Selecting for Innovation

Gustavo Manso*

February 8, 2026

PRELIMINARY AND INCOMPLETE - NOT FOR CIRCULATION

ABSTRACT

Previous research shows that pay-for-performance incentive schemes help firms screen productive workers. Yet attracting creative talent remains a central challenge for organizations. This paper develops a theoretical model showing that contracts that attract exploratory workers exhibit rewards for long-term success and tolerance (or even rewards) for early failure. Commitment to a long-term contract, job security, and timely performance feedback are also important to selecting for innovation. These results underscore the need to distinguish between selection and motivation when assessing incentives for innovation and help explain why organizations separate innovation from production activities.

Keywords: Innovation, Incentives, Contract Theory, Selection, Exploration, Exploitation.
JEL Codes: D86, J33, M52, O31.

*Haas School of Business, University of California, Berkeley. Email: manso@berkeley.edu.

I Introduction

Standard theory of incentives, backed by empirical literature, suggests that offering high–powered pay–for–performance contracts attracts more productive workers. In models of worker compensation and sorting, agents self–select into firms based on the sensitivity of pay to performance: high–ability workers are drawn to steeper incentive schemes, while low–ability workers opt for flatter ones (Salop and Salop, 1976; Harris and Raviv, 1978; Lazear, 1986). Consistent with this view, Lazear's (2000) study of Safelite Glass documents that switching from hourly wages to piece rates increased productivity by 44%, with roughly two–thirds of the gain attributable to worker selection rather than changes in effort.

Nonetheless, in practice, CEOs and other senior executives consistently rank "attracting and retaining creative talent" as a top-tier strategic bottleneck. Recent data indicates a sharp rise in "talent anxiety" regarding the scarcity of innovative thinkers; for instance, 75% of CEOs now identify the inability to hire and retain workers with high-level creative and adaptive skills as a primary threat to their three-year growth strategy (Oliver Wyman, 2025). This concern is echoed across the C-suite, with leaders reporting that while standard labor productivity has stabilized, the availability of "human-centric" skills—specifically original thinking and strategic reinvention—remains critically low (IBM, 2025; KPMG, 2024). Furthermore, research suggests that traditional corporate structures and rigid performance metrics are actively driving away the most creative workers, who increasingly prioritize autonomy and psychological safety over high-powered incentives (Deloitte, 2025).These patterns suggest a fundamental misalignment in current labor economics: widely used pay-for-performance schemes, though effective at selecting for high-effort "exploiters," may actually repel the "explorers" firms need for long-term viability. This raises a critical design challenge: how should firms structure contracts to screen for creativity rather than mere industriousness?

This paper addresses this question by developing a two–period principal–agent model with two types of workers—"explorers," who can invest in a novel, uncertain approach, and "exploiters," who can only apply a safe, well–understood method—and asking how a principal should design contracts when both motivation and selection matter. The analysis yields two contrasting prescriptions. On one hand, the optimal contract that attracts exploiters has the familiar structure of a high–powered piece–rate plan: it ties pay closely to contemporaneous output, penalizes early failures through low current wages and the threat of termination, and places little weight on long–term commitment or feedback. On the other hand, the optimal contract that attracts explorers is almost its mirror image: it backloads rewards on long–run success, softens or even reverses penalties for early failure, and bundles these payoffs with job security, long–term commitment, and timely feedback on performance.

In this respect, the selection problem turns out not to undermine but to reinforce the logic of innovation–motivating design. Whereas Manso (2011) shows that tolerance for early failure, backloaded rewards, job security, and timely feedback are necessary to induce a given agent to explore, the present analysis demonstrates that the same contractual features also perform a powerful screening function: they make the "explorer" contract attractive precisely to workers with a comparative advantage in experimentation and unattractive to exploiters. For executives who worry about how to enhance innovation, the combination of this paper and Manso (2011) therefore offers a clear recipe: the very instruments that foster exploration not only motivate innovative effort but also help attract the workers most inclined to innovate.

A large empirical literature has examined how compensation and governance practices affect innovation, typically interpreting "exploration–type" arrangements as mechanisms that relax short–term performance pressure and thereby encourage experimentation.[1] Most of this work is framed in

[1]See Section VIII for a review.

terms of the motivational channel and uses tools such as differences-in-differences and instrumental variables to identify causal effects on innovative output, but it is difficult to exclude the possibility that changes in incentives also reshape who chooses to work where. By showing that exploration–friendly contracts are powerful screening devices as well as motivators, the present model points to a complementary empirical agenda: in the spirit of Lazear (2000), one can ask whether individuals with more exploratory histories are disproportionately likely to enter, remain in, or exit firms, industries, or regions that adopt (or roll back) long–horizon, failure–tolerant, feedback–rich incentive schemes.

To model selection for innovation, I build on the class of Bayesian decision models known as bandit problems. In bandit problems, the agent is uncertain about the true distribution of the available actions' payoffs, and innovation is the discovery, through experimentation and learning, of actions that are superior to previously known actions. As in Manso (2011), the central trade-off is between the exploration of new untested actions and the exploitation of well–known actions. Exploration of new untested actions reveals information about potentially superior approaches, but is also likely to waste time with inferior ones. Exploitation of well–known actions ensures reasonable payoffs, but may prevent the discovery of superior actions.

This environment is embedded in a principal–agent model with hidden types. The principal hires an agent who may be either an "explorer" or an "exploiter." Explorers can either shirk or experiment with a novel method whose probability of success is initially unknown, while exploiters can either shirk or apply a conventional method with known success probability. The principal does not observe the agent's type or actions, only the sequence of successes and failures, and offers a menu of contracts from which the agent selects. The principal thus faces both the tension between exploration and exploitation inherited from bandit problems and the tension between working and shirking in standard agency models,

while simultaneously needing to design contracts that screen explorers from exploiters.

I solve the screening problem by reformulating it as an equivalent two–principal problem that separates incentive design by type and decouples the original optimization into two independent cost–minimization problems—one for a principal screening exploiters, and one for a principal screening explorers. Each contract is characterized by a three–step procedure: first, construct the least–cost contract that induces the desired behavior under moral hazard; second, adjust the contract to discourage imitation by the other type; and third, deliver any required surplus in a way that minimizes rent leakage. This procedure yields two distinct contracts, each tailored to its target type.

The optimal contract for an exploiter resembles a standard pay-for-performance scheme: it rewards each success promptly and penalizes failure with low current wages. In contrast, the contract that attracts explorers shifts rewards to the long run. Since exploration often involves early failures followed by uncertain success, the optimal scheme backloads compensation on favorable performance trajectories, sometimes even rewarding early failure if it is informative. These differences reflect both motivation and selection: the explorer contract supports risk-taking and experimentation while discouraging imitation by exploiters, for whom delayed contingent payoffs or rewards for early failure are less attractive. As a result, the principal can screen effectively through the shape of the payment profile.

Commitment plays very different roles across the two contracts. For the exploiter, commitment is largely irrelevant: the optimal contract relies heavily on short-term incentives that can be implemented period by period. For the explorer, commitment is essential. The principal must credibly promise to reward eventual success. Without this ability, the principal is forced to front-load incentives, making the contract more at-

tractive to exploiters and undermining separation. In the absence of commitment, exploration becomes more difficult to implement or altogether infeasible.

Termination similarly affects the two contracts in distinct ways. For the exploiter, the threat of termination after a failure is a powerful disciplinary tool that reduces the cost of providing effort incentives. It also strengthens screening, as explorers are more likely to face early failure and thus avoid such contracts. For the explorer, however, termination undermines the value of continued experimentation. While the option to terminate can be used to motivate effort or reduce shirking, it lowers the option value of learning and recovery and may distorts selection. In some cases, the principal may optimally commit to continuation even after failure, accepting short-run inefficiencies in order to encourage innovation.

The value of performance feedback also differs across contracts. In the exploitation contract, withholding interim feedback imposes no cost on the agent and helps deter explorers from mimicking: without knowledge of early outcomes, a misreporting explorer cannot adapt his strategy and receives lower expected payoffs. Witholding interim feedback thus strengthens screening without affecting incentives for the exploiter. In the exploration contract, by contrast, timely feedback is essential. It allows the agent to learn from failures and adjust course in the second period, increasing the expected payoff of experimentation. Providing feedback does not raise incentive costs, but improves efficiency and reinforces separation.

Together, these results show that the tools used to motivate exploration—backloaded rewards, tolerance for early failure, job security, commitment, and feedback—also help select for it. The contracts that induce innovation look fundamentally different from those that drive routine execution, and they appeal to different types of workers. In equilibrium, the principal offers a menu of contracts that motivates effort while sort-

ing agents according to their comparative advantage in exploration versus exploitation.

The idea that incentive pay can sort employees by ability and preferences has long been recognized. Early theories by Salop and Salop (1976), Harris and Raviv (1978), and Lazear (1986) showed that firms could use compensation contracts to induce self-selection of workers. A series of empirical studies confirms significant sorting effects of performance pay. For example, Lazear (2000) documents that switching to piece rates at Safelite Glass increased average productivity in part by attracting more productive workers, and Banker et al. (2001) find that a pay-for-performance plan in a retail firm attracted and retained higher-performing employees. Laboratory experiments by Eriksson and Villeval (2008), Cadsby et al. (2007), and Dohmen and Falk (2011) likewise show that workers self-select into incentive schemes based on their ability, risk attitudes, and other traits, leading to higher output under performance-based pay largely due to this sorting effect.

Another line of research examines how to design incentives to encourage innovation under moral hazard. Holmstrom (1989) argues that because innovative activities have noisier performance measures, the optimal scheme to motivate innovation should be more tolerant of early failures (i.e. "low-powered" incentives). In a related vein, Aghion and Tirole (1994) emphasize the difficulty of contracting on unpredictable innovation outcomes and derive an allocation of control rights that spurs innovation. More recent models explicitly consider the explore–exploit trade-off in incentive design: Manso (2011) shows that the optimal contract to motivate innovation involves substantial tolerance for early failure, reward for long-term success, and managerial job security – features that distinguish it from standard pay-for-performance contracts. In equilibrium, the principal offers a menu of contracts that sorts agents by their comparative advantage in exploration vs. exploitation. Relatedly, Hellmann (2007) and Hellmann and Thiele (2011) study innovation incentives in a

multi-tasking framework, highlighting the trade-offs between encouraging creative exploration and incentivizing routine execution.

This work also connects to the literature on bandit problems in economics, which formalize the exploitation–exploration tension. Multi-armed bandit models (see Berry and Fristedt (1985) for an introduction) analyze how decision-makers optimally experiment with uncertain options. Bergemann and Valimaki (2006) survey applications of bandit models across economics. In the context of innovation, Weitzman (1979) presents a classic model of searching for the best alternative, and March (1991) introduced the terminology of "exploration" and "exploitation" in organizational learning. A number of economic models have applied bandit frameworks to R&D and technology adoption problems. For example, Roberts and Weitzman (1981), Jensen (1982), Bhattacharya et al. (1986), and Moscarini and Smith (2001) use Bayesian learning models to study the sequential exploration of new projects and the adoption of innovations. In macroeconomic growth theory, models by Jovanovic and Rob (1990), Jovanovic and Nyarko (1996), and Aghion (2002) depict technological progress as the outcome of firms optimally balancing exploratory innovation and exploitative improvement of existing techniques. Bolton and Harris (1999) analyze strategic experimentation in a game with multiple firms, showing that information spillovers can lead to underexperimentation (free-riding) relative to the social optimum. Our analysis differs from these papers by focusing on the design of incentive contracts – rather than exogenously given strategies – to induce the desired balance between exploration and exploitation.

Finally, a growing recent literature extends these insights to dynamic incentive design with learning. Bergemann and Hege (2005) study the financing of innovation when investors learn about project quality over time; they show that staged funding with the option to abandon can provide efficient incentives, although their model does not feature an explicit explore/exploit choice by the agent. Several papers analyze optimal

long-term contracts that encourage experimentation. Halac et al. (2016) characterize the optimal contract for an agent who can experiment with a new project or stick to a status-quo project, and they find that incentive provision requires back-loaded rewards and the threat of termination to motivate experimentation. Halac et al. (2017) examine the design of contests for innovation, showing that prize-based tournaments can induce competing agents to experiment efficiently. Other work explores incentive design under private information in dynamic innovation settings: for instance, Gomes et al. (2016) analyze a model in which a firm has hidden information about its returns to exploration versus exploitation, and they show that the investor can screen types by offering contracts with different tolerance for failure (akin to our model's distinction between "explorer" and "exploiter" contracts).

More broadly, the analysis underscores the classic tension between exploiting well-established capabilities, which yields immediate returns, and exploring novel possibilities, which entails risky experimentation but is indispensable for long-run innovation. Because the practices that maximize exploitation often undermine those needed for exploration, firms typically adopt two broad approaches to this trade-off: integration and separation (Tushman and O'Reilly, 1996; O'Reilly and Tushman, 2013). Integration strategies embed exploratory activity within the firm's existing operations by giving employees slack time and backloaded rewards. For example, 3M's "15% time" and Google's "20% time" policies allow employees to pursue self-directed projects during part of their work hours, and many firms run open innovation contests with backloaded prizes on top of standard compensation. These policies create room for experimentation while keeping innovation inside the core organization. By contrast, separation strategies establish distinct units or laboratories, such as Bell Labs and Xerox PARC, focused exclusively on exploratory work and insulated from day-to-day operations. When motivation is the main constraint, integrated schemes can be effective: slack time and organization-

wide contests can spur exploration even within a single operational unit. When selection is the obstacle (exploration requires workers with different aptitudes), structural separation becomes essential, allowing the firm to attract and retain such "explorers" under tailored incentives while the rest of the organization remains focused on exploiting established capabilities.

The paper is organized as follows. Section II introduces the baseline principal–agent model of exploration and exploitation. Section III develops the solution approach, showing how the optimal contracts can be constructed using a three–step procedure based on decoupling and likelihood–ratio analysis. Section IV characterizes the optimal explorer and exploiter contracts and highlights the resulting selection effects. Section V studies implementation with and without commitment on the principal's side, while Section VI analyzes the optimal use of termination after failure. Section VII examines how the provision or withholding of feedback reshapes incentives and selection. Section VIII discusses empirical implications and applications to corporate governance, bankruptcy regimes, and intrapreneurship in large organizations. Section IX concludes. All proofs are in the Appendix.

II The Model

This section develops a two–period principal–agent model in which an agent may be either an explorer or an exploiter. In each period, an explorer can either exert effort in a novel, exploratory activity or shirk, while an exploiter can either operate a conventional technology or shirk. Effort generates a stochastic outcome, success or failure, and exploratory effort also updates beliefs about the profitability of the novel method, as in Manso (2011). The principal observes only the history of successes and failures, not the agent's type or effort, and must therefore design contracts that both motivate effort and separate explorers from exploiters.

To organise the analysis, I work with two closely related formulations of the contracting problem. The first is the original screening problem with a single principal, who offers a menu of contracts to a heterogeneous pool of agents. In this one–principal model, the principal chooses a pair of contracts and associated action plans to maximise expected profit subject to incentive compatibility, self–selection, and participation constraints.

The second formulation is an auxiliary two–principal problem. Here, each principal designs a contract with the goal of attracting and motivating one type—Principal E targets explorers and Principal X targets exploiters—while recognising that both types can in fact accept either contract and misreport their type. Each auxiliary problem is a cost–minimisation program for a given target type, subject to within–type incentive compatibility, ride–along constraints that limit the other type's payoff from misreporting, and participation constraints.

I show below that the two formulations are equivalent for separating allocations: any optimal menu in the one–principal problem can be supported by a pair of contracts that solve the two–principal problem, and any solution to the two–principal problem that is consistent with the induced utilities implements an optimal menu in the one–principal model. This equivalence allows me to characterise the optimal contracts by working with the decoupled explorer and exploiter subproblems in the remainder of the paper.

A One-Principal Model

Time is two periods. In each period the agent chooses an action and produces either a success (S) or a failure (F). If the agent takes action $i \in \{0, 1, 2\}$, the probability of success is p_i, which may be unknown to both parties. Let $E[p_i]$ denote the prior mean and $E[p_i \mid S, j]$, $E[p_i \mid F, j]$ the posterior means after observing S or F from action j. Learning is

action-specific:

$$E[p_j] = E[p_j \mid S, i] = E[p_j \mid F, i] \quad \text{for all } j \neq i.$$

Action 1 represents a conventional method with a known success probability

$$p_1 = E[p_1] = E[p_1 \mid S, 1] = E[p_1 \mid F, 1].$$

Action 2 is a novel method with an uncertain success probability p_2. Let $E[p_2]$ denote the prior mean and $E[p_2 \mid S, 2]$ the posterior mean after observing a success with the novel approach. I assume that success raises beliefs about the effectiveness of the new approach:

$$E[p_2] \;<\; E[p_2 \mid S, 2].$$

After a failure, the agent can learn from the unsuccessful attempt, abandon that particular variant of action 2, and try a fresh novel method drawn from the same prior. As a reduced-form representation, I denote by $E[p_2 \mid F, 2]$ the continuation success probability if the agent explores again after a first-period failure with action 2. In the baseline environment with feedback I assume that the agent can fully reset,

$$E[p_2 \mid F, 2] \;=\; E[p_2].$$

The exploratory ordering

$$E[p_2] \;<\; p_1 \;<\; E[p_2 \mid S, 2]$$

ensures that the novel approach is initially less attractive than the conventional one, but can become strictly superior following a success.

When I later consider environments in which the principal withholds feedback on interim performance, I will relax the reset assumption and allow impaired learning after failure by letting the continuation success

probability satisfy

$$E[p_2 \mid F, 2] \;<\; E[p_2],$$

capturing the idea that without feedback the agent may remain locked into a less promising variant of action 2.

Action 0 corresponds to shirking and yields success with probability $p_0 = 0$. All other actions considered in the model, including the conventional approach (action 1) and any version of the novel approach (action 2, both at the prior and after success or failure), have success probabilities strictly between zero and one.

Taking action $i \in \{0, 1, 2\}$ in a period entails a private cost $c_i \geq 0$, with $c_0 = 0$, $c_1 > 0$, and $c_2 > 0$. The principal and the agent are risk neutral and share a unit discount factor. Actions are unobservable (moral hazard).

A contract specifies wages $w = \{w_F, w_S, w_{SF}, w_{SS}, w_{FF}, w_{FS}\}$ contingent on the history of successes and failures. Limited liability imposes a minimum wage $w_x \geq \underline{w}$ for all outcomes $x \in \{F, S, SF, SS, FS, FF\}$. It is convenient to write

$$w_x = \underline{w} + \tilde{w}_x, \qquad \tilde{w}_x \geq 0,$$

and to refer to $\tilde{w} = \{\tilde{w}_F, \tilde{w}_S, \tilde{w}_{SF}, \tilde{w}_{SS}, \tilde{w}_{FF}, \tilde{w}_{FS}\}$ as the vector of *bonuses above the floor*. The principal can commit to a long-term contract. Later we will consider the case in which the principal cannot commit to a long-term contract.

An *action plan* is denoted $\langle {}_i{}^j_k \rangle$, where i is the first-period action, j the second-period action after a first-period success, and k the second-period action after a first-period failure. The principal's total expected revenue

and the agent's expected wage payments under plan $\langle {}_{i}{}^{j}_{k} \rangle$ are

$$R(\langle {}_{i}{}^{j}_{k} \rangle) = \{E[p_i]S + (1 - E[p_i])F\} + E[p_i]\{E[p_j \mid S, i]S + (1 - E[p_j \mid S, i])F\} + (1 - E[p_i])\{E[p_k \mid F, i]S + (1 - E[p_k \mid F, i])F\},$$

$$\begin{aligned} W(\tilde{w}, \langle {}_{i}{}^{j}_{k} \rangle) = {} & \{E[p_i]\tilde{w}_S + (1 - E[p_i])\tilde{w}_F\} \\ & + E[p_i]\{E[p_j \mid S, i]\tilde{w}_{SS} + (1 - E[p_j \mid S, i])\tilde{w}_{SF}\} \\ & + (1 - E[p_i])\{E[p_k \mid F, i]\tilde{w}_{FS} + (1 - E[p_k \mid F, i])\tilde{w}_{FF}\} + 2\underline{w}, \end{aligned}$$

and the agent's total effort cost is

$$C(\langle {}_{i}{}^{j}_{k} \rangle) = c_i + E[p_i]\, c_j + (1 - E[p_i])\, c_k.$$

Before contracting, the agent's type $\theta \in \{\mathrm{E}, \mathrm{X}\}$ is privately known. The prior probability of type E is $\Pr(\theta = \mathrm{E}) = \mu \in (0, 1)$. Types differ only in the costly actions they can perform. Explorers cannot exploit ($c_1 = +\infty$), and exploiters cannot explore ($c_2 = +\infty$). The corresponding sets of feasible action plans are

$$\mathcal{A}(\mathrm{E}) = \{\langle {}_{i}{}^{j}_{k} \rangle : i, j, k \in \{0, 2\}\}, \qquad \mathcal{A}(\mathrm{X}) = \{\langle {}_{i}{}^{j}_{k} \rangle : i, j, k \in \{0, 1\}\}.$$

Termination is excluded in the base model: the agent always remains employed for two periods. We will later consider what happens when termination is allowed.

Each type $\theta \in \{\mathrm{E}, \mathrm{X}\}$ has an exogenous reservation utility $\overline{U}_{\theta}^{\mathrm{com}} \geq 0$, which captures competition for talent and outside offers. An agent who declines all contracts receives $\overline{U}_{\theta}^{\mathrm{com}}$.

