}$ such that $\hat{a}_0 \succ a^*_\epsilon$ and $U_i(\hat{a}_0, a_0; w) - U_i(a^*_\epsilon, a_0; w) \geq U_i(\hat{a}_0, \bar{a}; w) - U_i(a^*_\epsilon, \bar{a}; w) > 0$ for any $a_0 \in \hat{A}$, where the first inequality follows from the property of decreasing differences and the second from a_0 being the smallest best-response to $\bar{a}$. Hence, $\hat{a}_0$ strictly dominates a^*_ϵ, contradicting the previous observation that a^*_ϵ is a best response to a^*_ϵ. As $(a^*_\epsilon, a^*_\epsilon)$ is a fixed point of $\widetilde{BR}$, $(a^*_\epsilon, a^*_\epsilon)$ is the limit found by iterating $\widetilde{BR}$ from $(\bar{a}, a^*_\epsilon)$ or $(a^*_\epsilon, \bar{a})$ in $\Gamma(w, \hat{A})$. By Lemma 5, it follows that $(a^*_\epsilon, a^*_\epsilon)$ is the unique Nash equilibrium of $\Gamma(w, \hat{A})$ and hence of $\Gamma(w, A_\epsilon)$.

Now, consider a sequence of strictly positive values ϵ_1, ϵ_2,... that converges to zero and for which there is a convergent sequence of fixed points $p(a^*_{\epsilon_1})$, $p(a^*_{\epsilon_2})$,... of the mappings T_{ϵ_1}, T_{ϵ_2},... . Since $[0,1]$ is a compact set, such a convergent sequence must exist. Moreover, its limit is the distribution

$$p(a^*) = \max_{a_0 \in A^0 \cup \{a_\emptyset\}} \left[p(a_0) - \frac{c(a_0)}{p(a^*)w_{11} + (1 - p(a^*))w_{10}} \right].$$

Let $\hat{a}_0 \in A^0 \cup \{a_\emptyset\}$ denote the maximizer on the right-hand side and define $\hat{\alpha} := p(a^*)w_{11} + (1 - p(a^*))w_{10}$. The principal's payoff in the unique equilibrium

$(a^*_{\epsilon_k}, a^*_{\epsilon_k})$ of $\Gamma(w, A_{\epsilon_k})$ as k grows large becomes arbitrarily close to

$$2\left[p(a^*)\right]\left[p(a^*)(1-w_{11})+(1-p(a^*))(1-w_{10})\right]=$$

$$2\left[p(\hat{a}_0)-\frac{c(\hat{a}_0)}{\hat{\alpha}}\right](1-\hat{\alpha})\leq 2\max_{\alpha\in[0,1],a_0\in A^0\cup\{a_\emptyset\}}\left[(1-\alpha)(p(a_0)-\frac{c(a_0)}{\alpha})\right],$$

where the inequality follows because $p(\hat{a}_0)-\frac{c(\hat{a}_0)}{\hat{\alpha}}\geq 0$ for all $\hat{\alpha}\geq 0$ and so it suffices to consider values of α between zero and one to maximize $(1-\alpha)(p(a_0)-\frac{c(a_0)}{\alpha})$ for any $a_0\in A^0\cup\{a_\emptyset\}$. But,

$$\begin{aligned}&2\max_{\alpha\in[0,1],a_0\in A^0\cup\{a_\emptyset\}}\left[(1-\alpha)(p(a_0)-\frac{c(a_0)}{\alpha})\right]\\=&2\max_{\alpha\in[0,1],a_0\in A^0}\left[(1-\alpha)(p(a_0)-\frac{c(a_0)}{\alpha})\right]\\=&v_{IPE},\end{aligned}$$

where v_{IPE} is the principal's payoff under the best IPE, because setting $\alpha=1$ yields the principal a payoff of zero given any action in A^0, the same payoff attained from choosing $a_\emptyset$ and any $\alpha\in[0,1]$.