Given a contract $\tilde{w}$ and an action plan $\langle {}_{i}{}^{j}_{k} \rangle$, the utility of an agent is

$$U(\tilde{w}, \langle {}_{i}{}^{j}_{k} \rangle) = W(\tilde{w}, \langle {}_{i}{}^{j}_{k} \rangle) - C(\langle {}_{i}{}^{j}_{k} \rangle),$$

and the principal's profit is

$$\Pi(\tilde{w}, \langle {}_{i}{}^{j}_{k} \rangle) = R(\langle {}_{i}{}^{j}_{k} \rangle) - W(\tilde{w}, \langle {}_{i}{}^{j}_{k} \rangle).$$

We will focus on separating equilibria of this game. In the original screening problem there is a single principal who offers a menu of contracts $\mathcal{M} = \{\tilde{w}^{\mathrm{E}}, \tilde{w}^{\mathrm{X}}\}$ and simultaneously chooses the action plan $a_\theta \in \mathcal{A}(\theta)$ that each type is expected to follow under its designated contract. The principal's problem can therefore be written as

$$\begin{aligned}
\max_{\tilde{w}^{\mathrm{E}}, \tilde{w}^{\mathrm{X}}, a_{\mathrm{E}}, a_{\mathrm{X}}} \quad & \mu\, \Pi\big(\tilde{w}^{\mathrm{E}}, a_{\mathrm{E}}\big) + (1-\mu)\, \Pi\big(\tilde{w}^{\mathrm{X}}, a_{\mathrm{X}}\big) \\
\text{s.t.} \quad & U_{\mathrm{E}}\big(\tilde{w}^{\mathrm{E}}, a_{\mathrm{E}}\big) \geq U_{\mathrm{E}}\big(\tilde{w}^{\mathrm{E}}, \langle {}_{l}{}^{m}_{n} \rangle\big), \quad \forall \langle {}_{l}{}^{m}_{n} \rangle \in \mathcal{A}(\mathrm{E}), && (\mathrm{IC})_E \\
& U_{\mathrm{X}}\big(\tilde{w}^{\mathrm{X}}, a_{\mathrm{X}}\big) \geq U_{\mathrm{X}}\big(\tilde{w}^{\mathrm{X}}, \langle {}_{l}{}^{m}_{n} \rangle\big), \quad \forall \langle {}_{l}{}^{m}_{n} \rangle \in \mathcal{A}(\mathrm{X}), && (\mathrm{IC})_X \\
& U_{\mathrm{E}}\big(\tilde{w}^{\mathrm{E}}, a_{\mathrm{E}}\big) \geq \sup_{\langle {}_{l}{}^{m}_{n} \rangle \in \mathcal{A}(\mathrm{E})} U_{\mathrm{E}}\big(\tilde{w}^{\mathrm{X}}, \langle {}_{l}{}^{m}_{n} \rangle\big) && (\mathrm{RAC})_E \\
& U_{\mathrm{X}}\big(\tilde{w}^{\mathrm{X}}, a_{\mathrm{X}}\big) \geq \sup_{\langle {}_{l}{}^{m}_{n} \rangle \in \mathcal{A}(\mathrm{X})} U_{\mathrm{X}}\big(\tilde{w}^{\mathrm{E}}, \langle {}_{l}{}^{m}_{n} \rangle\big), && (\mathrm{RAC})_X \\
& U_{\theta}\big(\tilde{w}^{\theta}, a_{\theta}\big) \geq \overline{U}_{\theta}^{\mathrm{com}}, \qquad \tilde{w}^{\theta} \geq 0, \qquad \theta \in \{\mathrm{E}, \mathrm{X}\}. && (\mathrm{PC})
\end{aligned}$$

Solving this problem directly is not very transparent, because every change in one contract affects the other type's misreporting incentives. The next subsection therefore introduces an auxiliary problem with two principals that captures the same economic trade-offs but separates the explorer and exploiter contracts.

B The Two–Principal Problem

There are two hidden types. Type E (explorer) can take costly action 2 (and 0), whereas type X (exploiter) can take costly action 1 (and 0). Their feasible action-plan sets, $\mathcal{A}(\mathrm{E})$ and $\mathcal{A}(\mathrm{X})$, are as in the model. In the original screening problem, a single principal wishes to implement a pair

of target plans $(a_{\mathrm{E}}, a_{\mathrm{X}}) \in \mathcal{A}(\mathrm{E}) \times \mathcal{A}(\mathrm{X})$ by offering a menu $\mathcal{M} = \{\tilde{w}^{\mathrm{E}}, \tilde{w}^{\mathrm{X}}\}$ and solving (**??**).

To highlight the structure of the optimal menu, consider instead an auxiliary environment with two distinct principals. Principal E faces only explorer agents and designs a contract $\tilde{w}^{\mathrm{E}}$ to screen among their feasible action plans $\mathcal{A}(\mathrm{E})$. Principal X faces only exploiter agents and designs a contract $\tilde{w}^{\mathrm{X}}$ for $\mathcal{A}(\mathrm{X})$. Each principal takes as given four numbers

$$(\overline{U}_{\mathrm{E}}^{\mathrm{own}}, \overline{U}_{\mathrm{E}}^{\mathrm{mis}}, \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \overline{U}_{\mathrm{X}}^{\mathrm{mis}}),$$

which summarise, for each type, the utilities that can be attained on either contract when following any feasible action plan. In this auxiliary problem, these numbers play the role of exogenous caps in the ride–along constraints.

Because the original principal faces both moral hazard and adverse selection, the full screening problem cannot be solved as a single linear program. It is convenient to decouple it into two simpler cost–minimisation problems—one for each type—linked by the requirement that neither type wishes to mimic the other. In each subproblem, the relevant principal minimises the expected wage cost of implementing the designated plan for that type subject to three families of constraints. The *incentive-compatibility* constraints (IC) ensure that the agent of a given type prefers the intended plan to all other feasible plans within its action set. The *ride-along* constraints (RAC) prevent the other type from benefiting by misreporting and then following any plan consistent with its own feasibility set. Finally, the *participation* constraints (PC) guarantee that each agent obtains at least its effective participation utility, taking into account both competitive offers and the option to imitate the other contract.

Given a candidate explorer contract $\tilde{w}^{\mathrm{E}}$, the cost–minimising contract that implements a_{X} for type X solves

$$\begin{aligned}
\min_{\tilde{w}^{\mathrm{X}}\geq 0} \quad & W(\tilde{w}^{\mathrm{X}}, a_{\mathrm{X}}) \\
\text{s.t.} \quad & U_{\mathrm{X}}(\tilde{w}^{\mathrm{X}}, a_{\mathrm{X}}) \geq U_{\mathrm{X}}(\tilde{w}^{\mathrm{X}}, \langle {}_{l}{}^{m}_{n} \rangle), \qquad \forall \langle {}_{l}{}^{m}_{n} \rangle \in \mathcal{A}(\mathrm{X}) \quad \text{(IC)} \\
& \sup_{\langle {}_{i}{}^{j}_{k} \rangle \in \mathcal{A}(\mathrm{E})} U_{\mathrm{E}}(\tilde{w}^{\mathrm{X}}, \langle {}_{i}{}^{j}_{k} \rangle) \leq \overline{U}_{\mathrm{E}}^{\mathrm{own}}, \qquad \text{(RAC)} \\
& U_{\mathrm{X}}(\tilde{w}^{\mathrm{X}}, a_{\mathrm{X}}) \geq \max\{\overline{U}_{\mathrm{X}}^{\mathrm{com}},\ \overline{U}_{\mathrm{X}}^{\mathrm{mis}}\}, \qquad \text{(PC)}.
\end{aligned}$$

Symmetrically, given a candidate exploiter contract $\tilde{w}^{\mathrm{X}}$, the cost–minimising contract that implements a_{E} for type E solves

$$\begin{aligned}
\min_{\tilde{w}^{\mathrm{E}}\geq 0} \quad & W(\tilde{w}^{\mathrm{E}}, a_{\mathrm{E}}) \\
\text{s.t.} \quad & U_{\mathrm{E}}(\tilde{w}^{\mathrm{E}}, a_{\mathrm{E}}) \geq U_{\mathrm{E}}(\tilde{w}^{\mathrm{E}}, \langle {}_{i}{}^{j}_{k} \rangle), \qquad \forall \langle {}_{i}{}^{j}_{k} \rangle \in \mathcal{A}(\mathrm{E}) \quad \text{(IC)} \\
& \sup_{\langle {}_{l}{}^{m}_{n} \rangle \in \mathcal{A}(\mathrm{X})} U_{\mathrm{X}}(\tilde{w}^{\mathrm{E}}, \langle {}_{l}{}^{m}_{n} \rangle) \leq \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \qquad \text{(RAC)} \\
& U_{\mathrm{E}}(\tilde{w}^{\mathrm{E}}, a_{\mathrm{E}}) \geq \max\{\overline{U}_{\mathrm{E}}^{\mathrm{com}},\ \overline{U}_{\mathrm{E}}^{\mathrm{mis}}\}, \qquad \text{(PC)}.
\end{aligned}$$

A menu $(\tilde{w}^{\mathrm{E}}, \tilde{w}^{\mathrm{X}})$ together with a set of four numbers $(\overline{U}_{\mathrm{E}}^{\mathrm{own}}, \overline{U}_{\mathrm{E}}^{\mathrm{mis}}, \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \overline{U}_{\mathrm{X}}^{\mathrm{mis}})$ are called *consistent* if those numbers coincide with the actual utilities generated by the menu in the original single–principal screening problem, in the sense that

$$\overline{U}_{\mathrm{E}}^{\mathrm{own}} = \sup_{\langle {}_{i}{}^{j}_{k} \rangle \in \mathcal{A}(\mathrm{E})} U_{\mathrm{E}}(\tilde{w}^{\mathrm{E}}, \langle {}_{i}{}^{j}_{k} \rangle), \quad \overline{U}_{\mathrm{E}}^{\mathrm{mis}} = \sup_{\langle {}_{i}{}^{j}_{k} \rangle \in \mathcal{A}(\mathrm{E})} U_{\mathrm{E}}(\tilde{w}^{\mathrm{X}}, \langle {}_{i}{}^{j}_{k} \rangle),$$

$$\overline{U}_{\mathrm{X}}^{\mathrm{own}} = \sup_{\langle {}_{l}{}^{m}_{n} \rangle \in \mathcal{A}(\mathrm{X})} U_{\mathrm{X}}(\tilde{w}^{\mathrm{X}}, \langle {}_{l}{}^{m}_{n} \rangle), \quad \overline{U}_{\mathrm{X}}^{\mathrm{mis}} = \sup_{\langle {}_{l}{}^{m}_{n} \rangle \in \mathcal{A}(\mathrm{X})} U_{\mathrm{X}}(\tilde{w}^{\mathrm{E}}, \langle {}_{l}{}^{m}_{n} \rangle).$$

These equalities define the fixed point for the four cap values used in the two linear programs; they do not require the corresponding ride-along *inequalities* to bind at the optimum.

For later reference, it is useful to introduce the *participation utility* of type $\theta \in \{\mathrm{E}, \mathrm{X}\}$,

$$\overline{U}_\theta \;\equiv\; \max\{\overline{U}_\theta^{\mathrm{com}},\; \overline{U}_\theta^{\mathrm{mis}}\},$$

which summarizes both competition for talent and the option to imitate the other contract. In the figures below, the horizontal axis is labelled by $\overline{U}_\theta$.

Given these definitions, the original one-principal's optimization problem can be approached separately for each type once the four constants $(\overline{U}_{\mathrm{E}}^{\mathrm{own}}, \overline{U}_{\mathrm{E}}^{\mathrm{mis}}, \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \overline{U}_{\mathrm{X}}^{\mathrm{mis}})$ are treated as parameters that summarise the interaction between the two subproblems.

LEMMA 1 (Equivalence of one– and two–principal formulations)*: Any optimal separating menu in the single–principal problem* (**??**) *induces a menu* $(\tilde{w}^{\mathrm{E}}, \tilde{w}^{\mathrm{X}})$ *and a vector* $(\overline{U}_{\mathrm{E}}^{\mathrm{own}}, \overline{U}_{\mathrm{E}}^{\mathrm{mis}}, \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \overline{U}_{\mathrm{X}}^{\mathrm{mis}})$ *that are feasible and consistent solutions for the two–principal problem. Conversely, any feasible and consistent solution to the two auxiliary cost–minimisation problems generates a separating menu that is optimal for* (**??**).

The proof, provided in Appendix A, is a straightforward mapping between the constraints of the original screening problem and the caps in the two auxiliary problems. Lemma 1 implies that solving the two–principal problem is the first step towards solving the original screening problem: the rest of the analysis therefore focuses on characterising the explorer and exploiter subproblems and tracing how their optimal contracts vary with the four cap values and with the participation utilities $\overline{U}_{\mathrm{E}}$ and $\overline{U}_{\mathrm{X}}$.

While Lemma 1 establishes a formal equivalence between the one– and two–principal formulations for separating allocations, it is important to emphasise that the two–principal problem is not a literal reformulation of the original environment. In particular, the equivalence holds at the level of outcomes—optimal menus and induced utilities—but not necessarily at

the level of decentralised procedures. The two principals in the auxiliary problems operate independently and do not coordinate their contracts. As a result, some separating outcomes that are implementable by a single coordinating principal may not be sustainable when each principal controls only one contract. For example, if separation requires delivering surplus on the misreporting type's contract to satisfy cross–type incentives, this adjustment is feasible under a unified menu but not within an isolated subproblem. The two–principal formulation is thus best viewed as an analytical device: it decomposes a complex screening problem into tractable subproblems while preserving global consistency through the utility caps. The equivalence ensures that solving these subproblems yields valid solutions to the original problem, but only when the resulting contracts jointly satisfy the fixed–point consistency conditions that stand in for coordination.

III Solution Approach

This section presents the tools used to solve the decoupled subproblems introduced above. The principal faces both hidden actions and hidden types, so the contracting problem combines moral hazard and adverse selection. The decoupling result allows me to treat the explorer and exploiter contracts separately, taking as given the four cross–type utility caps and two participation utilities:

$$(\overline{U}_{\mathrm{E}}^{\mathrm{own}}, \overline{U}_{\mathrm{E}}^{\mathrm{mis}}, \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \overline{U}_{\mathrm{X}}^{\mathrm{mis}}, \overline{U}_{\mathrm{E}}, \overline{U}_{\mathrm{X}}).$$

Each type–specific problem is then solved using a three–step procedure that progressively adjusts the wage schedule to satisfy incentive compatibility, separation, and participation at minimum cost.

Throughout this construction, I use likelihood–ratio analysis to determine which outcomes are most informative of the intended behavior. For

any two plans π and π' and any node h of the outcome tree, define

$$\Lambda_{\pi,\pi'}(h) \;=\; \frac{\Pr(h \mid \pi)}{\Pr(h \mid \pi')}.$$

Here π denotes the *target* plan (for the type whose contract is being designed), and π' denotes a comparison plan. In Step 1, π' represents a deviation by the *same* type. In Steps 2 and 3, π' represents the best–response plan of the *other* type when misreporting and accepting this contract. A higher value of $\Lambda_{\pi,\pi'}(h)$ means that outcome h is more indicative of the target plan π relative to the comparison plan π', since it occurs relatively more often under π than under π'. Events with high likelihood ratios will therefore be the most useful for allocating incentives and surplus.

Step 1: Baseline. The first step ignores participation and ride–along constraints and solves the principal's problem of motivating the intended type to follow the target plan π at minimum expected wage cost. Formally, the principal minimizes $W(\tilde{w}, a_\theta)$ subject only to type–θ's incentive–compatibility constraints. This yields a *baseline contract* that implements π for type θ, abstracting from selection considerations.

Likelihood–ratio analysis guides which outcomes should be rewarded in this step. For any within–type deviation $\pi' \in \mathcal{A}(\theta)$, the most cost–effective way to enforce π is to concentrate rewards on outcomes h with high $\Lambda_{\pi,\pi'}(h)$. Each dollar paid at such a node increases the agent's expected utility under π much more than under π', thereby tightening the relevant incentive constraint at minimal cost. In particular, if an outcome can occur under π but never under π' (so $\Pr(h \mid \pi') = 0$), then $\Lambda_{\pi,\pi'}(h) = +\infty$; these "infinite–LR" events are lexicographically most informative and are rewarded first when ruling out the deviation π'. Applying this logic across all relevant deviations yields the least–cost incentive–compatible wage schedule for type θ. If multiple baseline contracts attain the same mini-

mum cost, the principal breaks ties in favor of contracts that give lower utility to the misreporting type and that pay the agent earlier.

Step 2: Poison. The second step addresses cross–type misreporting. Starting from the baseline contract, the principal checks whether it violates the ride–along constraint for the other type θ'. If the baseline contract already satisfies

$$U_{\theta'}(\tilde{w}, a_\theta) \;\leq\; \overline{U}_{\theta'}^{\text{own}},$$

then it is already sufficiently unattractive to θ' and no poisoning is required. If instead the baseline delivers too much utility to θ', the principal must modify the wage schedule to make the contract less appealing to misreporters while preserving type–θ's incentive compatibility.

Let π denote the target plan for type θ and let π' denote the best–response plan of the other type θ' when misreporting and taking this contract. Likelihood–ratio analysis then identifies which outcomes are relatively easy for the misreporter to attain (low $\Lambda_{\pi,\pi'}(h)$) and which are relatively hard (high $\Lambda_{\pi,\pi'}(h)$). To *poison* the contract, the principal shifts payments away from outcomes with low likelihood ratios toward outcomes with high likelihood ratios. This reallocation preserves the intended type's incentives, because it moves rewards to events that are still likely under π, but it reduces the misreporter's expected utility, because those same events are relatively unlikely under π'. When the ride–along constraint is active, the poisoned contract is chosen so that the misreporter's utility is driven down exactly to the cap $\overline{U}_{\theta'}^{\text{own}}$.

In some cases, no such distortion exists that both keeps type θ on the target plan and satisfies the ride–along constraint for θ'. In the one–principal model, the burden of deterring misreporting can then be shifted to the other contract in the menu. In the two–principal formulation, by contrast, each principal controls a single contract and cannot adjust the other, so infeasibility of poisoning directly translates into the

absence of a separating equilibrium for that subproblem at the given parameter values.

Step 3: Deliver. The final step deals with participation. If the intended type θ already receives at least its participation utility, $U_\theta(\tilde{w}, a_\theta) \geq \overline{U}_\theta$, then the contract constructed in Steps 1–2 is feasible and cost–minimizing. Otherwise, the principal must raise U_θ up to $\overline{U}_\theta$ by delivering additional surplus in a way that preserves incentive compatibility and keeps misreporting unattractive.

Again, let π be the target plan and π' the misreporting type's best–response plan under the current contract. For any small adjustment $\Delta w(h)$ to the wage at node h, the marginal change in the intended type's expected utility is proportional to $\Pr(h \mid \pi)\Delta w(h)$, whereas the marginal change in the misreporter's utility is proportional to $\Pr(h \mid \pi')\Delta w(h)$. The ratio of these effects is exactly $\Lambda_{\pi,\pi'}(h)$. To raise U_θ at minimum cost while leaking as little as possible to θ', the principal therefore allocates additional payments to outcomes with high $\Lambda_{\pi,\pi'}(h)$: these are events the intended type reaches much more frequently than the misreporter. By concentrating surplus on such nodes and adjusting along IC–preserving directions (for example, along indifference arcs where the binding incentive constraints remain tight), the principal can satisfy the participation constraint *exactly* while keeping the ride–along constraint either slack or just binding.

When poisoning was necessary in Step 2, surplus delivery can be combined with further reallocation to even more strongly discourage misreporting, but the underlying principle is the same: surplus is delivered through outcomes that are most informative of the intended type's effort and least attractive to the misreporter.

Together, these three steps—baseline, poison, and deliver—decompose the complex screening and moral hazard problems into a sequence of likelihood–ratio–guided adjustments that can be applied to either type.

The resulting contract satisfies all incentive, ride–along, and participation constraints and minimizes the principal's expected wage bill. The next sections apply this construction to derive the exploiter and explorer contracts.

LEMMA 2 (Optimality of the Three–Step Construction)*: Fix a target type* $\theta \in \{\mathrm{E}, \mathrm{X}\}$ *and treat the constants*

$$(\overline{U}_{\mathrm{E}}^{\mathrm{own}}, \overline{U}_{\mathrm{E}}^{\mathrm{mis}}, \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \overline{U}_{\mathrm{X}}^{\mathrm{mis}}, \overline{U}_{\mathrm{E}}, \overline{U}_{\mathrm{X}})$$

as given parameters summarising cross–type incentives and participation. Assume the decoupled problem for type θ *is feasible at these parameters. Then the contract obtained through the three–step procedure— (i) baseline cost minimisation subject only to within–type IC; (ii) poisoning adjustments, if necessary, to satisfy the ride–along constraint; and (iii) surplus–delivery adjustments, if necessary, to satisfy the participation constraint— is an optimal solution to the decoupled problem. That is, the resulting contract minimises the principal's expected wage bill subject to implementing the target plan for type* θ*, deterring misreporting by the other type, and satisfying participation.*

IV Optimal contracts

This section derives the optimal explorer and exploiter contracts . Following the approach of Manso (2011), we begin by fixing the principal's target action plans and then solve for the least-cost incentive schemes that implement them under limited liability and hidden actions. Because the two tasks—exploitation and exploration—capture fundamentally different learning environments, they require very different compensation structures. Exploitation involves repeating a well-understood action whose probability of success is known, while exploration requires experimenting

with a new, uncertain method whose success rate can improve through learning. These two activities are represented respectively by the action plans

$$a_{\mathrm{X}} = \left\langle {}_{1}{}^{1}_{1} \right\rangle \qquad \text{and} \qquad a_{\mathrm{E}} = \left\langle {}_{2}{}^{2}_{2} \right\rangle,$$

which are the natural benchmarks for contrasting the provision of short-term effort versus long-term experimentation.

A Exploiter Contract

We first characterize the contract that implements exploitation, $a_X = \left\langle {}_{1}{}^{1}_{1} \right\rangle$. The principal faces a standard two–period moral–hazard problem under limited liability, with the additional complication that an explorer may misreport as an exploiter. The analysis therefore proceeds in three steps. We begin by deriving the least–cost contract that induces $\left\langle {}_{1}{}^{1}_{1} \right\rangle$ for a given exploiter. We then ask whether this baseline contract can be modified to deter a misreporting explorer without violating the exploiter's incentive constraints. Finally, we describe how surplus is delivered to the exploiter when his participation constraint binds, and how this surplus can be allocated so as to leak as little as possible to the explorer.

Since $p_0 = 0$, a success in any period occurs with strictly positive probability under effort and with zero probability under shirking. By the likelihood–ratio rule, the cheapest way to induce effort on any margin is to load the entire incentive differential on the corresponding success node. As a result, the one–period moral–hazard logic applies period by period: the principal pays a flat bonus for every success and nothing after failure. The ratio $\alpha_1 \equiv c_1/p_1$ represents the per–success bonus that makes the exploiter just willing to exert effort in a single period. The least–cost *baseline contract* that implements $\left\langle {}_{1}{}^{1}_{1} \right\rangle$ then pays

$$\tilde{w}_S = \tilde{w}_{SS} = \tilde{w}_{FS} = \frac{c_1}{p_1}, \qquad \tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0,$$

so that the exploiter earns the wage floor in each period and obtains baseline utility $2\underline{w}$, equivalent to earning the minimum wage each period. This contract simply repeats the one–period solution across time. If the ride–along constraint for the explorer and the exploiter's participation constraint are slack, the baseline contract is clearly optimal.

The next question is whether the principal can *poison* the exploiter's contract to deter a misreporting explorer. Intuitively, one might try to reduce the first–period bonus $\tilde{w}_S$ and compensate with a higher continuation bonus $\tilde{w}_{SS}$, thereby shifting rewards away from early success. However, any such rebalancing that leaves the exploiter indifferent along his incentive–compatibility frontier either leaves the explorer's payoff unchanged or strictly increases it, since it promises a larger payment in states (such as two consecutive successes) that are relatively attractive to the explorer. More generally, any attempt to reduce the explorer's value by lowering some success–contingent bonus must either violate the exploiter's incentive constraints or be offset by higher rewards in other success states that the explorer finds at least as attractive. Hence no feasible distortion of the baseline wage profile can reduce an explorer's ride–along payoff without breaking the exploiter's incentive compatibility. If the ride–along constraint binds, it must therefore be enforced on the explorer's own contract rather than through further distortions of the exploiter contract.

Finally, consider the case in which the exploiter's participation constraint binds and his reservation utility exceeds the baseline level, $\overline{U}_X > 2\underline{w}$. The principal must then deliver additional surplus $\Delta = \overline{U}_X - 2\underline{w} > 0$ in a way that maintains exploiter's incentives and minimizes rent leakage to a misreporting explorer. The allocation of this surplus follows the likelihood–ratio logic. If, at the baseline, an explorer who misreports as an exploiter would shirk on some margin (for example, shirk in period 1 or after a particular history), the principal first tightens that margin by raising the corresponding success bonus until the explorer is just willing to exert on that branch, while making the minimal compensating adjust-

ments needed to preserve the exploiter's incentive constraints. Because success against shirking has an effectively infinite likelihood ratio, this step delivers surplus to the exploiter with zero leakage of surplus to the explorer. Once the misreporting explorer would exert effort on all margins under the exploiter contract, any remaining surplus is delivered by increasing the first–period success payment $\tilde{w}_S$ only, keeping $\tilde{w}_{SS}$ and $\tilde{w}_{FS}$ at their incentive–minimal levels. First–period success is then the most diagnostic non–shirking outcome, so concentrating additional rewards on $\tilde{w}_S$ provides the exploiter with the required rent while leaking as little surplus as possible to the explorer.

PROPOSITION 1 (Exploiter Contract)*: The least–cost baseline contract that implements exploitation* $a_X = \langle {}_1{}^1_1 \rangle$ *pays a constant bonus for each success and zero after failure:*

$$\tilde{w}_S = \tilde{w}_{SS} = \tilde{w}_{FS} = \frac{c_1}{p_1}, \qquad \tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0.$$

This baseline contract is optimal if the explorer's ride–along constraint and the exploiter's participation constraint are slack.

No feasible modification of this wage profile can reduce an explorer's ride–along payoff without violating the exploiter's incentive–compatibility constraints. If the ride–along constraint binds, it must therefore be enforced on the explorer's own contract rather than by distorting the exploiter contract.

If the exploiter's participation constraint binds and his reservation utility satisfies $\overline{U}_X > 2\underline{w}$*, the principal delivers the required surplus* $\Delta = \overline{U}_X - 2\underline{w} > 0$ *as follows. First, if a misreporting explorer would shirk on any margin under the baseline contract, the success bonus on that margin is raised, with minimal compensating adjustments to preserve exploiter's incentive constraints, until exertion is optimal for the misreporting explorer. Once the misreporting explorer would exert on all margins, any*

remaining surplus is delivered by increasing $\tilde{w}_S$ only, keeping $\tilde{w}_{SS}$ and $\tilde{w}_{FS}$ at their incentive–minimal levels.

Figure 1 illustrates the optimal exploiter contract for the benchmark parameters.[2] The least–cost baseline contract in Proposition 1 sets all success–contingent bonuses equal to

$$\tilde{w}_S = \tilde{w}_{SS} = \tilde{w}_{FS} = \frac{c_1}{p_1} = \frac{4}{5}, \qquad \tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0,$$

so that under the target plan $\langle {}_{1}{}^{1}_{1} \rangle$ the exploiter's expected utility equals $2\underline{w} = 0.4$.

For this calibration, the misreporting explorer already exerts on all margins at the baseline contract, so the first surplus-delivery step in Proposition 1 is vacuous: no bonus needs to be raised simply to eliminate shirking. When the participation constraint binds and additional surplus $\Delta = \overline{U}_{\mathrm{X}} - 2\underline{w} > 0$ must be delivered, likelihood–ratio comparisons between the target exploiter plan and the explorer's best–response deviation imply that first–period success is the most diagnostic event among non–shirking outcomes; the least–leaky way to raise the exploiter's utility is therefore to increase $\tilde{w}_S$ only, keeping $\tilde{w}_{SS}$ and $\tilde{w}_{FS}$ fixed at their baseline level c_1/p_1. In the numerical solution, the optimal contract thus holds $\tilde{w}_{SS} = \tilde{w}_{FS} = c_1/p_1$ constant and raises $\tilde{w}_S$ linearly with $\overline{U}_{\mathrm{X}}$ once $\overline{U}_{\mathrm{X}}$ exceeds $2\underline{w}$. Figure 1 shows the resulting wage schedule: $\tilde{w}_S$ has a kink at $2\underline{w}$, while the continuation bonuses remain flat.

In summary, to attract and motivate exploiters the principal must offer a contract that closely resembles a standard pay–for–performance scheme. Because exploitation involves repetition of a known technology with predictable outcomes, the optimal incentive relies on immediate and proportional rewards for observable success. The resulting contract pro-

[2]The parameters are $p_0 = 0$, $E[p_2] = 0.3$, $E[p_2 \mid S, 2] = 0.7$, and $p_1 = 0.5$, with $c_1 = 0.4$, $c_2 = 0.2$, and a wage floor $\underline{w} = 0.2$.

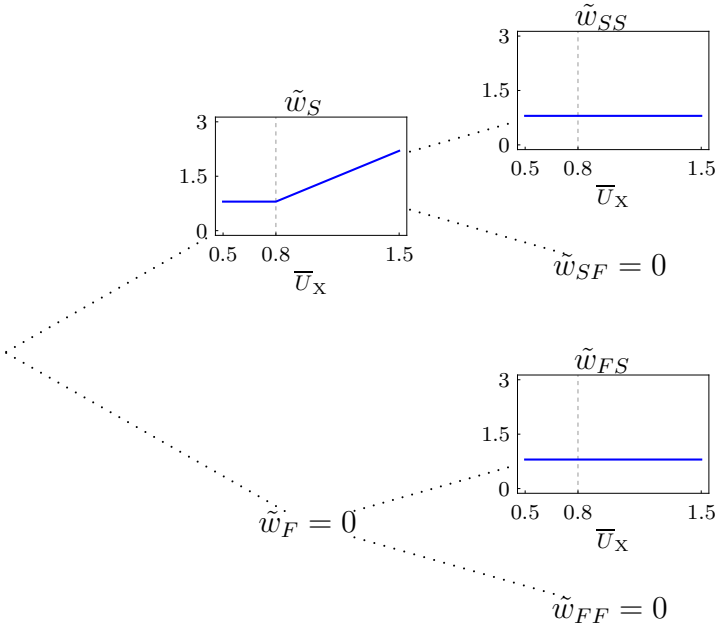


Figure 1. Optimal contract that implements exploitation as a function of the minimal rent $\overline{U}_X$ under the benchmark parameters. Total pay resembles a repetition of standard pay–for–performance across periods; when participation binds, extra surplus is delivered by raising only the first–period success bonus $\tilde{w}_S$.

vides high–powered, short–term incentives that tie compensation tightly to current performance, ensuring continuous effort while leaving only small rents beyond those required for participation. In equilibrium, such a contract attracts agents who are confident in their ability to deliver consistent results and discourages explorers, for whom the same bonus structure yields insufficient expected utility given the high probability of early failures. The exploitation contract therefore performs a dual role: it motivates steady effort and simultaneously screens out agents inclined toward experimentation, mirroring the pay structures commonly observed in routine operational or production tasks.

B Explorer Contract

We now characterize the least–cost contract that implements exploration $a_E = \langle {}_2{}^2_2 \rangle$. The principal faces a two–period moral–hazard problem

under limited liability, with the added complication that an exploiter may misreport as an explorer. As in the exploiter case, the analysis proceeds in three steps. We begin by deriving the least–cost baseline contract that induces $\langle 2\frac{2}{2} \rangle$ for a given explorer. We then ask whether this baseline contract can be modified to deter a misreporting exploiter without violating the explorer's incentive constraints. Finally, we describe how surplus is delivered to the explorer when his participation constraint binds, and how this surplus can be allocated so as to leak as little as possible to the exploiter.

In the baseline case where participation and ride–along constraints are slack, all failure-branch wages are set at the floor, and effort is induced through success-contingent bonuses that separate cleanly by history. After a first-period failure, continuation effort requires a recovery bonus $\tilde{w}_{FS} = c_2/E[p_2]$; after a success, continuation requires a double–success bonus $\tilde{w}_{SS} = (1 + E[p_2])c_2/(E[p_2]\,E[p_2 \mid S, 2])$. As incentive constraints are binding under this contract, the explorer earns baseline utility $2\underline{w}$.

The first–period success payment $\tilde{w}_S$ does not enter any binding incentive constraint and is therefore set at the floor. In principle, period–1 effort could be induced at the same cost via a combination of $\tilde{w}_S$ and $\tilde{w}_{SS}$, but doing so would concentrate rents on early success and make the contract strictly more attractive to a misreporting exploiter, for whom S is more diagnostic. Anticipating this, the baseline loads all first–period incentives on $\tilde{w}_{SS}$ and keeps $\tilde{w}_S$ at the minimum.

Because these wages are already at their incentive-minimal levels for the explorer, there is no feasible "poisoning" of the contract: any rebalancing that preserves $\langle 2\frac{2}{2} \rangle$ either leaves the misreporting exploiter better off or violates the explorer's incentive constraints. Hence, if ride–along binds, it must be enforced on the exploiter's menu, not by distorting the explorer contract.

When the participation constraint binds and the required utility $\overline{U}_E > 2\underline{w}$, the principal must deliver additional surplus $\Delta = \overline{U}_E - 2\underline{w} > 0$. The allocation of this surplus follows a two–step rule. First, if a misreporting exploiter would shirk on any margin, the principal tightens that margin by raising the corresponding success bonus—first $\tilde{w}_{SS}$, and then, if necessary, $\tilde{w}_{FS}$—until exertion is optimal on all branches. This step delivers surplus to the explorer with zero leakage. Once the misreporter exerts everywhere, any remaining surplus is delivered along the least–leaky direction identified by likelihood–ratio comparisons between SS and F.

PROPOSITION 2 (Explorer Contract)*: The least–cost baseline contract that implements exploration* $a_E = \langle {}_2{}^2_2 \rangle$ *pays:*

$$\tilde{w}_S = \tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0, \qquad \tilde{w}_{FS} = \frac{c_2}{E[p_2]}, \qquad \tilde{w}_{SS} = \frac{(1 + E[p_2])\, c_2}{E[p_2]\, E[p_2 \mid S, 2]}.$$

This contract is optimal if the exploiter's ride–along constraint and the explorer's participation constraint are slack.

No feasible modification of this wage profile can reduce an exploiter's ride–along payoff without violating the explorer's incentive–compatibility constraints. If the ride–along constraint binds, it must therefore be enforced on the exploiter's own contract rather than by distorting the explorer contract.

If the explorer's participation constraint binds and his reservation utility satisfies $\overline{U}_E > 2\underline{w}$*, the principal delivers the required surplus* $\Delta = \overline{U}_E - 2\underline{w} > 0$ *as follows. First, if a misreporting exploiter would shirk on any margin under the baseline contract, the success bonus on that margin is raised—first* $\tilde{w}_{SS}$*, then if needed* $\tilde{w}_{FS}$*—until exertion is optimal for the misreporting exploiter. Once the misreporting exploiter would exert on all margins, any remaining surplus is delivered by increasing either:*

- *$\tilde{w}_{SS}$ only, if double success is more diagnostic of exploration, that is, if*

$$\Lambda_{\langle 1\,1\rangle,\langle 2\,2\rangle}(SS) = \frac{p_1^2}{E[p_2]\,E[p_2 \mid S,2]} \;\leq\; \Lambda_{\langle 1\,1\rangle,\langle 2\,2\rangle}(F) = \frac{1-p_1}{1-E[p_2]}\,,$$

- *both $\tilde{w}_{SS}$ and $\tilde{w}_F$ jointly, along the explorer's and exploiter's incentive–compatibility frontiers, if failure is more diagnostic, that is, if*

$$\Lambda_{\langle 1\,1\rangle,\langle 2\,2\rangle}(SS) > \Lambda_{\langle 1\,1\rangle,\langle 2\,2\rangle}(F).$$

Figure 2 illustrates the optimal explorer contract under the benchmark parameters. At the baseline, the contract uses only the second–period bonuses $\tilde{w}_{FS}$ and $\tilde{w}_{SS}$ to induce exploration. As $\overline{U}_E$ rises, the contract first follows Step 1 of Proposition 2, tightening incentives by raising $\tilde{w}_{SS}$ (and, if needed, $\tilde{w}_{FS}$) just enough to induce exertion by a misreporting exploiter. Once effort is optimal on all margins, Step 2 allocates remaining surplus based on likelihood–ratio comparisons: under the benchmark parameters, failure is more informative of exploration than SS, so $\tilde{w}_F$ and $\tilde{w}_{SS}$ rise together along the exploiter IC arc. The resulting wage paths for $\tilde{w}_F$ and $\tilde{w}_{SS}$ exhibit kinks at the point where the contract moves along eliminating shirking at different nodes and minimizing surplus leaks to an exerting misreporting exploiter.

In summary, to attract and motivate explorers the principal must offer a contract that departs sharply from standard pay–for–performance. Because exploration is risky and produces long stretches of early failure, the optimal scheme back–loads rewards onto histories that signal persistent experimentation rather than only early wins. Failure–branch wages sit at the floor whenever they would encourage shirking, but when likelihood–ratio comparisons make first–period failure more informative of exploration than success, the surplus–delivery step can even tilt pay toward F, rewarding short–term failure relative to success. This tolerance

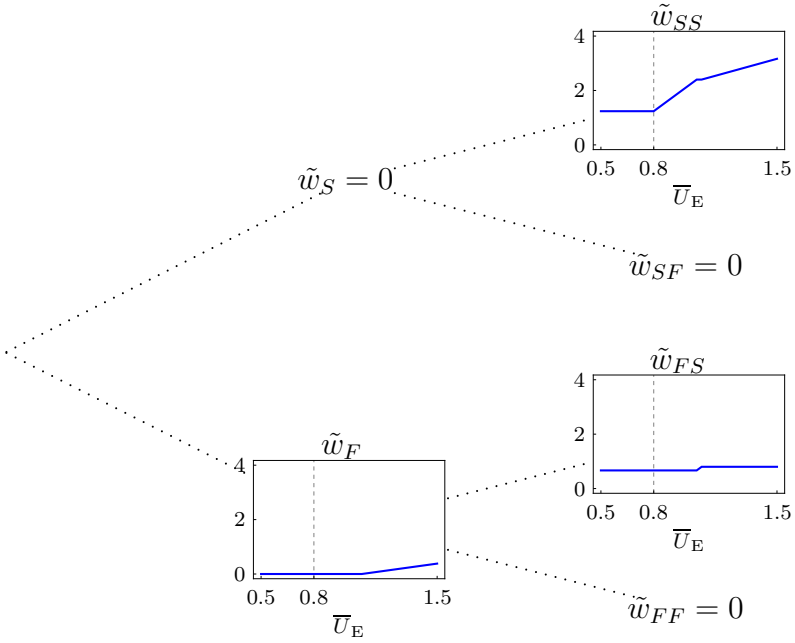


Figure 2. The optimal explorer contract for the target plan $a_{\mathrm{E}} = \langle 2\,2\,0 \rangle$ under limited liability ($w_{\min} = 1$) and parameter values $p_0 = 0.35$, $E[p_2] = 0.45$, $E[p_2 \mid S, 2] = 0.67$, $p_1 = 0.60$, and $c_1 = c_2 = 0.5$. The baseline contract sets $\tilde{w}^0_{SS} = 4.05$, yielding a baseline utility of $\overline{U}^{\mathrm{E}}_{\mathrm{base}} \approx 2.5$. Because $r_{SS} > r_F$, the failure outcome is the more informative signal of exploration: initially, both $\tilde{w}_F$ and $\tilde{w}_{SS}$ rise together along the explorer's IC–preserving ray until the exploiter's constraint binds at $\overline{U}^{*}_{\mathrm{E}} \approx 3.1$. Beyond this point, only $\tilde{w}_{SS}$ increases. The kink in the bonus paths marks this shift from the explorer's to the exploiter's binding constraint.

for early failure, combined with high continuation bonuses (w_{FS}, w_{SS}), makes the contract attractive to explorers—who expect many setbacks on the way to success—while leaving exploiters relatively worse off, since for them the same pattern of payoffs provides weak incentives for steady, predictable performance. In equilibrium, the exploration contract both sustains experimenting effort and helps select agents inclined toward innovation rather than routine execution, complementing the exploitation contract analyzed above.

V Lack of Commitment

In this section we assume that the principal cannot commit to a long–term contract and is restricted to a sequence of one–period contracts. This lack of commitment has different implications for the contracts derived in Section IV. We will consider both the exploiter and explorer contracts.

A Exploitation without commitment

Recall that the baseline contract that implements exploitation relies only on short–term incentives: the principal rewards the agent for current success and pays nothing after a failure. As shown in Manso (2011), a sequence of short–term contracts implements this contract exactly. Consequently, the baseline exploiter contract derived in Proposition 1 does not require commitment.

When the outside option of the exploiter is high, however, the principal must provide additional surplus to satisfy the agent's participation constraint. Proposition 1 describes three ways to provide this surplus. One is to raise the wage after first–period success, $\tilde{w}_S$. Because this payment is made contemporaneously, the resulting contract can still be implemented by a sequence of short–term contracts. The alternatives, which are only used when the misreporting explorer shirks in the second period under the baseline contract, involve increasing the second–period payments after success in the second period following success in the first period or following a failure in the first period, via $\tilde{w}_{SS}$ or $\tilde{w}_{FS}$. Such back–loaded bonuses must be promised at time 0 and are not renegotiation–proof: without commitment, the principal would wish to remove them ex post, and anticipating this, the agent would refuse the contract. We summarise this discussion in the following result.

PROPOSITION 3 (Exploitation without commitment)*: Suppose the principal cannot commit to a long–term contract and wishes to implement exploitation. The baseline contract described in Proposition 1 can be implemented via a sequence of short–term contracts. If the participation constraint requires additional surplus, the principal will raise the payment for a first–period success,* $\tilde{w}_S > 0$*, and this modification continues to be implementable via short–term contracts. In contrast, providing extra surplus through the second–period payments* $\tilde{w}_{SS}$ *or* $\tilde{w}_{FS}$ *is not renegotiation–proof and therefore cannot be implemented without commitment.*

In summary, commitment is irrelevant for the baseline exploitation contract and matters only when the optimal committed contract would back–load surplus onto $\tilde{w}_{FS}$ or $\tilde{w}_{SS}$. In those cases, a principal without commitment must deliver the same surplus through $\tilde{w}_S$ instead, slightly increasing the contract's attractiveness to a misreporting explorer. In the benchmark calibration, surplus is already delivered through $\tilde{w}_S$, so lack of commitment has no bite for exploitation—unlike in the exploration problem, where deferred rewards will be essential to motivate experimentation.

B Exploration without commitment

We now examine how the lack of commitment affects the contract that selects an explorer when the target is $\langle {}_2{}^2_2 \rangle$ (explore in periods 1 and 2). Under commitment, the baseline exploration contract defers compensation to the second period: the agent receives no pay after first success and is rewarded only for second–period success, and obtains a back–loaded surplus w_{SS} to compensate for the high exploration cost. The agent also receives compensation $\tilde{w}_{FS}$ to induce exertion after failure in the first period.

When commitment is not possible, the principal cannot credibly promise back–loaded *rents* after the first period. As a result, w_{SS} can no longer deliver deferred surplus; it must be set at the *minimal* level that still induces second–period exploration *after a first–period success*. The principal must therefore use a first–period success bonus $\tilde{w}_S > 0$ to induce period 1 experimentation. This restores incentives to experiment but makes the contract more attractive to the exploiter, weakening the selection margin. After a first–period failure the target prescribes *exertion*, so wages should be just enough to make exploration optimal at that node.

If the explorer's outside option $\overline{U}_{\mathrm{E}}$ exceeds her baseline utility $2\underline{w}$, the principal must deliver surplus $\Delta = \overline{U}_{\mathrm{E}} - 2\underline{w} > 0$. Without commitment, the only instruments available to the principal are first period bonuses $\tilde{w}_S$ and $\tilde{w}_F$. The optimal delivery strategy depends on the misreporting exploiter's best response.

If the misreporting exploiter exerts effort, the principal increases both $\tilde{w}_F$ and $\tilde{w}_S$ at the same rate. Increasing $\tilde{w}_F$ faster than $\tilde{w}_S$ would induce shirking by the explorer, while increasing $\tilde{w}_S$ faster than $\tilde{w}_F$ would deliver more surplus to the misreporting exploiter than to the explorer. Adding the same increment to both wages keeps the explorer's first–period incentive constraint exactly binding while limiting the rent leakage to the exerting misreporter.

If the misreporting exploiter shirks, the principal initially raises the first–period success wage $\tilde{w}_S$ alone until the exploiter becomes indifferent between shirking and exerting effort. Once the exploiter reaches indifference, additional surplus is delivered by increasing $\tilde{w}_F$ and $\tilde{w}_S$ in parallel. This strategy minimizes leakage to a shirking misreporter and preserves the explorer's first–period incentive compatibility.