References

Alchian, Armen A. and Harold Demsetz (1972), "Production, information costs, and economic organization." *American Economic Review*, 62(5), 777–795. [5]

Antic, Nemanja (2015), "Contracting with unknown technologies." *Unpublished manuscript, Princeton University*. [5]

Bell, Brian, Simone Pedemonte, and John Van Reenen (2021), "Ceo pay and the rise of relative performance contracts: A question of governance?" *Journal of the European Economic Association*, 19(5), 2513–2542. [2, 4]

Benson, Alan M and Sima Sajjadiani (2018), "Are bonus pools driven by their incentive effects? evidence from fluctuations in gainsharing incentives." *ILR Review*, 71(3), 567–599. [2, 4]

Brandenburger, Adam and Eddie Dekel (1987), "Rationalizability and correlated equilibria." *Econometrica*, 1391–1402. [8]

Carroll, Gabriel (2015), "Robustness and linear contracts." *American Economic Review*, 105(2), 536–563. [3, 5, 9, 12, 19, 25]

Chassang, Sylvain (2013), "Calibrated incentive contracts." *Econometrica*, 81(5), 1935–1971. [5]

Che, Yeon-Koo and Seung-Weon Yoo (2001), "Optimal incentives for teams." *American Economic Review*, 91(3), 525–541. [5, 7]

Dai, Tianjiao and Juuso Toikka (2022), "Robust incentives for teams." *Econometrica*, 90(4), 1583–1613. [5, 6]

Fleckinger, Pierre (2012), "Correlation and relative performance evaluation." *Journal of Economic Theory*, 147(1), 93 – 117. [5, 15]

Fleckinger, Pierre, David Martimort, and Nicolas Roux (2023), "Should they compete or should they cooperate? the view of agency theory." *Working paper*. [15]

Frankel, Alexander (2014), "Aligned delegation." *American Economic Review*, 104(1), 66–83. [5]

Garrett, Daniel F. (2014), "Robustness of simple menus of contracts in cost-based procurement." *Games and Economic Behavior*, 87, 631 – 641. [5]

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci (2004), "Differentiating ambiguity and ambiguity attitude." *Journal of Economic Theory*, 118(2), 133–173. [11]

Hackman, J. Richard (2002), *Leading Teams: Setting the Stage for Great Performances*. Harvard Business Press. [3, 6]

Holmström, Bengt (1979), "Moral hazard and observability." *Bell Journal of Economics*, 10(1), 74–91. [2]

Holmström, Bengt (1982), "Moral hazard in teams." *Bell Journal of Economics*, 13(2), 324–340. [2]

Hurwicz, Leonid and Leonard Shapiro (1978), "Incentive structures maximizing residual gain under incomplete information." *Bell Journal of Economics*, 180–191. [5]

Itoh, Hideshi (1991), "Incentives to help in multi-agent situations." *Econometrica*, 59(3), 611–636. [5]

Kambhampati, Ashwin (2023), "Randomization is optimal in the robust principal-agent problem." *Journal of Economic Theory*, 207, 105585. [7]

Laffont, Jean-Jacques and David Martimort (2009), *The Theory of Incentives*. Princeton university press. [2]

Lazear, Edward P. (1989), "Pay equality and industrial politics." *Journal of Political Economy*, 97(3), 561–580. [5]

Marku, Keler, Sergio Ocampo Diaz, and Jean-Baptiste Tondji (2023), "Robust contracts in common agency." *Rand Journal of Economics (forthcoming)*. [5]

Milgrom, Paul and John Roberts (1990), "Rationalizability, learning, and equilibrium in games with strategic complementarities." *Econometrica*, 1255–1277. [10, 23, 27, 31]

Rees, Daniel I., Jeffrey S. Zax, and Joshua Herries (2003), "Interdependence in worker productivity." *Journal of Applied Econometrics*, 18(5), 585–604. [2]

Rosenthal, Maxwell (2020), "Robust incentives for risk." *Unpublished manuscript, Georgia Institute of Technology*. [5]

Shavell, Steven (1979), "On moral hazard and insurance." *Quarterly Journal of Economics*, 93(4), 541–562. [2]

Vives, Xavier (1990), "Nash equilibrium with strategic complementarities." *Journal of Mathematical Economics*, 19(3), 305 – 321. [10, 23, 31]