PROPOSITION 4 (Exploration without commitment)*: Under no commitment, the optimal baseline contract that implements the explorer's tar-*

get plan $\langle {}_2{}^2_2 \rangle$ *pays* $\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0$*, and:*

$$\tilde{w}^{\,0}_{FS} = \frac{c_2}{E[p_2]}, \qquad \tilde{w}^{\,0}_{SS} = \frac{c_2}{E[p_2 \mid S, 2]}, \qquad \tilde{w}^{\,0}_{S} = \frac{c_2}{E[p_2]}.$$

If the explorer's outside option $\overline{U}_{\mathrm{E}}$ *exceeds the baseline utility* $2\underline{w}$*, the principal needs to provides the additional surplus* $\Delta = \overline{U}_{\mathrm{E}} - \underline{w}$ *as follows:*

(A1) If the misreporting exploiter shirks*, first increase* $\tilde{w}_S$ *alone until the exploiter is indifferent between shirking and exerting, and then continue by raising both* $\tilde{w}_S$ *and* $\tilde{w}_F$ *in parallel.*

(A2) If the misreporting exploiter exerts effort*, increase both* $\tilde{w}_F$ *and* $\tilde{w}_S$ *in parallel.*

This two–stage adjustment preserves the explorer's first–period incentive constraint while minimizing the cost of surplus delivery under no commitment.

Figure 3 summarizes how losing commitment changes both the structure and the selection properties of the explorer contract. With commitment, the principal can implement $\langle {}_2{}^2_2 \rangle$ by back–loading all rents on second–period outcomes $(\tilde{w}_F, \tilde{w}_{SS})$, keeping first–period bonuses at the floor. Without commitment, $\tilde{w}_{FS}$ and $\tilde{w}_{SS}$ must be fixed at their IC–minimal levels and any additional surplus for the explorer has to be front–loaded on first–period wages $(\tilde{w}_S, \tilde{w}_F)$. The optimal surplus–delivery policy depends on whether a misreporting exploiter shirks or exerts: when he shirks, the principal first raises $\tilde{w}_S$ alone to keep imitation unprofitable and then moves $\tilde{w}_S$ and $\tilde{w}_F$ in parallel; when he exerts, both bonuses are increased from the start. This front–loading restores incentives for exploration but makes the explorer menu more attractive to exploiters, weakening the selection margin relative to the committed benchmark.

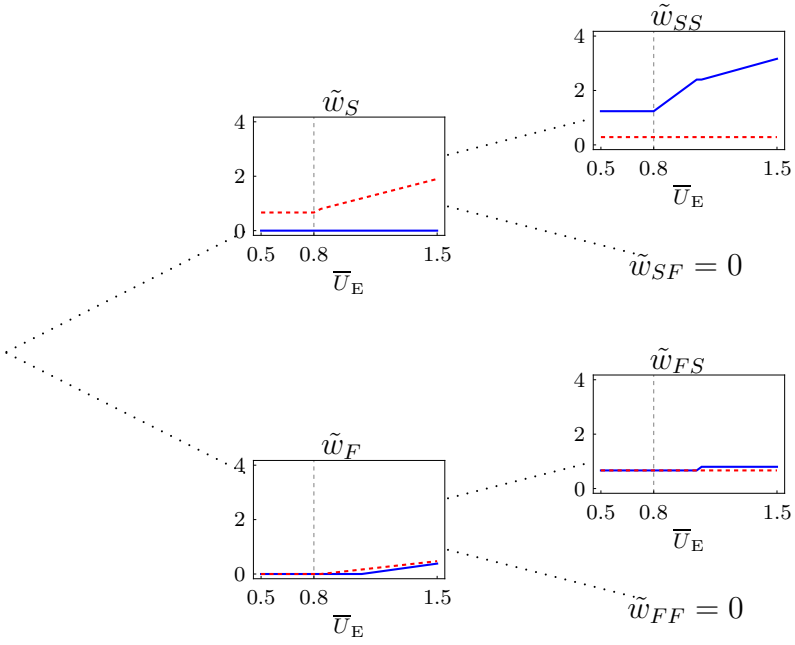


Figure 3. Explorer contract as a function of the minimal rent $\overline{U}_E$ under the benchmark parameters, with (solid blue curves) and without (dashed red curves) commitment. With commitment, extra surplus is back–loaded on second–period bonuses $(\tilde{w}_F, \tilde{w}_{SS})$, whereas without commitment it must be front–loaded on first–period wages (w_S, w_F), making the explorer menu more attractive to exploiters.

Figure 4 plots the ride–along payoff of a misreporting exploiter under commitment (solid) and no commitment (dashed) as a function of $\overline{U}_{\mathrm{E}}$. For low values of $\overline{U}_{\mathrm{E}}$, while the misreporter still shirks, the two schedules overlap. Once the no commitment contract has raised $\tilde{w}_S$ enough that the misreporter starts exerting, the curves diverge: commitment allows the principal to back–load rents onto histories that are informative of exploration through $(\tilde{w}_F, \tilde{w}_{SS})$, whereas without commitment the same surplus must be front–loaded on $(\tilde{w}_S, \tilde{w}_F)$, yielding a strictly higher ride–along value. The gap between the two ride–along schedules shows where commitment has bite—precisely when the contract must simultaneously motivate exploration and deter imitation by an exerting exploiter.

Taken together, the analysis shows that commitment plays very different roles for exploitation and exploration. For the exploiter contract, the baseline implementation of $\left\langle {}_{1}{}^{1}_{1} \right\rangle$ never needs commitment: as long as

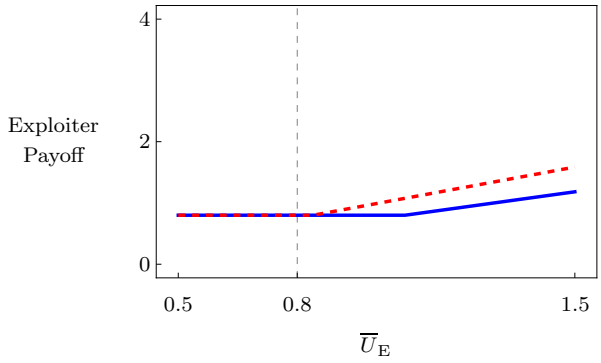


Figure 4. Explorer ride–along payoff under the explorer contract with (solid blue curves) and without (dashed red curves) commitment as a function of the minimal rent $\overline{U}_E$. Commitment allows the principal to defer rents to informative second–period histories, keeping the mimicking payoff low; without commitment, surplus is front–loaded on first–period wages and the exploiter's ride–along value rises, weakening selection.

only effort provision constraints bind, the principal can replicate the optimal scheme period by period. Commitment becomes useful only when extra surplus must be delivered and the misreporting explorer shirks in the second period, allowing the principal to defer rents to histories that are unlikely under shirking and thereby reduce the ride–along payoff. By contrast, under the explorer contract commitment already improves selection at the baseline whenever the misreporting exploiter exerts, because back–loading rents onto informative second–period signals lowers his mimicking value relative to the no–commitment benchmark; if the misreporter shirks at the baseline, commitment starts to matter precisely when the required surplus is high enough to induce him to exert, at which point deferred rewards again strictly tighten screening.

VI Termination

In this section the principal is allowed to terminate the agent after a failure in the first period. Termination can be interpreted as the decision to fire a manager or discontinue funding for a project following poor interim performance. The possibility of termination affects both motivation

and selection. On the motivation side, the threat of dismissal disciplines the agent by increasing the cost of early failure, thereby facilitating effort provision. On the selection side, termination alters the attractiveness of the contract across types: while the threat of job loss deters exploratory individuals who value continuity, it may appeal to exploiters who prefer high–powered, short–term incentives. Hence, termination can serve as a screening device, making the contract less attractive to explorers and more appealing to exploiters. At the same time, termination prevents continuation after an initial setback and may destroy valuable opportunities for learning and recovery. The optimal termination policy therefore trades off its disciplinary and screening benefits against the efficiency losses from premature termination.

A Exploiter Contract with Termination

Allowing the principal to terminate the agent after a first–period failure adds a new disciplinary and selection margin to the exploitation contract. On the incentive side, the threat of dismissal strengthens first–period effort: because failure implies job loss, the agent faces an additional implicit penalty, which allows the principal to reduce monetary rewards. On the selection side, termination changes the composition of self–selecting agents: exploiters, who value high–powered short–term incentives, find such contracts more attractive, while explorers, who value continuation after early setbacks, are discouraged from applying. Termination thus disciplines behavior and helps screen out exploratory types, but it also sacrifices efficient continuation after failure.

The structure of the contract is otherwise similar to the baseline exploitation contract. Since the agent works in the second period only after a first–period success, the "success–after–failure" branch disappears. The continuation bonus $\tilde{w}_{SS}$ remains at its baseline level $\alpha_1 = c_1/p_1$, while the first–period success bonus $\tilde{w}_S$ falls below its previous value, as the

prospect of termination already deters shirking. The overall cost of providing incentives is therefore lower than without termination: the principal rewards success less because poor performance now carries a stronger implicit penalty.

If the exploiter's participation constraint binds, additional surplus $\Delta = \overline{U}_X^{\text{mis}} - (1+p_1)\underline{w}$ must be provided at minimum cost. The optimal adjustment follows the likelihood–ratio rule and depends on the misreporting explorer's behavior at the baseline contract. When the explorer shirks, the principal first raises $\tilde{w}_S$ until the explorer is just willing to exert in the first period, then shifts to $\tilde{w}_{SS}$ if needed. When the explorer exerts at baseline, separation relies directly on the relative likelihood ratios of S and SS. Since $\Lambda_{\langle 1\,{}^1_t\rangle,\langle 2\,{}^2_t\rangle}(S)$ dominates, the principal raises w_S alone.

PROPOSITION 5 (Exploiter contract with termination)*: When the principal can terminate the agent after a first–period failure, the threat of dismissal acts as an additional disciplinary device. The least–cost contract that implements exploitation* $a_{\mathrm{X}} = \langle {}^1_t\rangle$ *under limited liability satisfies*

$$\tilde{w}_{SS} = \frac{c_1}{p_1}, \qquad \tilde{w}_S = \max\left\{\frac{c_1}{p_1} - \underline{w}, 0\right\}, \qquad \tilde{w}_F = \tilde{w}_{FS} = \tilde{w}_{SF} = \tilde{w}_{FF} = 0.$$

If the explorer's outside option U_E^{own} *exceeds the misreport payoff, no adjustment is required. Otherwise, adjustment must be made in the explorer contract.*

If the exploiter's participation constraint binds and his outside utility $\overline{U}_X^{\text{mis}}$ *exceeds the baseline rent* $(1+p_1)\underline{w}$*, the principal must deliver additional surplus* $\Delta = \overline{U}_X^{\text{mis}} - (1+p_1)\underline{w}$ *at minimal cost. The optimal adjustment follows two steps:*

(i) Explorer shirks. *The principal first raises* $\tilde{w}_S$ *and then* $\tilde{w}_{SS}$ *if needed until the explorer is indifferent between exerting and shirking in periods 1 and 2.*

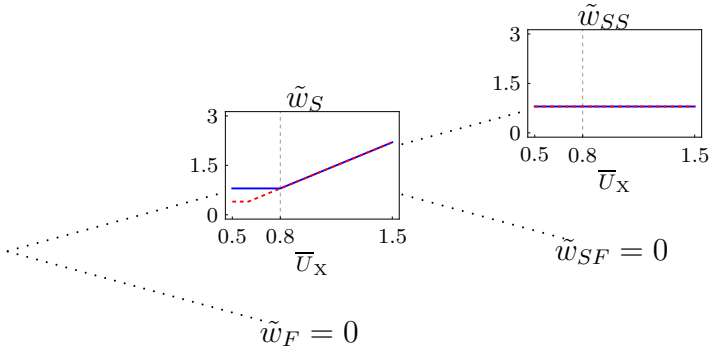


Figure 5. The optimal contracts that implement exploitation without termination (blue solid curves) and with termination (red dashed curves) as a function of the minimal rent $\overline{U}_X$ under the benchmark parameters. With termination, the failure branch is shut down, w_{FS} is set to zero, and incentives are concentrated on the surviving success branch: the first–period success wage w_S falls, while the continuation bonus w_{SS} remains at $\alpha_1 = c_1/p_1$. The threat of dismissal thus substitutes for part of the monetary reward needed to implement exploitation.

(ii) Explorer exerts. *At that point the relevant comparison switches to* $\langle {}_1{}^1_t \rangle$ *versus* $\langle {}_2{}^2_t \rangle$*, for which* $\Lambda_{\langle {}_1{}^1_t \rangle, \langle {}_2{}^2_t \rangle}(S) > \Lambda_{\langle {}_1{}^1_t \rangle, \langle {}_2{}^2_t \rangle}(SS)$*. The remaining surplus is then delivered by increasing* $\tilde{w}_S$*, leaving* $\tilde{w}_{SS}$ *unchanged.*

Figure 5 displays the optimal exploiter contract as a function of the minimal rent $\overline{U}_X$ that must be delivered. Relative to the no–termination benchmark, allowing dismissal after failure lowers the kink at which the exploiter's participation binds, reduces the first–period success bonus w_S for any given rent, keeps the continuation bonus w_{SS} at its baseline level $\alpha_1 = c_1/p_1$, and shuts down w_{FS}. The firing threat thus substitutes for part of the monetary carrot on the $\{S, SS\}$ branch.

Figure 6 plots the expected wage cost of implementing exploitation on the surviving success branch $\{S, SS\}$, with and without termination, as a function of $\overline{U}_X$. Termination shuts down continuation after failure, so the only relevant wage nodes are S and SS. For any given rent requirement, the firing threat allows the principal to lower the first–period success bonus, shifting the entire cost schedule on $\{S, SS\}$ downward relative to the no–termination benchmark. Beyond the kink where participation

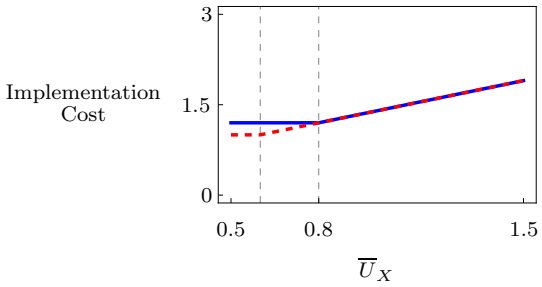


Figure 6. The expected wage cost of implementing exploitation on the success branch $\{S, SS\}$ without termination (blue solid line) and with termination (red dashed line) as a function of the minimal rent $\overline{U}_X$. Termination removes continuation pay after failure and strengthens first–period discipline, shifting the entire cost schedule on $\{S, SS\}$ downward at the baseline. For higher rent requirements, additional surplus is delivered one–for–one through success pay, so the two cost schedules are the same beyond the second kink.

binds, additional surplus is delivered one–for–one through success pay, so the two schedules are parallel: termination keeps implementation cheaper on this branch whenever the threat of dismissal remains operative.

Allowing dismissal after an early failure lets the principal substitute the threat of job loss for part of the monetary carrot. With termination, the period–2 branch disappears following F, so only the S– and SS/SF–nodes matter for implementation. The firing threat tightens the period–1 margin and lowers the bonus that must be paid upon first–period success; at the baseline, the present value reduction in expected wages paid over $\{S, F, SS, SF\}$ equals $p_1 \min\{\alpha_1, \underline{w}\}$, where $\alpha_1 = c_1/p_1$. For higher participation levels, extra surplus can be delivered such that termination never raises implementation cost on these nodes and weakly lowers it whenever the first–period stick still substitutes for the carrot.

Termination is also selection–enhancing. Removing the failure branch eliminates the continuation value that both types could secure after F, but this loss is typically larger for the explorer: the explorer reaches F more often and, absent termination, can insure that branch by shirking in period 2 to collect (at least) the floor. Consequently, termination depresses the misreporting explorer's value more than the exploiter's, tightening

ride–along on the explorer side. Finally, if the exploiter's participation binds under termination, surplus can be restored at least cost and with minimal leakage by raising w_S only: each dollar delivered at S accrues with probability p_1 to the exploiter, while the misreporter's success probability is strictly lower (either p_0 or $E[p_2]$), so the same transfer raises the explorer's payoff by strictly less. The following corollary states these three points.

COROLLARY 1 (Termination and selection on the exploiter contract)*: Termination after early failure affect the implementation and selection of exploitation in three ways:*

- ***(i)*** *The threat of termination disciplines the agent after early failure and substitutes for monetary incentives, allowing the principal to reduce first–period bonus pay relative to the no–termination baseline.*
- ***(ii)*** *Because explorers fail more frequently in early periods, termination typically reduces their continuation value by more than that of exploiters, making the exploiter contract less attractive to exploratory types.*
- ***(iii)*** *If the exploiter's participation constraint binds, the principal can restore the required surplus while increasing the explorer's expected utility by strictly less, preserving screening against explorers.*

Taken together, Proposition 5 and Corollary 1 show that early termination is a powerful tool for implementing and selecting exploiters. On the incentive side, the threat of dismissal allows the principal to cut back on monetary rewards while preserving effort, lowering the baseline cost of exploitation and keeping the termination contract uniformly cheaper whenever the firing margin binds. On the selection side, removing continuation pay after failure typically hurts explorers more than exploiters and, when additional surplus must be delivered, raising $\tilde{w}_S$ alone restores

the exploiter's rent while leaking strictly less to the misreporting explorer. Termination thus simultaneously strengthens discipline, sharpens separation, and tilts self–selection in favor of exploiters.

B Explorer Contract with Termination

Allowing the principal to terminate the agent after a first–period failure affects exploration in two ways. On the incentive side, the threat of dismissal strengthens effort in period 1: because early failure leads to job loss and the forfeiture of the second–period wage floor, the prospect of continuation becomes valuable, allowing the principal to motivate exploration at a lower monetary cost. On the selection side, termination discourages explorers who expect frequent failures, making the contract relatively more attractive to exploiters. Termination therefore serves as a disciplinary device but may be detrimental to screening.

The structure of the contract is similar to the baseline explorer contract, except that the "failure" branch is truncated: the agent works in the second period only after a first–period success. As in the baseline model, the first–period deviations determine the required bonus. For moderate floors the relevant constraints are those against shirking in one or both periods, yielding a minimal double–success bonus $\tilde{w}^0_{SS}$ that declines with $\underline{w}$. When $\underline{w}$ is high, the first–period constraints relax, and the second–period incentive constraint binds instead.

As before, the ride–along constraint cannot be relaxed by lowering $\tilde{w}_{SS}$ and compensating with $\tilde{w}_S$, since increasing $\tilde{w}_S$ leaks strictly more rent to a misreporting exploiter. The baseline contract already minimizes the exploiter's payoff, so any binding RAC must be enforced on the exploiter's side of the menu.

When the explorer's participation constraint binds, the principal provides additional surplus $\Delta = \bar{U}_{\mathrm{E}} - (1 + E[p_2])\underline{w}$ through the most diag-

nostic outcome according to the likelihood–ratio rule. Comparing

$$\Lambda_{\langle 1\,{}^1_1\rangle,\langle 2\,{}^2_t\rangle}(SS) = \frac{p_1^2}{E[p_2]E[p_2 \mid S,2]} \quad \text{and} \quad \Lambda_{\langle 1\,{}^1_1\rangle,\langle 2\,{}^2_t\rangle}(F) = \frac{1-p_1}{1-E[p_2]},$$

the principal raises $\tilde{w}_{SS}$ alone when the double–success event is more diagnostic of exploration, or uses a combination of $\tilde{w}_{SS}$ and a small failure bonus $\tilde{w}_F$ when failure is relatively more informative. The adjustment proceeds along incentive–indifference arcs that keep the explorer's and exploiter's constraints binding, ensuring that any required surplus is delivered at minimal cost while maintaining separation between types.

PROPOSITION 6 (Optimal explorer contract with termination)*: Consider an explorer with target plan* $\langle {}_{2\,t}^{2}\rangle$ *(termination after a first–period failure). The optimal baseline contract sets*

$$w_F = w_{SF} = w_{FF} = w_S = \underline{w}, \qquad w_{SS} = \underline{w} + \tilde{w}^0_{SS},$$

with a single success–contingent bonus

$$\tilde{w}^0_{SS} = \max\left\{ \frac{c_2}{E[p_2 \mid S,2]}\,,\ \frac{(1+E[p_2])\,c_2 - E[p_2]\ \underline{w}}{E[p_2]\ E[p_2 \mid S,2]} \right\}.$$

If the participation constraint binds, the deliver surplus $\Delta = \overline{U}_{\mathrm{E}} - (1 + E[p_2])\underline{w} > 0$ *as follows.*

Step A *If the misreporting exploiter shirks on any margin, raise* $\tilde{w}_{SS}$ *until he exerts in both periods.*

Step B *Once the misreporting exploiter exerts on all margins, compare the likelihood ratios of exploitation vs. exploration:*

$$\Lambda_{SS} = \Lambda_{\langle 1\,{}^1_t\rangle,\langle 2\,{}^2_t\rangle}(SS) = \frac{p_1^2}{E[p_2]\ E[p_2 \mid S,2]}, \qquad \Lambda_F = \Lambda_{\langle 1\,{}^1_t\rangle,\langle 2\,{}^2_t\rangle}(F) = \frac{1-p_1}{1-E[p_2]}.$$

- *If $\Lambda_{SS} \leq \Lambda_F$ (double success more diagnostic of exploration), raise $\tilde{w}_{SS}$ only.*
- *If $\Lambda_{SS} > \Lambda_F$ (failure more diagnostic), share surplus between F and SS along IC indifference arcs: first, while misreporting exploiter strictly prefers shirking in period 1, move along the* explorer *IC,*

  $$E[p_2]E[p_2 \mid S, 2]\, d\tilde{w}_{SS} \;=\; E[p_2]\, d\tilde{w}_F \iff E[p_2 \mid S, 2]\, d\tilde{w}_{SS} = d\tilde{w}_F,$$

  then, once misreporting exploiter is indifferent between shirking and exerting in period 1, move along the exploiter *IC under termination,*

  $$p_1^2\, d\tilde{w}_{SS} = p_1\, d\tilde{w}_F \iff p_1\, d\tilde{w}_{SS} = d\tilde{w}_F.$$

Termination after a first–period failure truncates the F-branch and removes the second–period wage floor on that branch. This creates an additional *implicit penalty* for early failure and therefore relaxes the explorer's incentive constraints: the principal can substitute the firing threat for part of the monetary reward (e.g., a lower w_{SS} suffices at the baseline). However, termination also destroys *insurance* after failure. When the explorer's participation constraint binds, the principal must compensate the explorer for the lost continuation value, and that restoration must be done through the remaining nodes in an incentive–feasible way. As a result, although termination lowers incentive costs at the baseline, it can *increase* the overall cost of implementing exploration once participation is binding.

The selection effect works through *exposure to the failure branch* together with the continuation surplus that is lost there. Under exploration, first–period failure occurs with probability $1-E[p_2]$, whereas a misreporting exploiter reaches failure with probability $1-p_1$ if he exerts in period 1 (and $1-p_0$ if he shirks). In the typical region in which the misreporting exploiter's continuation surplus after failure is pinned down at the floor

(e.g., he shirks after F or his post–failure incentive constraint binds), the per–failure loss from termination is essentially the same across types, so the explorer's higher failure exposure implies a larger expected loss. More generally, the comparison depends on both failure probabilities and the continuation surplus that each type enjoys after F, which is why the corollary states the selection effect as *typical* rather than universal.

Finally, when the explorer's participation constraint binds under termination, the principal must restore the explorer's lost utility while limiting rent leakage to a misreporting exploiter. The least–leaky restoration uses the most diagnostic outcomes, as identified by likelihood ratios. If failure is the more informative history ($\Lambda_F < \Lambda_{SS}$), incentive feasibility forces restoration to rely on a combination of w_F and w_{SS} rather than on w_F alone. In that region, full restoration of the explorer's pre–termination utility necessarily raises the misreporting exploiter's expected utility *above* his own pre–termination level, weakening selection against exploiters.

COROLLARY 2 (Termination and selection on the explorer contract)*: Termination after early failure can distort the implementation and selection of exploration in three ways.*

- ***(i)*** *Although termination relaxes incentive constraints, it can increase the cost of implementing exploration once the explorer's participation constraint binds, because the principal must compensate for the loss of continuation value.*
- ***(ii)*** *Termination typically reduces the continuation value of exploratory agents more than that of exploitative agents, since explorers face early failure more frequently and therefore lose more insurance from continuation.*
- ***(iii)*** *In some environments, restoring the explorer's lost utility after termination cannot be done without raising the exploiter's expected utility above his pre-termination level.*

Figures 7 and 8 depict how termination modifies the structure and the cost of the contract that implements exploration. When termination is available, a first–period failure ends employment and the agent forfeits the second–period floor $\underline{w}$. This loss of continuation value relaxes the first–period incentive constraint in the baseline: the principal can induce initial exploration with a smaller double–success bonus w_{SS}, so the baseline implementation cost is lower under termination. Once the participation constraint binds, however, part of the surplus the agent would have received through the guaranteed floor after failure must be delivered explicitly. In line with Proposition **??**, the principal allocates this extra rent between w_{SS} and the failure payment w_F so as to minimize leakage to a misreporting exploiter, using the likelihood–ratio rule. Because the floor is lost when the agent is fired, rewarding failure becomes useful sooner under termination and w_F must rise to offset the missing continuation floor.

The cost comparison mirrors this logic. For small participation requirements, termination lowers the cost of implementation because the threat of losing $\underline{w}$ after failure substitutes for part of the monetary incentive. As the required surplus $\overline{U}_{\mathrm{E}}^{\mathrm{mis}}$ increases, however, the need to compensate the agent for the lost floor through w_F makes the marginal cost under termination steeper. Consequently, the termination schedule starts below the no–termination schedule but eventually overtakes it as $\overline{U}_{\mathrm{E}}^{\mathrm{mis}}$ grows, reflecting that termination is discipline–enhancing at the baseline yet rent–intensive when substantial surplus must be delivered.

Taken together, the results highlight an important asymmetry in how termination operates under exploitation and exploration. In both cases, the threat of dismissal after early failure substitutes job loss for part of the monetary carrot, lowering the baseline cost of implementation. However, termination is a more powerful instrument on the exploiter side. Under the exploiter contract, it cuts more into the misreporting explorer's continuation value than into the exploiter's own payoffs and, when participation

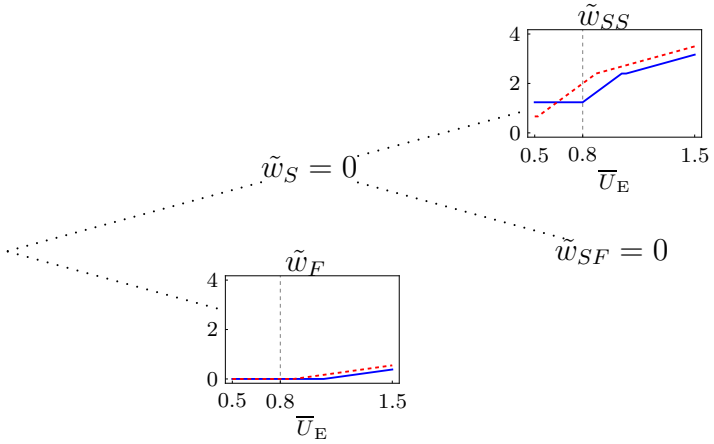


Figure 7. The optimal explorer contracts that implement exploration without termination (blue solid curves) and with termination after a first–period failure (red dashed curves), as a function of the required rent $\overline{U}_{\mathrm{E}}$ under the benchmark parameters. With termination, a first–period failure ends employment and the agent loses the second–period floor on the failure branch, so the double–success bonus w_{SS} can be reduced at the baseline. When participation binds, the principal delivers extra surplus by adjusting w_{SS} and, when failure is informative, w_F along the likelihood–ratio frontier.

binds, the principal can restore the exploiter's rent at least cost while leaving the explorer no better off than before. Under the explorer contract, by contrast, termination also reduces the explorer's continuation value but its selection benefits are weaker: when the misreporting exploiter exerts and failure is the most diagnostic outcome, any full restoration of the explorer's rent necessarily makes the exploiter strictly better off as well. Thus, dismissal after failure is effective for motivating and selecting exploiters, while it can even work against separation under exploration.

VII Feedback

In this section the principal privately observes interim performance at the end of the first period but can choose whether to reveal it to the agent. The decision to provide feedback affects the agent's ability to condition future actions on past outcomes and, therefore, the structure and cost of the optimal contract. When feedback is withheld, the agent cannot distinguish between success and failure and must choose a single

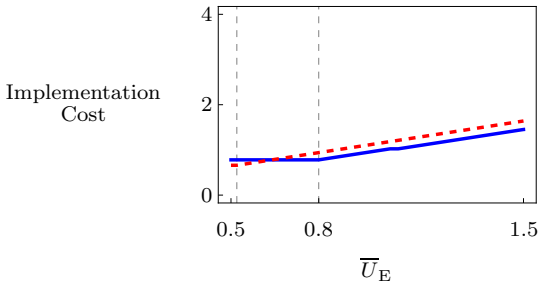


Figure 8. Expected wage cost of implementing exploration on the nodes $\{S, F, SS, SF\}$ without termination (blue solid line) and with termination (red dashed line), as a function of the required rent $\overline{U}_{\mathrm{E}}$. By removing the second–period floor after early failure, termination sharpens first–period discipline and lowers the baseline implementation cost when rent requirements are low. For higher rent requirements, additional surplus is delivered one–for–one through success and failure pay, so the two cost schedules are parallel and the termination schedule remains above the no–termination benchmark.

continuation strategy for both histories. Providing feedback, in contrast, allows the principal to tailor incentives to each realized outcome and the agent to condition his second-period action on first-period output.

A Exploiter contract without feedback

When the principal observes interim performance but withholds it from the agent, the agent cannot distinguish whether the first period resulted in success or failure. Second–period actions must therefore be chosen without regard to past outcomes. This restriction removes all interim conditioning from the contract and effectively reduces the environment to one of repeated effort with hidden outcomes. From the agent's perspective, the employment relationship proceeds "in the dark": he receives no signal about his early performance and thus faces the same continuation value after both success and failure. When withholding feedback, the principal must choose $w_S = w_F$.

In Manso (2011), withholding feedback helps to economize on incentives: when effort is costly and $p_0 > 0$, concealing interim outcomes prevents the agent from gaming second–period behavior and allows the

principal to reuse the same reward across histories, lowering implementation cost.[3] In the present framework, however, there is no such motivational gain. Because shirking never produces success when $p_0 = 0$, the principal can already implement repeated effort at minimal rent $2\underline{w}$ to the agent. Therefore, hiding the interim signal does not reduce the cost of implementing exploitation. The no–feedback contract simply replicates a pay–for–performance scheme that pays a constant bonus per success, yielding the same total expected wage as under full feedback.

The value of withholding feedback instead comes from selection. By keeping the agent uninformed about early results, the principal prevents a misreporting explorer from adjusting actions after a first–period failure. An explorer who accepts the exploitation contract under no feedback must choose the same second–period action after both success and failure and is therefore stuck with a low expected return after failure, since $E[p_2 \mid F, 2] < E[p_2]$. This loss of flexibility makes the contract strictly less attractive to explorers while leaving exploiters indifferent. Hence withholding feedback "poisons" the exploitation contract as a screening device. The principal will therefore always choose not to provide feedback when implementing exploitation.

PROPOSITION 7 (Optimal exploiter contract without feedback)*: (i) Baseline. The least–cost contract that implements exploitation under limited liability satisfies*

$$\tilde{w}_S = \tilde{w}_F = \tilde{w}_{FF} = 0, \qquad \tilde{w}_{SF} = \tilde{w}_{FS} = \frac{c_1}{p_1}, \qquad \tilde{w}_{SS} = \frac{2c_1}{p_1}.$$

(ii) Ride–along and selection. The ride–along constraint is therefore strictly easier to satisfy under no feedback: the explorer's expected utility falls while the exploiter's incentives and cost remain unchanged. In this sense,

[3] A similar result appears in Lizzeri, Meyer, and Persico (2002), who show that suppressing performance evaluations can reduce incentive costs in repeated moral–hazard settings.

withholding feedback strictly poisons *the exploitation contract for the explorer, making misreporting unattractive without affecting implementation costs for the target type.*

(iii) Surplus delivery. If the exploiter's participation constraint binds, additional surplus can be delivered through the mixed–success outcomes $\{SF, FS\}$*, which are less informative of misreporting than double success* SS*. Along the exploiter's incentive–compatible arc,*

$$p_1^2\, d\tilde{w}_{SS} + p_1(1-2p_1)\, d\tilde{w}_{SF} = 0,$$

so increases in $\tilde{w}_{SF} = \tilde{w}_{FS}$ *must be accompanied by offsetting increases (or decreases) in* $\tilde{w}_{SS}$ *to preserve incentives. This adjustment raises expected utility at minimum cost and leaks the least rent to a misreporting explorer.*

Figures 9 and 10 illustrate how withholding feedback affects the exploiter contract and the explorer's ride–along payoff. In Figure 9, the blue solid curves show the full–feedback contract from Section IV, while the red dashed curves plot the optimal no–feedback contract from Proposition 7, as functions of the minimal rent $\overline{U}_X$ required for the exploiter. Under no feedback, the agent cannot condition on the first–period outcome, so $w_S = w_F$ and the principal replicates a simple pay–for–performance scheme that pays a constant bonus per success. The overall slope of the schedules is the same in both regimes: each additional unit of rent is delivered one–for–one through success–contingent pay, so the expected implementation cost is unchanged.

Figure 10 focuses on selection by plotting the explorer's ride–along payoff when he misreports and accepts the exploiter contract, with feedback (blue solid) and without feedback (red dashed), as a function of $\overline{U}_{\mathrm{E}}$. With feedback, the explorer can react to a first–period failure by switching to exploration in period 2 and enjoys a high continuation value on the failure branch. Without feedback, he receives no information about interim

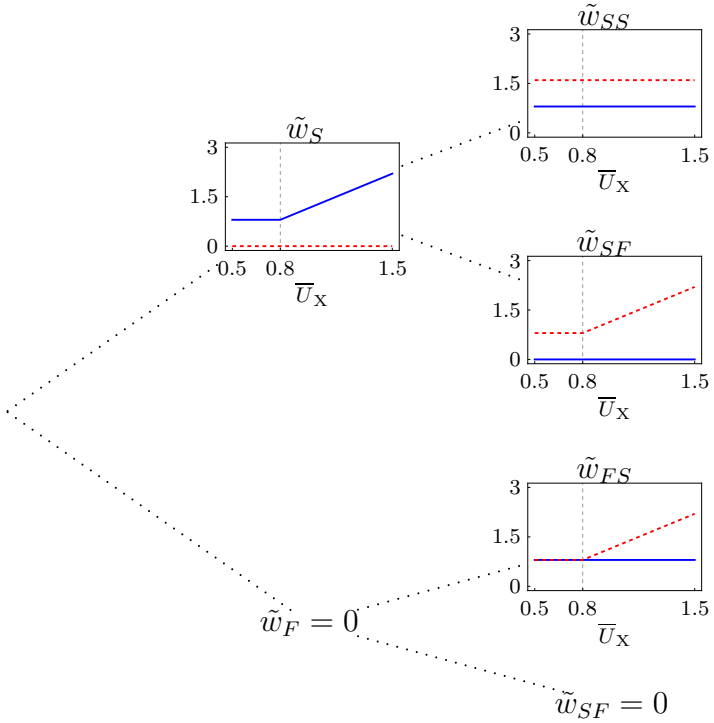


Figure 9. The optimal contracts that implement exploitation with feedback (blue solid curves) and without feedback (red dashed curves), as a function of the minimal rent $\overline{U}_X$ under the benchmark parameters. Under no feedback the agent cannot distinguish success from failure, so $w_S = w_F$ and the principal uses a simple pay–for–performance scheme that pays the same bonus per success across histories. The pattern of wages across nodes changes, but the total expected payment per unit of rent is the same as under full feedback.

performance and must choose the same second–period action after both success and failure; he is therefore stuck with the lower expected return $E[p_2 \mid F, 2] < E[p_2]$ after failure. This loss of flexibility shifts the entire ride–along schedule downward for the explorer, while the exploiter's expected payoff and incentive costs remain the same.

In sum, when $p_0 = 0$ the role of feedback in the exploiter contract is purely about selection, not motivation. With or without feedback, the principal can implement repeated exploitation at the same minimal expected cost by paying a constant bonus per success. What changes is the value of misreporting: suppressing the interim signal forces an explorer who rides along on the exploiter contract to choose the same sec-

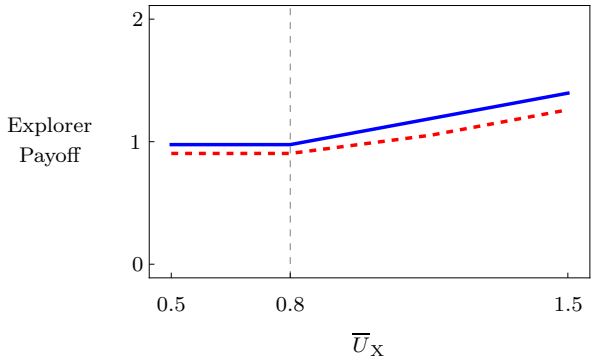


Figure 10. Explorer ride–along payoff under the exploiter contract with feedback (blue solid line) and without feedback (red dashed line), as a function of the required exploiter rent $\overline{U}_{\mathrm{X}}$. When feedback is provided, a misreporting explorer can respond to a first–period failure by tilting second–period effort toward exploration and enjoys a high continuation value. Without feedback, he must choose the same second–period action after both success and failure and is stuck with a low expected return after failure, so his ride–along payoff is strictly lower at all rent levels. The exploiter's incentives and implementation cost are unaffected, so withholding feedback strictly poisons the exploiter contract for explorers.

ond–period action after success and failure and thus forgo the option value of switching to exploration after an early setback. Withholding feedback therefore leaves the exploiter indifferent while strictly reducing the explorer's ride–along payoff, making the exploitation menu less attractive to explorers without altering implementation costs for the target type.

B Explorer contract without feedback

When feedback is withheld, the agent cannot distinguish whether the first-period experiment succeeded or failed. As a result, he must commit to a fixed second-period action regardless of the realized outcome. In particular, if the agent explores again after failure, he continues with the same instance of action 2 and faces a lower success probability $E[p_2 \mid F, 2] < E[p_2]$.

By contrast, if feedback is provided, the agent learns whether the original approach was successful and can adapt after failure. In that case,

he reoptimizes and selects a new instance of action 2, restoring the continuation success probability to $E[p_2]$. An incentive contract with the same implementation cost can support this adaptive plan, and the agent remains indifferent across regimes. However, the expected project value is strictly higher when feedback is available.

PROPOSITION 8 (Feedback strictly preferred for exploration)*: Suppose the principal implements exploration using either: (i) a no-feedback contract that induces plan* $\langle {}_2 2_2^2 \rangle$*, or (ii) the feedback contract derived in Proposition 1 that implements adaptive continuation via* $\langle {}_2 2_2^2 \rangle$*. In both cases, the agent's expected utility and the principal's incentive cost are identical. However, with feedback the agent resets action 2 after failure and restores the success probability to* $E[p_2]$*, whereas without feedback he is locked into a fixed continuation probability* $E[p_2 \mid F, 2] < E[p_2]$*. Feedback therefore increases expected project value. The principal strictly prefers to provide feedback when implementing exploration.*

Proposition 8 shows that, unlike in the exploiter case, withholding feedback is never attractive when the principal wishes to implement exploration. Because the agent cannot update his beliefs about the success of the initial experiment, he is forced to continue with the same instance of action 2 after failure and is stuck with the lower continuation success probability $E[p_2 \mid F, 2]$. The principal can still induce $\langle {}_2 2_2^2 \rangle$ at the same expected cost as under feedback, but the value of the project falls: the contract induces exploration in both periods but destroys the agent's option to abandon unsuccessful trajectories and restart the search.

Taken together, the analysis of feedback reveals a sharp asymmetry between exploitation and exploration. For the exploiter contract, withholding feedback leaves implementation costs unchanged but strictly poisons the menu for explorers, since it prevents a misreporting explorer from conditioning second–period behavior on early failure. The princi-

pal therefore prefers no feedback when implementing exploitation. For the explorer contract, by contrast, feedback does not change incentives or rents but raises expected project value by allowing the agent to adapt after a failed experiment. The principal strictly prefers to provide feedback when implementing exploration. In equilibrium, feedback is used to preserve option value in exploration and withheld to harden screening on the exploitation side.

VIII Empirical implications

The model yields empirical predictions about how different incentive practices shape the composition of the workforce engaged in exploration versus exploitation. The central mechanism is selection: contracts that tolerate early failure and reward long-term success attract individuals with a comparative advantage in exploration, whereas standard pay–for–performance schemes appeal more to exploiters. As a result, changes in incentive regimes should be reflected not only in innovative output, but also in who chooses to work under those regimes. These predictions are relevant in a variety of settings, including executive compensation and corporate governance, bankruptcy codes and creditor rights, social insurance and unemployment benefits, venture capital and entrepreneurial finance, and academic funding and evaluation systems. In addition, the analysis has implications for organizational design, shedding light on when firms should integrate exploration and exploitation within the same units and when they should separate them into distinct structures.

A Empirical Tests of the Model

The selection model suggests two complementary empirical tests. First, one can examine whether incentive practices used in settings where innovation is more (less) important resemble the exploration (exploitation)

contracts derived in the theory, and whether adopting such practices is associated with higher (lower) innovation. Second, one can study whether the adoption of exploration–type incentive schemes attracts and retains more exploratory workers, that is, whether these contracts act as selection devices for creative talent. The first test is inherently a joint test of motivation and selection, since higher innovation under exploration contracts may arise either because incumbent workers experiment more, or because more innovative workers self–select into these environments. The second test isolates the selection margin by focusing on who joins, leaves, or relocates when incentive practices change.

This distinction is closely related to the decomposition in Lazear (2000), who studies a change from fixed wages to piece rates at Safelite Glass. He shows that the resulting 44% increase in productivity can be decomposed into roughly one–third coming from increased effort by incumbent workers (a pure motivational effect) and about two–thirds from the selection of more productive workers into the firm. In the present context, a similar decomposition is needed: exploration–friendly contracts may raise innovation because they motivate given agents to explore, but also because they attract agents with a comparative advantage in exploration. Empirical designs that track both performance and worker flows around changes in contract form are therefore informative about the relative importance of these two components.

A growing empirical literature examines the motivational effect of exploration-type contracts across diverse institutional settings. Stock options and long-term executive compensation (Laux, 2012; Baranchuk, Kieschnick, and Moussawi, 2014; Chang et al., 2015; Gonzalez-Uribe and Groen-Xu, 2017; Ladika and Sautner, 2020), takeover defenses and board independence (Atanassov, 2013; Chemmanur and Tian, 2018; Bhojraj, Sengupta, and Zhang, 2017; Balsmeier, Fleming, and Manso, 2017), analyst coverage and short-term market pressure (He and Tian, 2013; Guo, Perez-Castrillo, and Toldra-Simats, 2019), shareholder litigation risk (Lin, Liu,

and Manso, 2019; Hassan, Houston, and Karim, 2021; Mbanyele and Muchenje, 2022), bankruptcy codes and creditor rights (Acharya and Subramanian, 2009; Mann, 2018; Fare, Dejardin, and Toulemonde, 2024), unemployment insurance and social insurance (Hombert et al., 2020; Gottlieb, Townsend, and Xu, 2022), venture capital and investor tolerance for failure (Tian and Wang, 2014; Chemmanur, Loutskina, and Tian, 2014; Barrot, 2017), public listing status (Bernstein, 2015; Cong and Howell, 2021; Zhang, Wang, and Gao, 2023), labor laws and job security (Acharya, Baghai, and Subramanian, 2013, 2014; Bradley, Kim, and Tian, 2017; Bena, Ortiz-Molina, and Simintzi, 2021), tenure and academic employment (Azoulay, Graff Zivin, and Manso, 2011; Brogaard, Engelberg, and Van Wesep, 2018; Tripodi et al., 2025), performance evaluation and feedback systems (Gross, 2017; Chu, Tian, and Wang, 2019; Howell, 2021), and financial development or wealth shocks (Hsu, Tian, and Xu, 2014; **?**) all provide settings in which the predictions of the model have been tested.

A particularly informative example is Azoulay, Graff Zivin, and Manso (2011), who study the Howard Hughes Medical Institute (HHMI) investigator program in the life sciences. The HHMI offers long–term, flexible funding that tolerates early failure and emphasizes high–impact, long–horizon research, thus closely resembling the exploration contract developed in the theory. Comparing HHMI investigators to carefully chosen control scientists funded by standard NIH grants, they find that HHMI investigators produce more highly cited and more "radical" work, suggesting that long–term, failure–tolerant incentives stimulate exploration. However, at the same time, according to the predictions of the current paper, the HHMI grant attracts scientists based on exploratory promise, making it natural to ask whether the observed differences are due to motivation, selection, or both. By exploiting pre– and post–appointment trajectories and narrow selection margins, their analysis attempts to focus on motivation but cannot rule out the presence of selection.

Randomised experiments provide perhaps cleaner evidence on the motivational channel, holding selection fixed by design. Ederer and Manso (2013) conduct a laboratory experiment in which subjects manage a stylised "lemonade stand" and must choose between fine–tuning an existing strategy (exploitation) and experimenting with new product–mix and location combinations (exploration). Participants are randomly assigned to a fixed wage, a standard pay–for–performance contract, or an exploration contract that rewards profits only in later periods and is tolerant of early low performance. Consistent with the model, subjects under the exploration contract experiment more and achieve higher long–run performance, whereas standard pay–for–performance contracts bias behaviour towards exploitation. Because contracts are randomly assigned, these differences identify a pure motivational effect of exploration–type incentives. An interesting direction for future work is to extend such experimental designs to study selection by allowing subjects to choose among contracts. Observing which individuals opt into exploration–friendly versus exploitation–friendly schemes would provide direct evidence on whether the most creative, risk–tolerant workers self–select into contracts that display tolerance for early failure and reward for long–term success.

B Ambidexterity and Organizational Design

The model developed in this paper also sheds light on a classic challenge in organizational theory: how to balance the demands of exploration and exploitation. This tension, famously articulated by March (1991), has led to the concept of organizational ambidexterity—the ability of a firm to pursue both incremental improvements and radical innovation simultaneously. While exploitation relies on refining existing capabilities to maximize efficiency, exploration involves risk-taking, experimentation, and a tolerance for failure. Managing both effectively is difficult because the structures, incentives, and cultures that support each tend to con-

flict. O'Reilly and Tushman (2004) define ambidextrous organizations as those that "host multiple contradictory structures, processes, and cultures within the same firm" and show that such firms are more likely to outperform their peers in dynamic environments.

Firms typically adopt one of two broad approaches to achieve ambidexterity. Under "structural ambidexterity," exploration and exploitation are housed in separate organizational units. These units differ not only in function but also in personnel, processes, and culture. Examples include Lockheed Martin's Skunk Works, Alphabet's Google X lab, and IBM's Emerging Business Organization—all dedicated structures designed to shield innovative efforts from the short-term pressures of the core business. In contrast, "contextual ambidexterity" enables individuals within the same unit to engage in both activities, supported by systems and norms that balance alignment and adaptability (Gibson and Birkinshaw, 2004). Google's 20% time and 3M's 15% time exemplify this approach: employees remain embedded in their day-to-day roles but are granted freedom and encouragement to pursue new ideas on a fraction of their time. This enables firms to incorporate innovation into the fabric of daily operations, without formal structural separation.

The appropriate organizational form depends on the nature of the constraint. If the main challenge is motivation—encouraging existing employees to take exploratory action—then contextual ambidexterity can be effective. Granting autonomy, ensuring tolerance for early failure, and rewarding long-term success, as in the policies of 3M and Google, can activate exploration even within operational units (Manso, 2011). However, if the constraint is selection—if individuals differ fundamentally in their aptitude for exploration versus exploitation—then structural separation is more appropriate. This allows firms to attract, retain, and manage explorers under systems tailored to their preferences, while keeping exploiters focused on operational efficiency. The theory of selection for innovation thus provides a microfoundation for understanding when firms

should integrate exploration and exploitation within a unified structure and when they should separate them into distinct units with differentiated incentives and norms.

IX Conclusion

This paper develops a framework to study the selection of exploratory versus exploitative workers. The analysis shows that the optimal incentive scheme designed to attract an explorer-type employee is fundamentally different from the standard high-powered contract that would attract an exploiter. Tolerance (or even reward) for early failure, backloaded rewards for long-term success, strong job security, commitment to a long-term horizon, and timely feedback on performance are all critical features of a contract that screens for innovative, exploratory talent. By contrast, the contract that attracts exploitative workers closely resembles a traditional pay-for-performance plan: it ties pay tightly to immediate output, penalizes early failures (through low current wages and the threat of termination), and places little weight on long-term commitment or feedback.

These findings have implications for organizational design and the classic trade-off between exploration and exploitation. They highlight an ambidexterity choice between integrating exploration into the core business versus separating it into distinct units or spin-offs. The model accordingly draws a sharp distinction between motivation and selection constraints in innovation, each calling for a different organizational approach. When the primary challenge is motivation, integrated schemes can be effective: organization-wide innovation contests and slack-time policies (e.g., Google's "20% time" or 3M's "15% time") spur experimentation within the core business without requiring formal structural separation. Conversely, when the obstacle is selection (i.e. when effective exploration demands individuals with distinct aptitudes) structural sep-

aration is preferable, allowing the firm to attract and cultivate dedicated "explorers" under tailored incentives while keeping exploiters focused on operational efficiency. Thus, the model offers a micro-foundational explanation for when firms should weave exploratory initiatives into their existing operations and when they should isolate them in dedicated units, tightly linking the choice of ambidexterity strategy to the nature of the underlying constraint.

There are several possible extensions of the model. For instance, one could analyze a setting in which workers endogenously select among projects that differ in their level of exploratory potential. This scenario adds a new layer of complexity to the selection problem: the principal must design incentives not only to distinguish explorers from exploiters, but also to guide an explorer's project choice—ensuring that inherently innovative employees undertake sufficiently exploratory ventures rather than gravitating toward safer projects. Another interesting extension would be to study competition in the labor market for exploratory talent. If multiple firms compete to attract explorers, equilibrium contracts and the allocation of innovative workers across organizations might change, raising new questions about efficiency and the design of optimal incentives. I leave these issues for future research.

Finally, the model's implications suggest new avenues for empirical work. Many empirical studies of innovation incentives focus on motivation while treating the workforce as fixed, potentially overlooking important selection effects. Future research could examine how changes in incentive schemes, regulation, or organizational form influence which individuals enter or exit firms, industries, or occupations. In the spirit of Lazear (2000), one could ask whether individuals with more exploratory tendencies disproportionately join (or remain in) organizations that adopt long-horizon, failure-tolerant, feedback-rich incentives, and whether they are more likely to leave when such practices are rolled back. Investigating who self-selects into (or out of) innovative environments would help dis-

entangle motivation versus selection and provide a more complete picture of how incentives affect innovation.

References

Acharya, Viral, Ramin Baghai, and Krishnamurthy Subramanian. 2013. "Labor Laws and Innovation." *Journal of Law and Economics* 56 (4): 997–1037.

Acharya, Viral, Ramin Baghai, and Krishnamurthy Subramanian. 2014. "Wrongful Discharge Laws and Innovation." *Review of Financial Studies* 27 (1): 301–346.

Acharya, Viral, and Krishnamurthy Subramanian. 2009. "Bankruptcy Codes and Innovation." *The Review of Financial Studies* 22 (12): 4949–4988.

Aghion, Philippe, 2002, "Schumpeterian growth theory and the dynamics of income inequality," *Econometrica* 70, 855–882.

Aghion, Philippe, and Jean Tirole. 1994. "The Management of Innovation." *Quarterly Journal of Economics* 109 (4): 1185–1209.

Atanassov, Julian. 2013. "Do Hostile Takeovers Stifle Innovation? Evidence from Antitakeover Legislation and Corporate Patenting." *Journal of Finance* 68 (3): 1097–1131.

Azoulay, Pierre, Joshua Graff Zivin, and Gustavo Manso. 2011. "Incentives and Creativity: Evidence from the Academic Life Sciences." *RAND Journal of Economics* 42 (3): 527–554.

Balsmeier, Benjamin, Lee Fleming, and Gustavo Manso. 2017. "Independent Boards and Innovation." *Journal of Financial Economics* 123(3): 536–557.

Banker, Rajiv D., Seok-Young Lee, Gordon Potter, and Dhinu Srinivasan, 2001, "An empirical analysis of continuing improvements following the implementation of a performance-based compensation plan," *Journal of Accounting and Economics* 30, 315–350.

Baranchuk, Nina, Robert Kieschnick, and Rabih Moussawi. 2014. "Motivating Innovation in Newly Public Firms." *Journal of Financial Economics* 111(3): 578–588.

Barrot, Jean-No"el. 2017. "Investor Horizon and the Life Cycle of Innovative Firms: Evidence from Venture Capital." *Management Science* 63(9):2773–3145.

Bena, Jan, Hernan Ortiz-Molina, and Elena Simintzi. 2021. "Shielding Firm Value: Employment Protection and Intra-firm Reallocation." *Journal of Financial Economics* 139 (1): 161–189.

Bergemann, Dirk, and Ulrich Hege, 2005, "The financing of innovation: learning and stopping," *RAND Journal of Economics* 36, 719–752.

Bergemann, Dirk, and Juuso Valimaki, 2006, "Bandit problems," in *The New Palgrave Dictionary of Economics*, 2nd ed., S. Durlauf and L. Blume (eds.) (Palgrave Macmillan).

Bernstein, Shai. 2015. "Does Going Public Affect Innovation?" *Journal of Finance* 70 (4): 1365–1403.

Bernstein, Shai, Timothy McQuade, and Richard Townsend. 2021. "Do Household Wealth Shocks Affect High-Growth Entrepreneurship?" *The Review of Financial Studies* 34 (3): 1189–1238.

Berry, Donald A., and Bert Fristedt, 1985, *Bandit Problems: Sequential Allocation of Experiments* (London: Chapman and Hall).

Bhattacharya, Sudipto, Kalyan Chatterjee, and Larry Samuelson, 1986, "Sequential research and the adoption of innovations," *Oxford Economic Papers* 38 (Suppl.), 219–243.

Bhojraj, Sanjeev, Partha Sengupta, and Suning Zhang. 2017. "Takeover Defenses: Entrenchment and Efficiency." *Journal of Accounting and Economics* 63: 142–160.

Bolton, Patrick, and Christopher Harris, 1999, "Strategic experimentation," *Econometrica* 67, 349–374.

Bradley, Daniel, Incheol Kim, and Xuan Tian. 2017. "Do Unions Affect Innovation?" *Management Science* 63(7): 2251–2271.

Brogaard, Jonathan, Joseph Engelberg, and Edward Van Wesep. 2018. "Do Economists Swing for the Fences after Tenure?" *Journal of Economic Perspectives* 32: 179—94.

Cadsby, C. Bram, Fei Song, and Francis Tapon, 2007, "Sorting and incentive effects of pay-for-performance: an experimental investigation," *Academy of Management Journal* 50, 387–405.

Castro-Pires, Henrique, and George Georgiadis. 2025. "Incentives and Selection." Working Paper, Northwestern University.

Chang, Xin, King-Yeung Fu, Angie Low, and Wenrui Zhang. 2015. "Non-Executive Stock Options and Corporate Innovation." *Journal of Financial Economics* 115 (1): 168–188.

Chemmanur, Thomas, Elena Loutskina, and Xuan Tian. 2014. "Corporate Venture Capital, Value Creation, and Innovation." *Review of Financial Studies* 27 (8): 2434–2473.

Chemmanur, Thomas, and Xuan Tian. 2018. "Do Anti-Takeover Provisions Spur Corporate Innovation?" *Management Science* 64 (11): 5378–5400.

Chen, Zhibin, Yibin Xu, Zongtao Tian, and Xu Meng. 2022. "The Impact of Social Insurance Law on Corporate Innovation: Evidence from a Quasi-Natural Experiment." *Economic Modelling* 111: 105830.

Chief Executive (2025). CEO Briefing: Challenges in Attracting Creative Professionals. https://chiefexecutive.net.

Chu, Jun, Xuan Tian, and Wenyu Wang. 2019. "Corporate Innovation Along the Supply Chain" *Management Science* 65 (6): 2445–2945.

Cong, Lin William, and Sabrina T. Howell. 2021. "Policy Uncertainty and Innovation: Evidence from Initial Public Offerings Interventions in China." *Management Science* 67: 6629–7289.

Deloitte (2025). 2025 Global Human Capital Trends: The Imagination Imperative. https://www2.deloitte.com/us/en/insights/focus/human-capital-trends.html.

Dohmen, Thomas, and Armin Falk, 2011, "Performance pay and multi-dimensional sorting: productivity, preferences, and gender," *American Economic Review* 101, 556–590.

Ederer, Florian, and Gustavo Manso. 2013. "Is Pay for Performance Detrimental to Innovation?" *Management Science* 59 (7): 1496–1513.

Edmans, Alex, Vivian W. Fang, and Katharina A. Lewellen. 2017. "Equity Vesting and Investment." *Review of Financial Studies* 30(7): 2229–2271.

Eriksson, Tor, and Marie-Claire Villeval, 2008, "Performance-pay, sorting and social motivation," *Journal of Economic Behavior & Organization* 68, 412–421.

Farè, Marie, Marcus Dejardin, and Eric Toulemonde. 2024. "Creditor Rights and Innovation: International Evidence." *Journal of Economic Behavior & Organization* 216: 356–375.

Feltham, Gerald A., and Martin G. H. Wu. 2001. "Incentive Efficiency of Stock versus Options." *Review of Accounting Studies* 6 (1): 7–28.

Francis, Bill B., Iftekhar Hasan, Zenu Sharma, and Maya Waisman. 2019. "Motivating High-Impact Innovation: Evidence from Managerial Compensation Contracts." *Financial Markets, Institutions & Instruments* 28(3): 291–318.

Francis, Bill B., Incheol Kim, Bin Wang, and Zhengyi Zhang. 2018. "Labor Law and Innovation Revisited." *Journal of Banking & Finance* 94: 1–15.

Gibson, Cristina B., and Julian Birkinshaw. 2004. "The Antecedents, Consequences, and Mediating Role of Organizational Ambidexterity." *Academy of Management Journal* 47(2): 209–226.

Gomes, Renato, Daniel Gottlieb, and Lucas Maestri, 2016, "Experimentation and project selection: screening and learning," *Games and Economic Behavior* 96, 145–169.

Gonzalez-Uribe, Juanita, and Michael Groen-Xu. 2017. "CEO Contract Horizon and Innovation." *Working Paper*.

Gottlieb, Joshua D., Richard R. Townsend, and Ting Xu. 2022. "Does Career Risk Deter Potential Entrepreneurs?" *Review of Financial Studies* 35(9): 3973–4015.

Gross, Daniel. 2017. "Performance Feedback in Competitive Product Development." *Rand Journal of Economics* 48 (2): 438–466.

Guo, Bing, David Perez-Castrillo, and Antoni Toldra-Simats. 2019. "Firm's Innovation Strategy Under the Shadow of Analyst Coverage." *Journal of Financial Economics* 131 (1): 456–483.

Halac, Marina, Navin Kartik, and Qingmin Liu, 2016, "Optimal contracts for experimentation," *Review of Economic Studies* 83, 1040–1091.

Halac, Marina, Navin Kartik, and Qingmin Liu, 2017, "Contests for experimentation," *Journal of Political Economy* 125, 1523–1569.

Harris, Milton, and Artur Raviv, 1978, "Some results on incentive contracts with applications to education and employment, health insurance, and law enforcement," *American Economic Review* 68, 20–30.

Hassan, M. Kabir, Reza Houston, and M. Sydul Karim. 2021. "Courting Innovation: The Effects of Litigation Risk on Corporate Innovation." *Journal of Corporate Finance* 71.

He, Jie, and Xuan Tian. 2013. "The Dark Side of Analyst Coverage: The Case of Innovation." *Journal of Financial Economics* 109: 856–878.

Hellmann, Thomas, 2007, "When do employees become entrepreneurs?," *Management Science* 53, 919–933.

Hellmann, Thomas, and Veikko Thiele, 2011, "Incentives and innovation: a multi-tasking approach," *American Economic Journal: Microeconomics* 3, 78–128.

Holmstrom, Bengt, 1989, "Agency costs and innovation," *Journal of Economic Behavior & Organization* 12, 305–327.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, & Organization* 7 (Special Issue): 24–52.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar. 2020. "Can Unemployment Insurance Spur Entrepreneurial Activity?" *Journal of Finance* 75 (3): 1247–1285.

Howell, Sabrina. 2021. "Learning from Feedback: Evidence from New Ventures." *Review of Finance* 25: 595–627.

Hsu, Po-Hsuan, Xuan Tian, and Yan Xu. 2014. "Financial Development and Innovation: Cross-Country Evidence." *Journal of Financial Economics* 112 (1): 116–135.

IBM Institute for Business Value (2025). CEO Study 2025: Scaling Generative AI and the Creative Mindshift. https://www.ibm.com/thought-leadership/institute-business-value/en-us/report-ceo-2025.

Jensen, Richard A., 1982, "Adoption and diffusion of an innovation of uncertain profitability," *Journal of Economic Theory* 27, 182–193.

Jovanovic, Boyan, and Yaw Nyarko, 1996, "Learning by doing and the choice of technology," *Econometrica* 64, 1299–1310.

Jovanovic, Boyan, and Rafael Rob, 1990, "Long waves and short waves: growth through intensive and extensive search," *Econometrica* 58, 1391–1409.

KPMG (2024). 2024 U.S. CEO Outlook. https://home.kpmg/xx/en/home/insights/2024/09/2024-ceo-outlook.html.

Ladika, Tomislav, and Zacharias Sautner. 2020. "Executive Compensation and Corporate Innovation." *Journal of Financial Economics* 135 (2): 521–545.

Lambert, Richard A. 1986. "Executive Effort and Selection of Risky Projects." *RAND Journal of Economics* 17 (1): 77–88.

Laux, Volker. 2012. "Stock Option Vesting Conditions, CEO Turnover, and Myopic Investment." *Journal of Financial Economics* 106 (3): 513–526.

Lazear, Edward P. 1981. "Agency, Earnings Profiles, Productivity, and Hours Restrictions." *American Economic Review* 71, 606–620.

Lazear, Edward P. 1986. "Salaries and piece rates," *Journal of Business* 59, 405–431.

Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review* 90 (5): 1346–1361.

Leaver, Clare, Owen Ozier, Pieter Serneels, and Andrew Zeitlin. 2021. "Recruitment, Effort, and Retention Effects of Performance Contracts

for Civil Servants: Experimental Evidence from Rwandan Primary Schools." *American Economic Review* 111 (7): 2213–2246.

Lin, Chen, Siqi Liu, and Gustavo Manso. 2019. "Shareholder Litigation and Corporate Innovation." *Management Science* 65 (5): 2043–2062.

Mann, William. 2018. "Creditor Rights and Innovation: Evidence from Patent Collateral." *Journal of Financial Economics* 130 (1): 25–47.

Manso, Gustavo. 2011. "Motivating innovation," *Journal of Finance* 66, 1823–1860.

March, James G. 1991. "Exploration and exploitation in organizational learning," *Organization Science* 2, 71–87.

Mbanyele, Wilson, and Luke Muchenje. 2022. "Legal Risk and Innovation." *Journal of Corporate Finance* 72: 102139.

McKinsey & Company (2025). Global Survey: Hiring and Retaining Innovation Talent. https://www.mckinsey.com.

Moscarini, Giuseppe, and Lones Smith, 2001, "The optimal level of experimentation," *Econometrica* 69, 1629–1644.

Nanda, Ramana, and Matthew Rhodes-Kropf. 2017. "Financing Risk and Innovation." *Management Science* 63(4): 901–918.

Oliver Wyman Forum (2025). 2025 CEO Survey: The Great Reinvention. https://www.oliverwymanforum.com/leadership-ceo-survey-2025.html.

O'Reilly, Charles A., III, and Michael L. Tushman. 2004. "The Ambidextrous Organization." *Harvard Business Review* 82 (4): 74–81.

O'Reilly, Charles A., and Michael L. Tushman, 2013, "Organizational ambidexterity: past, present, and future," *Academy of Management Perspectives* 27, 324–338.

Roberts, Kevin, and Martin L. Weitzman, 1981, "Funding criteria for research, development, and exploration projects," *Econometrica* 49, 1261–1288.

Salop, Joanne, and Steven C. Salop, 1976, "Self-Selection and Turnover in the Labor Market," *Quarterly Journal of Economics* 90, 619–627.

Sapra, Haresh, Ajay Subramanian, and Krishnamurthy V. Subramanian. 2014. "Corporate Governance and Innovation: Theory and Evidence." *Journal of Financial and Quantitative Analysis* 49 (4): 957–1003.

Sauermann, Henry, and Wesley M. Cohen. 2010. "What Makes Them Tick? Employee Motives and Firm Innovation." *Management Science* 56 (12): 2134–2153.

Tian, Xuan, and Tracy Yue Wang. 2014. "Tolerance for Failure and Corporate Innovation." *Review of Financial Studies* 27 (1): 211–255.

Tripodi, Giorgio, Xiang Zheng, Yifan Qian, Dakota Murray, Benjamin F. Jones, Chaoqun Ni, and Dashun Wang. 2025. "Tenure, Security, and Scientific Innovation." Working Paper.

Tushman, Michael L., and Charles A. O'Reilly III, 1996, "Ambidextrous organizations: managing evolutionary and revolutionary change," *California Management Review* 38 (4), 8–30.

Weitzman, Martin L., 1979, "Optimal search for the best alternative," *Econometrica* 47, 641–654.

Zhang, Wei, Yue Wang, and Rui Gao. 2023. "Public Listing and Innovation: Evidence from Deregulation." *Journal of Corporate Finance* 80: 102277.

A Online Appendix

Proof of Lemma 1. We establish the equivalence between the single–principal and two–principal formulations by showing each direction separately.

(i) Single–Principal ⇒ Two–Principal. Let $(\tilde{w}^{\mathrm{E}}, \tilde{w}^{\mathrm{X}})$ be an optimal separating menu in the single–principal problem (**??**), with associated action plans $(a_{\mathrm{E}}, a_{\mathrm{X}})$. Define:

$$\overline{U}_{\mathrm{E}}^{\mathrm{own}} := U_{\mathrm{E}}(\tilde{w}^{\mathrm{E}}, a_{\mathrm{E}}), \qquad \overline{U}_{\mathrm{E}}^{\mathrm{mis}} := \sup_{a' \in \mathcal{A}(\mathrm{E})} U_{\mathrm{E}}(\tilde{w}^{\mathrm{X}}, a'),$$
$$\overline{U}_{\mathrm{X}}^{\mathrm{own}} := U_{\mathrm{X}}(\tilde{w}^{\mathrm{X}}, a_{\mathrm{X}}), \qquad \overline{U}_{\mathrm{X}}^{\mathrm{mis}} := \sup_{a' \in \mathcal{A}(\mathrm{X})} U_{\mathrm{X}}(\tilde{w}^{\mathrm{E}}, a').$$

These four values describe each type's utility on their own and the other contract.

We verify that each contract satisfies the constraints of its corresponding auxiliary problem:

- *Incentive compatibility:* In the original problem, each type prefers its assigned plan over all alternatives in their feasible action set. This ensures that $\tilde{w}^{\mathrm{E}}$ and $\tilde{w}^{\mathrm{X}}$ satisfy the IC constraints of the two–principal problems.
- *Ride–along constraints:* By construction, the explorer's utility on $\tilde{w}^{\mathrm{X}}$ is at most $\overline{U}_{\mathrm{E}}^{\mathrm{own}}$, and similarly the exploiter's utility on $\tilde{w}^{\mathrm{E}}$ is at most $\overline{U}_{\mathrm{X}}^{\mathrm{own}}$.
- *Participation:* From the single–principal solution, each type obtains at least their participation utility $\overline{U}_{\theta}^{\mathrm{com}}$, and also weakly prefers their own contract to mimicking the other type. Thus,

$$U_{\theta}(\tilde{w}^{\theta}, a_{\theta}) \geq \max\{\overline{U}_{\theta}^{\mathrm{com}}, \overline{U}_{\theta}^{\mathrm{mis}}\}, \quad \theta \in \{\mathrm{E}, \mathrm{X}\}.$$

Finally, by definition, the values $(\overline{U}_{\mathrm{E}}^{\mathrm{own}}, \overline{U}_{\mathrm{E}}^{\mathrm{mis}}, \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \overline{U}_{\mathrm{X}}^{\mathrm{mis}})$ coincide with the actual utilities induced by the menu, so consistency holds. Moreover, if $\tilde{w}^{\mathrm{E}}$ or $\tilde{w}^{\mathrm{X}}$ were not cost-minimising given the relevant constraints, then the original menu would not be optimal—a contradiction.

(ii) Two–Principal ⇒ Single–Principal. Suppose we are given a pair $(\tilde{w}^{\mathrm{E}}, \tilde{w}^{\mathrm{X}})$ that solves the two auxiliary problems for target plans $(a_{\mathrm{E}}, a_{\mathrm{X}})$ at utility caps $(\overline{U}_{\mathrm{E}}^{\mathrm{own}}, \overline{U}_{\mathrm{E}}^{\mathrm{mis}}, \overline{U}_{\mathrm{X}}^{\mathrm{own}}, \overline{U}_{\mathrm{X}}^{\mathrm{mis}})$ and satisfies the consistency conditions.

We claim that this menu is feasible for the single–principal problem:

- *Within-type IC:* By the two–principal solution, each type weakly prefers their designated plan to all alternatives within their own action set.
- *No-mimicking:* The ride–along and participation constraints ensure that

$$U_{\mathrm{E}}(\tilde{w}^{\mathrm{E}}, a_{\mathrm{E}}) \geq \sup_{a' \in \mathcal{A}(\mathrm{E})} U_{\mathrm{E}}(\tilde{w}^{\mathrm{X}}, a'), \qquad U_{\mathrm{X}}(\tilde{w}^{\mathrm{X}}, a_{\mathrm{X}}) \geq \sup_{a' \in \mathcal{A}(\mathrm{X})} U_{\mathrm{X}}(\tilde{w}^{\mathrm{E}}, a').$$

- *Participation:* Follows from the participation constraints in the two–principal formulation.

Lastly, optimality follows by contradiction: if there existed a strictly better feasible menu, then at least one of the two subproblems could be solved at strictly lower cost, violating the cost-minimisation in the auxiliary formulation. Hence the menu is optimal for the single–principal problem. □

Proof of Lemma 2. Fix a type $\theta \in \{\mathrm{E}, \mathrm{X}\}$ and a target plan $\pi \in \mathcal{A}(\theta)$. Assume the decoupled subproblem for type θ is feasible at the given utility caps and participation threshold. We show that the contract produced

by the three–step procedure—*baseline*, *poison*, and *deliver*—is an optimal solution to the cost–minimisation problem.

Step 1: Baseline. In the first step, the principal solves the constrained minimisation problem

$$\min_{\tilde{w}\geq 0} W(\tilde{w},\pi) \quad \text{s.t. } U_\theta(\tilde{w},\pi) \geq U_\theta(\tilde{w},\pi') \text{ for all } \pi' \in \mathcal{A}(\theta),$$

thereby obtaining the least–cost contract that implements π for type θ alone. Likelihood–ratio analysis ensures that the optimal contract concentrates wages on outcomes that are most informative of plan π, i.e., those with high likelihood ratios $\Lambda_{\pi,\pi'}(h)$ relative to within–type deviations π'. Any alternative wage schedule either violates some IC constraint or increases total expected payments. Thus, the baseline contract is cost–minimising among all contracts that implement π for type θ.

Step 2: Poison. If the baseline contract violates the ride–along constraint for the other type θ', the principal poisons the contract by reallocating payments toward outcomes h with high $\Lambda_{\pi,\pi'}(h)$, where π' is the best–response plan of θ' when misreporting. This reallocation preserves incentive compatibility for type θ—since the rewards remain on outcomes frequent under π—but reduces the misreporter's utility by shifting weight to outcomes unlikely under π'. If feasible, this process drives the misreporter's utility down to $\overline{U}_{\theta'}^{\text{own}}$ at minimal cost. The resulting contract satisfies both IC and the cross–type ride–along constraint with weak or exact equality.

Step 3: Deliver. If the poisoned contract from Step 2 gives type θ less than their required utility $\overline{U}_\theta$, the principal delivers surplus by increasing payments at outcomes h with high $\Lambda_{\pi,\pi'}(h)$. These are events that benefit type θ more than the misreporter, ensuring minimal leakage of surplus.

By adjusting along IC–preserving directions, the principal reaches $U_\theta = \overline{U}_\theta$ while maintaining all other constraints. If the ride–along constraint was binding, any necessary additional poisoning is combined with surplus delivery to maintain feasibility.

Each step of the procedure strictly minimises cost conditional on the prior constraints as well as leakage to the misreporting type. No feasible contract can achieve a lower expected wage bill while satisfying IC, RAC, and PC. Therefore, the contract obtained through the three–step construction is optimal for the decoupled problem. □

Proof of Proposition 1. Fix the target action plan $\langle {}_1{}^1_1 \rangle$. Under limited liability, write each wage as $w_x = w + \tilde{w}_x$ with $\tilde{w}_x \geq 0$ for each outcome $x \in \{S, F, SS, SF, FS, FF\}$. Incentive compatibility depends only on these bonus components $\tilde{w}_x$. Let $\Lambda_{a,a'}(x) := \Pr(x \mid a)/\Pr(x \mid a')$ denote the likelihood ratio of outcome x under plan a versus plan a'. With $p_0 = 0$, any incentive comparison against a shirking margin yields *polar* LRs on that branch. For example,

$$\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_0{}^j_k \rangle}(S) = \frac{p_1}{0} = +\infty, \qquad \Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_0{}^j_k \rangle}(F) = \frac{1-p_1}{1} = 1 - p_1.$$

Similarly, for second-period deviations we have:

$$\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_i{}^0_k \rangle}(SS) = \frac{p_1}{0} = +\infty, \quad \Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_i{}^0_k \rangle}(SF) = \frac{1-p_1}{1} = 1 - p_1,$$

$$\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_i{}^j_0 \rangle}(FS) = \frac{p_1}{0} = +\infty, \quad \Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_i{}^j_0 \rangle}(FF) = \frac{1-p_1}{1} = 1 - p_1.$$

Moreover, between the exploiter's plan $\langle {}_1{}^1_1 \rangle$ and the explorer's plan $\langle {}_2{}^2_2 \rangle$ (who exerts on all margins), we have:

$$\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(S) \;=\; \frac{p_1}{E[p_2]}, \qquad \Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(SS) \;=\; \frac{p_1^2}{E[p_2] \cdot E[p_2 \mid S, 2]},$$

$$\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(FS) \;=\; \frac{p_1(1-p_1)}{E[p_2]\big(1 - E[p_2]\big)}.$$

Thus, relative to any deviation that involves shirking on a given margin, the success node on that margin is infinitely more likely under the intended plan (most diagnostic, likelihood ratio is $+\infty$), while the failure node is least diagnostic (finite likelihood ratio). By the likelihood–ratio principle, cost minimization loads the entire incentive differential on the most diagnostic outcome for each margin. Thus, in the optimal baseline contract we must have

$$(p_1{-}0)\,(\tilde{w}_{SS}{-}\tilde{w}_{SF}) = c_1, \qquad (p_1{-}0)\,(\tilde{w}_{FS}{-}\tilde{w}_{FF}) = c_1, \qquad (p_1{-}0)\,\tilde{w}_S = c_1.$$

Solving these gives $\tilde{w}_S = \tilde{w}_{SS} = \tilde{w}_{FS} = \frac{c_1}{p_1}$ and $\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0$. Equivalently, the principal pays a constant bonus c_1/p_1 for each success and nothing after failures—exactly the one-period moral-hazard solution replicated over the two-period tree.

Notably, while an alternative wage profile could achieve the same implementation cost by shifting part of the first-period success bonus to the second-period success-after-success node (i.e. reducing $\tilde{w}_S$ and increasing $\tilde{w}_{SS}$ accordingly), our tie-breaking rule selects the above baseline with $\tilde{w}_S = \tilde{w}_{SS}$. Such a shift would only raise the explorer's rent without lowering cost, because success in period 1 is more diagnostic than success after success. (indeed $\Lambda_{\langle 1\,1\rangle,\langle 2\,2\rangle}(S) > \Lambda_{\langle 1\,1\rangle,\langle 2\,2\rangle}(SS)$). The principal optimally refrains from this distortion.

The exploiter's baseline expected utility is $2\underline{w}$. A misreporting explorer's payoff on the exploiter's contract is (weakly) increasing in each success-contingent bonus $(\tilde{w}_S, \tilde{w}_{SS}, \tilde{w}_{FS})$, since any increase in these bonuses raises the misreporter's value. The only feasible way to reduce any of them while still implementing $\langle {}_1{}_1^1\rangle$ is to lower $\tilde{w}_S$ and simultaneously raise $\tilde{w}_{SS}$ along the first-period IC indifference arc. However, as discussed above, this adjustment keeps the misreporting explorer indifferent if he shirks and makes him strictly better off if he exerts (shifting weight from the more diagnostic S outcome to the less diagnostic SS outcome increases

the explorer's expected payoff). Therefore, if the ride–along constraint binds, it must be enforced on the explorer's own menu rather than by tweaking the exploiter's wages.

Now suppose the exploiter's participation constraint binds, and let $\Delta := \overline{U}_X - 2\underline{w} > 0$ denote the additional surplus that must be delivered.

Step A (tighten any shirking margin first). If the misreporting explorer shirks on some margin under the current contract, the principal uses the most diagnostic success node on that branch to eliminate shirking at least cost. Proceed as follows:

(i) Shirking in period 1. Raise $\tilde{w}_S$ until the period–1 exploration IC for the explorer is tight. This uses the first–period success node, which has infinite likelihood–ratio advantage against shirking when $p_0 = 0$, and therefore eliminates period–1 shirking with zero leakage.

(ii) Shirking in period 2 after S. Raise $\tilde{w}_{SS}$ until the explorer is just willing to exert in period 2 following a first–period success. Increasing $\tilde{w}_{SS}$ strengthens both the explorer's and exploiter's period–1 incentives and therefore does not endanger any earlier IC.

(iii) Shirking in period 2 after F. Increase $\tilde{w}_{FS}$ while holding $\tilde{w}_S$ fixed *as long as* the exploiter's period–1 IC remains slack. This raises the explorer's continuation value after failure just enough to make exertion optimal in the second period, without affecting the exploiter's willingness to exert in period 1. If further increases in $\tilde{w}_{FS}$ are needed after the exploiter's period–1 IC becomes binding, continue by adjusting $(\tilde{w}_{FS}, \tilde{w}_S)$ *jointly* along the exploiter's period–1 IC indifference arc, so that exerting in period 1 remains optimal for the exploiter. This preserves all ICs while using the success–after–failure node—which has infinite likelihood–ratio against shirking when $p_0 = 0$—to eliminate shirking with minimal leakage.

After completing the appropriate substeps above, the misreporting explorer exerts on all margins, and all shirking–related ICs are tight at minimum cost.

Step B (once the explorer exerts on all margins). After Step A, all relevant deviations for the misreporter involve exertion on every margin; the polar $\Lambda_{=,+}(\infty)$ advantages disappear. The least-cost way to deliver any remaining surplus Δ must keep all of the exploiter's IC constraints intact while leaking as little as possible to the misreporter. Among all such IC-preserving directions, increasing $\tilde{w}_S$ *alone* is optimal. When comparing $\langle {}_1{}^1_1 \rangle$ to a misreporter who also exerts on every margin (e.g. $\langle {}_2{}^2_2 \rangle$), the first-period success outcome S has the highest diagnosticity available under the IC manifold. Indeed, $\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(S)$ exceeds both $\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(SS)$ and $\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(FS)$. Thus, any remaining Δ is optimally delivered by raising $\tilde{w}_S$ only.

Combining the baseline contract construction, the impossibility of relaxing the ride–along constraint via the exploiter's wage schedule, and the two-step surplus–delivery rule completes the proof. □

Proof of Proposition 2. Fix the target plan $a_{\mathrm{E}} = \langle {}_2{}^2_2 \rangle$. Under limited liability write $w_x = \underline{w} + \tilde{w}_x$ with $\tilde{w}_x \geq 0$ for $x \in \{F, S, SF, SS, FS, FF\}$, so incentives depend only on $\tilde{w}$.

Since rewards for failure induce shirking, failure–branch wages do not help and can be set at the floor: $w_F = w_{SF} = w_{FF} = \underline{w}$. The second–period ICs decouple cleanly:

(i) *After F:* to induce exploration in period 2,

$$E[p_2]\ \tilde{w}_{FS} \ \geq \ c_2 \quad \Rightarrow \quad \tilde{w}^0_{FS} = \frac{c_2}{E[p_2]}.$$

(ii) *After S:* to induce exploration in period 2,

$$E[p_2 \mid S, 2]\ \tilde{w}_{SS} \ \geq\ c_2 \quad \Rightarrow \quad \tilde{w}_{SS} \ \geq\ \frac{c_2}{E[p_2 \mid S, 2]}.$$

Period 1 exploration requires that choosing 2 in period 1 beats the nearest deviations. Relative to $\langle {}_0{}^2_2 \rangle$ (shirk in period 1, then explore), the *only* incremental benefit from choosing 2 in period 1 is the option value of reaching S and then SS:

$$E[p_2]\ E[p_2 \mid S, 2]\ \tilde{w}_{SS} \ \geq\ c_2 \quad \Rightarrow \quad \tilde{w}_{SS} \ \geq\ \frac{c_2}{E[p_2]\ E[p_2 \mid S, 2]}.$$

Relative to $\langle {}_0{}^0_2 \rangle$ (shirk both periods), the expected bonus under $\langle {}_2{}^2_2 \rangle$ is $E[p_2]E[p_2 \mid S, 2]\tilde{w}_{SS}$ and must cover costs $(1 + E[p_2 \mid S, 2])c_2$. Therefore,

$$\tilde{w}_{SS} \ \geq\ \frac{(1 + E[p_2])\, c_2}{E[p_2]\ E[p_2 \mid S, 2]}.$$

The two lower bounds coincide, yielding the minimal bonuses

$$\tilde{w}^0_{FS} = \frac{c_2}{E[p_2]}, \qquad \tilde{w}^0_{SS} = \frac{(1 + E[p_2])\, c_2}{E[p_2]\ E[p_2 \mid S, 2]}.$$

Finally, w_S does not tighten any IC at these primitives and can be kept at the floor: $w_S = \underline{w}$. This gives the stated baseline and

$$U_{\mathrm{E}}^{\mathrm{base}} = 2\underline{w}.$$

A misreporting exploiter's payoff on the explorer contract is (weakly) increasing in w_{SS} and w_{FS}. Both stand at their IC–minimal levels for $\langle {}_2{}^2_2 \rangle$; any attempt to "poison" by lowering w_{SS} and compensating with w_S or w_{FS} either violates the explorer's ICs or strictly raises the exploiter's value by the likelihood–ratio rule. Thus, if the ride–along cap binds, it must be enforced on the exploiter's menu, not here.

If participation binds, set $\Delta = \overline{U}_{\mathrm{E}}^{\mathrm{mis}} - U_{\mathrm{E}}^{\mathrm{base}} > 0$. By the LR principle comparing $\langle {}_1{}^1_1 \rangle$ to $\langle {}_2{}^2_2 \rangle$, the two informative directions that preserve the explorer's IC manifold are SS and (if needed) F. Define

$$\Lambda_{SS} = \Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_0 \rangle}(SS) = \frac{p_1^2}{E[p_2]\ E[p_2 \mid S, 2]}, \qquad \Lambda_F = \Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_0 \rangle}(F) = \frac{1 - p_1}{1 - E[p_2]}.$$

Step A (if the misreporting exploiter shirks on some margin). Tighten that margin first using the instrument with the highest LR against shirking: raise $\tilde{w}_S$ until the exploiter exerts in the first period; if shirking remains after S or F, then raise $\tilde{w}_{SS}$ and $\tilde{w}_{FS}$ until exertion after S and F is also optimal. This piecewise adjustment minimizes leakage while preserving $\langle {}_2{}^2_2 \rangle$.

Step B (if the misreporting exploiter already exerts $\langle 1, 1, 1 \rangle$). Keep $\tilde{w}_{FS} = \tilde{w}_{FS}^0$ and deliver any remaining Δ along the most diagnostic IC–preserving direction: if $\Lambda_{SS} \geq \Lambda_F$, raise $\tilde{w}_{SS}$ alone; otherwise increase $(\tilde{w}_F, \tilde{w}_{SS})$ along the exploiter's IC–indifference ray so that the second–period ICs remain tight (i.e., offset a rise in w_F with the exact increase in w_{SS} that keeps $\langle {}_1{}^1_1 \rangle$ incentive compatible). This minimizes leakage to the exploiter.

Combining the baseline construction, the impossibility of relaxing the ride–along on the explorer contract, and the surplus–delivery rule completes the proof. □

Proof of Proposition 3. Fix the target action plan $\langle {}_1{}^1_1 \rangle$ and write $w_x = \underline{w} + \tilde{w}_x$ with $\tilde{w}_x \geq 0$. Under no commitment, the principal offers a sequence of one–period contracts and may renegotiate at the beginning of period 2 after observing the period–1 outcome.

By Proposition 1 (with $p_0 = 0$), the least–cost long–term contract that implements $\langle {}_1{}^1_1 \rangle$ pays a constant bonus $\alpha_1 = c_1/p_1$ for every success: $\tilde{w}_S = \tilde{w}_{SS} = \tilde{w}_{FS} = \alpha_1$ and $\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0$. This scheme is a

repetition of the one–period moral–hazard solution and therefore admits a short–term implementation: in period $t \in \{1, 2\}$ the principal offers a one–period contract that pays $\underline{w} + \alpha_1$ upon success and $\underline{w}$ upon failure. Each period's effort decision depends only on the contemporaneous success bonus; the resulting sequence of short–term contracts induces effort in both periods and replicates the baseline at the same expected cost.

If the participation constraint requires additional surplus, the only instruments available are $\tilde{w}_S$ and $\tilde{w}_F$ since any promise to provide extra surplus in the second period is not renegotiation-proof. If the misreporting explorer shirks or exerts in the first period, raising the bonus on S is the least leaky option. □

Proof of Proposition 4. Fix the target plan $a_{\mathrm{E}} = \langle {}_2{}^2_2 \rangle$ and write $w_x = \underline{w} + \tilde{w}_x$ with $\tilde{w}_x \geq 0$ for $x \in \{F, S, SF, SS, FS, FF\}$. Without commitment, the principal offers a sequence of one–period contracts and can renegotiate at the beginning of period 2 after observing the period–1 history $h \in \{S, F\}$.

At the period–2 renegotiation, the period–1 action is sunk; the principal minimizes the *current* expected wage bill subject to inducing exploration in period 2, given h. By the one–period logic, the minimal continuation contracts are:

$$\text{after } F: \quad E[p_2]\,\tilde{w}_{FS} = c_2 \;\Rightarrow\; \tilde{w}^0_{FS} = \frac{c_2}{E[p_2]}, \qquad \text{after } S: \quad E[p_2 \mid S, 2]\,\tilde{w}_{SS} = c_2 \;\Rightarrow\; \tilde{w}^0_{SS} = \frac{c_2}{E[p_2 \mid S, 2]},$$

with all failure–pay outcomes at the floor ($\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0$). Any back–loaded surplus embedded in w_{FS} or w_{SS} above these minima is strictly dominated at renegotiation by a cheaper contract that delivers the same (or higher) continuation utility to the agent at lower expected cost; hence such promises are not renegotiation–proof. Therefore, in equilibrium, the period–2 bonuses must equal the minimal values $\tilde{w}^0_{FS}$ and $\tilde{w}^0_{SS}$.

Given the renegotiation–proof continuation, the explorer's *expected* continuation utility after S or F is zero (the period–2 bonus just covers c_2). Hence the period–1 exploration IC depends only on period–1 wages:

$$\underbrace{p_2(w_S - \underline{w}) + (1-p_2)(w_F - \underline{w}) - c_2}_{\text{choose 2 in period 1}} \geq \underbrace{(w_F - \underline{w})}_{\text{choose 0 in period 1}} \quad \Longleftrightarrow \quad p_2\,(w_S - w_F) \geq c_2.$$

Under limited liability the cost–minimizing baseline sets $w_F = \underline{w}$ and chooses $\tilde{w}_S$ at equality:

$$\tilde{w}_S^0 = \frac{c_2}{E[p_2]}.$$

Collecting terms yields the stated baseline schedule:

$$\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0, \qquad \tilde{w}_{FS}^0 = \frac{c_2}{E[p_2]}, \qquad \tilde{w}_{SS}^0 = \frac{c_2}{E[p_2 \mid S, 2]}, \qquad \tilde{w}_S^0 = \frac{c_2}{E[p_2]},$$

which is implementable via a sequence of one–period contracts: in period 1 pay $\underline{w} + \tilde{w}_S^0$ upon success and $\underline{w}$ upon failure; in period 2, conditional on $h \in \{S, F\}$, pay the corresponding minimal continuation bonus derived above.

Suppose the participation constraint binds and set $\Delta = \overline{U}_{\mathrm{E}}^{\mathrm{mis}} - U_E^0 > 0$.

(a) Misreporting exploiter shirks at the current schedule. To tighten the exploiter's period–1 IC at least cost, increase $\tilde{w}_S$ alone until he is indifferent between shirking and exerting. This raises $(w_S - w_F)$ and, by the likelihood–ratio rule with $p_0 = 0$, leaks weakly less to a shirker than rewarding failure. Once the exploiter exerts on the first–period margin, *keep* the explorer's period–1 IC at equality by moving (w_S, w_F) in parallel:

$$p_2\,(w_S - w_F) = c_2 \quad \Rightarrow \quad dw_S = dw_F.$$

This parallel move delivers additional surplus while preserving the explorer's period–1 IC and the renegotiation–proof period–2 implementation.

(b) Misreporting exploiter already exerts. From the outset move (w_S, w_F) in parallel, $dw_S = dw_F$, so that $p_2(w_S - w_F) = c_2$ remains tight. Any attempt to deliver surplus via w_{SS} or w_{FS} would be undone by period–2 renegotiation, while any asymmetric change to (w_S, w_F) that relaxes the explorer's IC would require compensating adjustments in period 2 and is therefore dominated by the parallel move.

The baseline schedule is implementable by short–term contracts and is unique among renegotiation–proof implementations. Any additional surplus must be delivered through period–1 nodes. The two–stage rule above (tighten with w_S if needed, then move (w_S, w_F) in parallel) preserves the explorer's period–1 IC and minimizes cost under no commitment. This proves the proposition. □

Proof of Proposition 5. Fix the target $\langle {}_1{}^1_t \rangle$ (termination after a first–period failure). Write $w_x = \underline{w} + \tilde{w}_x$ with $\tilde{w}_x \geq 0$.

With termination, only the S-branch continues to period 2. The period–2 IC is the standard one–period effort condition:

$$(p_1 - 0)\,(\tilde{w}_{SS} - \tilde{w}_{SF}) \;\geq\; c_1, \qquad \Rightarrow \qquad \tilde{w}_{SS} = \alpha_1 = \frac{c_1}{p_1}, \;\; \tilde{w}_{SF} = 0.$$

Hence the continuation value above the floor after S is

$$\tilde{V}_2 \;=\; p_1 \tilde{w}_{SS} - c_1 \;=\; p_1 \frac{c_1}{p_1} - c_1 \;=\; 0.$$

Under $p_0 = 0$, a shirker fails for sure and is terminated, earning only the floor in period 1. An exerting exploiter's above–floor value in period 1 is $p_1(\tilde{w}_S + \tilde{V}_2) - c_1 = p_1 \tilde{w}_S - c_1$. Adding the floors,

$$\underbrace{(1 + p_1)\underline{w} + p_1 \tilde{w}_S - c_1}_{\text{exert}} \;\geq\; \underbrace{\underline{w}}_{\text{shirk}} \quad \Longleftrightarrow \quad p_1 \tilde{w}_S \;\geq\; c_1 - p_1 \underline{w}.$$

Thus the least–cost choice is

$$\tilde{w}_S \;=\; \max\left\{\frac{c_1}{p_1} - \underline{w},\; 0\right\} \;=\; \max\{\alpha_1 - \underline{w},\; 0\}, \qquad \tilde{w}_F = 0.$$

Intuitively, termination makes early failure costly (loss of the second–period job and floor), so the principal can substitute some of the monetary carrot in period 1 with the firing stick. When $\underline{w} \geq \alpha_1$, no above–floor first–period bonus is needed.

A misreporting explorer's payoff is weakly increasing in success–contingent bonuses. Here $\tilde{w}_{SS}$ and $\tilde{w}_S$ already sit at their IC–minimal levels for $\langle {}_1{}^1_t \rangle$; any reduction violates the target IC, any increase raises the misreporter's value. Hence if a ride–along cap binds, it must be enforced on the explorer side.

If the exploiter's participation binds, let $\Delta > 0$ be the extra rent to deliver.

Step A (if the misreporting explorer shirks anywhere). Tighten the shirking margin using the success node with the infinite LR against shirking under $p_0 = 0$:

- If he shirks in period 1, raise $\tilde{w}_S$ until his period–1 exploration IC is tight.
- If he then still shirks in period 2 after S, raise $\tilde{w}_{SS}$ (keeping the exploiter's period–2 IC at equality) until he is indifferent after S.

Step B (once the misreporter exerts $\langle {}_2{}^2_2 \rangle$). With all margins exerting, no LR is 0 or $+\infty$. Keep the target's continuation payment ($\tilde{w}_{SS}$) constant and deliver any remaining Δ by increasing $\tilde{w}_S$ only—the first–period success node is the most diagnostic remaining direction against exploration on the IC manifold.

This establishes $\tilde{w}_{SS} = \alpha_1$ and $\tilde{w}_S = \max\{\alpha_1 - \underline{w}, 0\}$ at baseline, the impossibility of relaxing ride–along on the exploiter schedule, and the two–step surplus rule under termination. □

Proof of Corollary 1. We prove each item in turn.

Part (i). Under termination after a first–period failure, failure carries an additional implicit penalty: the agent loses the second–period continuation opportunity and, with limited liability, forgoes at least the second–period floor $\underline{w}$ that would otherwise be available. This "firing stick" substitutes for part of the first–period monetary incentive. In the baseline exploiter contract, the first–period success wage $\tilde{w}_S$ can therefore be reduced by $\min\{\alpha_1, \underline{w}\}$ (where $\alpha_1 \equiv c_1/p_1$), so the reduction in expected implementation cost is

$$p_1 \min\{\alpha_1, \underline{w}\}.$$

For higher required participation (the exploiter's outside option above the baseline), any additional surplus could be delivered for example as an unconditional first–period transfer until that transfer reaches $\min\{\alpha_1, \underline{w}\}$; beyond this point, the termination and no–termination implementations coincide on the remaining relevant branches. Hence termination weakly lowers the implementation cost of exploitation over the relevant nodes, consistent with item (i) of the corollary.

Part (ii). Item (ii) is a selection statement: termination truncates the failure branch and thereby reduces the value of the exploiter contract for a misreporting explorer. The magnitude of this selection effect depends on how much continuation surplus each type obtains on the failure branch under the *no–termination* exploiter contract.

Fix any exploiter contract that implements exertion in period 1, and compare the (no–termination) contract to the same wage schedule with

termination added after a first–period failure. For a type $\theta \in \{\mathrm{E}, \mathrm{X}\}$ and an induced action profile π_θ, define the continuation surplus after failure as the type's expected second–period utility (net of the second–period cost) conditional on a first–period failure under the no–termination contract:

$$S_\theta(F \mid \pi_\theta) \;\equiv\; \mathbb{E}\big[u(w_{F\cdot}) - c(\cdot) \,\big|\, F, \pi_\theta\big].$$

Termination removes this continuation value entirely. Hence the utility loss from introducing termination for type θ can be written as

$$\Delta_\theta \;\equiv\; U_\theta^0 - U_\theta^t \;=\; \Pr(F \mid \pi_\theta)\, S_\theta(F \mid \pi_\theta).$$

By limited liability and the option to shirk in period 2, we always have $S_\theta(F \mid \pi_\theta) \geq \underline{w}$: conditional on F, any type can guarantee the floor by shirking.

(A) The baseline / "no rents behind failure" region. In the baseline cost–minimizing implementation of exploitation (and more generally whenever the post–failure second–period incentive constraint for the exploiter binds, or he optimally shirks after F), the exploiter earns no continuation rent after failure:

$$S_{\mathrm{X}}(F \mid \pi_{\mathrm{X}}) = \underline{w}, \qquad \Rightarrow \qquad \Delta_{\mathrm{X}} = (1 - p_1)\underline{w}.$$

For a misreporting explorer, period–1 exploration succeeds with probability $E[p_[]p_2] \equiv E[p_2]$, so $\Pr(F \mid \pi_{\mathrm{E}}) \geq 1 - E[p_2]$, and $S_{\mathrm{E}}(F \mid \pi_{\mathrm{E}}) \geq \underline{w}$ (with equality whenever the explorer shirks in period 2 after a first–period failure, or is just indifferent there). Hence

$$\Delta_{\mathrm{E}} \;=\; \Pr(F \mid \pi_{\mathrm{E}})\, S_{\mathrm{E}}(F \mid \pi_{\mathrm{E}}) \;\geq\; (1 - E[p_2])\, \underline{w}.$$

Under the maintained ordering that exploitation is safer than exploration, $p_1 > E[p_2]$, so $(1 - E[p_2]) > (1 - p_1)$ and therefore

$$\Delta_{\mathrm{E}} \;\geq\; (1 - E[p_2])\underline{w} \;>\; (1 - p_1)\underline{w} \;=\; \Delta_{\mathrm{X}}.$$

Thus, in this baseline region, termination reduces the misreporting explorer's value from accepting the exploiter contract by more than it reduces the exploiter's, making the exploiter contract less attractive to exploratory types—the selection effect described in item (ii).

(B) When the no–termination contract parks rents on the recovery node FS. If, instead, the no–termination exploiter contract delivers additional surplus on the post–failure branch (e.g. through a higher recovery payment w_{FS}), then the exploiter's continuation surplus after failure satisfies

$$S_{\mathrm{X}}(F \mid \pi_{\mathrm{X}}) = p_1 w_{FS} + (1-p_1) w_{FF} - c_1 = \underline{w} + p_1(w_{FS} - \underline{w}) - c_1 \;>\; \underline{w}$$

(using $w_{FF} = \underline{w}$ under limited liability). In this region the comparison becomes

$$\Delta_{\mathrm{E}} > \Delta_{\mathrm{X}} \quad \Longleftrightarrow \quad \Pr(F \mid \pi_{\mathrm{E}})\, S_{\mathrm{E}}(F \mid \pi_{\mathrm{E}}) \;>\; (1-p_1)\, S_{\mathrm{X}}(F \mid \pi_{\mathrm{X}}).$$

A particularly transparent sufficient condition obtains in the common case where the misreporting explorer still chooses to shirk in period 2 after a first–period failure (or is just indifferent there), so that $S_{\mathrm{E}}(F \mid \pi_{\mathrm{E}}) = \underline{w}$ and $\Pr(F \mid \pi_{\mathrm{E}}) = 1 - E[p_2]$. Then $\Delta_{\mathrm{E}} > \Delta_{\mathrm{X}}$ holds whenever the failure–branch rent on the exploiter contract is not too large:

$$(p_1 - E[p_2])\,\underline{w} \;>\; (1-p_1)\big(S_{\mathrm{X}}(F \mid \pi_{\mathrm{X}}) - \underline{w}\big) = (1-p_1)\big(p_1(w_{FS} - \underline{w}) - c_1\big).$$

In particular, if w_{FS} is increased at most up to the point at which a misreporting explorer is just indifferent between shirking and exerting after failure (so that his continuation surplus remains $\underline{w}$), then $w_{FS} - \underline{w} \le c_2 / E[p_2]$, implying

$$S_{\mathrm{X}}(F \mid \pi_{\mathrm{X}}) \;\le\; \underline{w} + p_1 \frac{c_2}{E[p_2]} - c_1,$$

and the preceding inequality yields an explicit parameter region under which the "termination hurts the explorer more" conclusion continues to hold. When this condition fails (i.e. when sufficiently large rents are

deliberately placed behind failure in the no–termination contract), the truncation effect in item (ii) can be weakened or even reversed; this is why the corollary phrases the claim as "typically."

Part (iii). Suppose the exploiter's participation constraint binds under termination, so the principal must deliver additional surplus $\Delta = \overline{U}_{\mathrm{X}} - U^t_{\mathrm{X}} > 0$. Consider restoring this surplus by increasing the first–period success wage w_S only. Increasing w_S by Δ/p_1 raises the exploiter's expected utility by exactly Δ (since $\Pr(S \mid \langle 1, 1, t\rangle) = p_1$). A misreporting explorer's period–1 success probability on the exploiter contract is at most $q_{\mathrm{E}} \le E[p_2]$, so the same adjustment raises his expected utility by at most

$$q_{\mathrm{E}} \cdot \frac{\Delta}{p_1} \;\le\; \frac{E[p_2]}{p_1}\Delta \;<\; \Delta \qquad (\text{since } p_1 > E[p_2]).$$

Thus the required surplus can be restored while increasing the explorer's expected utility by strictly less than Δ. Equivalently, restoring surplus through w_S *widens* the utility gap between the exploiter and a misreporting explorer on the exploiter contract, preserving (and strengthening) screening against explorers, as claimed in item (iii). □

Proof of Proposition 6. Fix the target plan $a_{\mathrm{E}} = \langle {}_{2}{}^{2}_{t}\rangle$ and write $w_x = \underline{w} + \tilde{w}_x$ with $\tilde{w}_x \ge 0$ for $x \in \{F, S, SF, SS, FS, FF\}$. Under termination, employment in period 2 occurs only after S in period 1.

Because continuation happens only after S, any payment on the failure branch cannot help incentives and is set at the floor: $w_F = w_{SF} = w_{FF} = \underline{w}$. Likewise, $w_S = \underline{w}$: the period–1 exploration margin (with $p_0 = 0$) is driven entirely by the option value of reaching SS; w_S does not tighten the IC and is set to its minimum.

Second period (after S). To induce exploration in period 2,

$$E[p_2 \mid S, 2]\,\tilde{w}_{SS} \;\ge\; c_2 \quad\Rightarrow\quad \tilde{w}_{SS} \;\ge\; \frac{c_2}{E[p_2 \mid S, 2]}. \qquad (\mathrm{IC}_{\langle {}_{2}{}^{0}_{t}\rangle})$$

First period. Compare $\langle {}_{2}{}^{2}_{t} \rangle$ to the two relevant deviations, both of which terminate immediately when $p_0 = 0$:

(i) $\langle {}_{0}{}^{2}_{t} \rangle$ (shirk in $t = 1$, explore in $t = 2$): with $p_0 = 0$, the deviation yields no continuation (failure for sure), so

$$U_{\mathrm{E}}^{\langle {}_{0}{}^{2}_{t} \rangle} = \underline{w}.$$

Under the target,

$$U_{\mathrm{E}}^{\langle {}_{2}{}^{2}_{t} \rangle} = \underline{w}(1 + E[p_2]) + E[p_2]E[p_2 \mid S, 2]\ \tilde{w}_{SS} - (1 + E[p_2])c_2.$$

Thus

$$E[p_2]E[p_2 \mid S, 2]\ \tilde{w}_{SS} \ \geq\ (1 + E[p_2])c_2 - E[p_2]\underline{w}. \qquad (\mathrm{IC}_{\langle {}_{0}{}^{2}_{t} \rangle})$$

(ii) $\langle {}_{0}{}^{0}_{t} \rangle$ (shirk in both periods): again $U_{\mathrm{E}}^{\langle {}_{0}{}^{0}_{t} \rangle} = \underline{w}$, so the same inequality as in (**??**) is obtained:

$$E[p_2]E[p_2 \mid S, 2]\ \tilde{w}_{SS} \ \geq\ (1 + E[p_2])c_2 - E[p_2]\underline{w}. \qquad (\mathrm{IC}_{\langle {}_{0}{}^{0}_{t} \rangle})$$

Collecting the constraints, the minimal success bonus is

$$\tilde{w}^{0}_{SS} = \max\left\{ \frac{c_2}{E[p_2 \mid S, 2]}\ ,\ \frac{(1 + E[p_2])\, c_2 - E[p_2]\underline{w}}{E[p_2]\ E[p_2 \mid S, 2]} \right\},$$

with $w_F = w_{SF} = w_{FF} = w_S = \underline{w}$ and $w_{SS} = \underline{w} + \tilde{w}^{0}_{SS}$, as stated.

Suppose the explorer's participation binds and define $\Delta = \overline{U}_{\mathrm{E}} - U^{0}_{\mathrm{E}} > 0$.

Step A (eliminate any shirking by the misreporting exploiter). If a misreporting exploiter shirks on any margin at the current schedule, increase $\tilde{w}_{SS}$ until he is just indifferent. With $p_0 = 0$, success nodes have infinite likelihood–ratio against shirking, so loading SS first minimizes leakage while preserving $\langle {}_{2}{}^{2}_{t} \rangle$.

Step B (choose the least–leaky IC–preserving direction). Once the misreporter exerts on all margins, compare the likelihood ratios of exploitation relative to exploration at the two informative nodes:

$$\Lambda_{SS} = \Lambda_{\langle 1\,\frac{1}{t}\rangle,\langle 2\,\frac{2}{t}\rangle}(SS) = \frac{p_1^2}{E[p_2]\,E[p_2 \mid S,2]}, \qquad \Lambda_F = \Lambda_{\langle 1\,\frac{1}{t}\rangle,\langle 2\,\frac{2}{t}\rangle}(F) = \frac{1-p_1}{1-E[p_2]}.$$

If $\Lambda_{SS} \leq \Lambda_F$ (double success more diagnostic of exploration), raise $\tilde{w}_{SS}$ alone. If instead $\Lambda_{SS} > \Lambda_F$ (failure more diagnostic), share surplus between F and SS along IC–indifference arcs: first move along the *explorer* IC (maintain the explorer's $t=2$ and $t=1$ ICs),

$$E[p_2]E[p_2 \mid S,2]\; d\tilde{w}_{SS} \;=\; E[p_2]\; d\tilde{w}_F \quad\Longleftrightarrow\quad E[p_2 \mid S,2]\; d\tilde{w}_{SS} = d\tilde{w}_F,$$

until the misreporting exploiter is just indifferent in $t=1$; then continue along the *exploiter* IC under termination,

$$p_1^2\, d\tilde{w}_{SS} = p_1\, d\tilde{w}_F \quad\Longleftrightarrow\quad p_1\, d\tilde{w}_{SS} = d\tilde{w}_F.$$

These paths preserve the binding ICs and minimize leakage to the misreporter. □

Proof of Corollary 2. We establish each item and highlight the parameter regions in which the claimed effects arise.

Part (i). Item (i) is about the *net* cost effect of termination once the explorer's participation constraint binds: termination relaxes incentive constraints (a cost saving), but it also destroys continuation value after early failure (a participation cost).

(A) Incentive effect at the baseline (participation slack). Let $W^{\mathrm{NT}}_{S,F,SS,SF}$ and $W^{\mathrm{T}}_{S,F,SS,SF}$ denote the explorer's *baseline* expected wage costs *restricted to the nodes* $\{S,F,SS,SF\}$, without and with termination, respectively. Under termination, the $F \to \{FS, FF\}$ branch is removed, so

on $\{S, F, SS, SF\}$ the only difference is the double–success bonus. Hence

$$W^{\mathrm{NT}}_{S,F,SS,SF} - W^{\mathrm{T}}_{S,F,SS,SF} = \underbrace{\Pr(SS \mid \langle_{2}{}^{2}_{\cdot}\rangle)}_{=\,E[p_2]\,E[p_2|S,2]} \big(\tilde{w}^{\mathrm{NT}}_{SS} - \tilde{w}^{\mathrm{T}}_{SS}\big). \tag{A.1$'$}$$

With $p_0 = 0$ and $E[p_2 \mid F, 2] = E[p_2]$, the baseline double–success bonuses are

$$\tilde{w}^{\mathrm{NT}}_{SS} = \frac{(1 + E[p_2])\, c_2}{E[p_2]\, E[p_2|S,2]}, \qquad \tilde{w}^{\mathrm{T}}_{SS} = \max\left\{ \frac{c_2}{E[p_2|S,2]},\ \frac{(1 + E[p_2])\, c_2 - E[p_2]\, \underline{w}}{E[p_2]\, E[p_2|S,2]} \right\}.$$

The same two–case algebra as in the original proposition then yields

$$W^{\mathrm{NT}}_{S,F,SS,SF} - W^{\mathrm{T}}_{S,F,SS,SF} = \min\{E[p_2]\, \underline{w},\ c_2\}.$$

Thus, *holding participation fixed at the baseline level*, termination reduces the principal's incentive cost on the surviving implementation nodes, because the firing threat substitutes for part of the monetary incentive (a lower $\tilde{w}_{SS}$).

(B) Participation effect when the explorer's PC binds. Termination also removes continuation value after a first–period failure. Let

$$\Delta_{\mathrm{E}} \ \equiv\ U^0_{\mathrm{E}} - U^t_{\mathrm{E}}$$

denote the explorer's utility loss from introducing termination (computed below in part (ii)); in particular, $\Delta_{\mathrm{E}} \geq (1 - E[p_2])\underline{w}$. If the explorer's participation constraint binds at some required utility $\overline{U}_{\mathrm{E}}$, the principal must replace this lost continuation value through explicit transfers. Relative to the no–termination contract, the *amount of surplus that must be delivered* is therefore higher by (at least) Δ_{E}.

Crucially, restoring utility under termination is not free from an incentive standpoint: because the post–failure continuation branch is truncated, compensation must be loaded onto first–period failure (F) and/or double success (SS). Raising w_F relaxes the explorer's incentives, so

maintaining exploration requires compensating increases in w_{SS}; consequently, the marginal cost of delivering participation utility under termination can be higher than under no termination. In particular, for participation requirements in the region where (i) termination forces a positive restoration $\overline{U}_{\mathrm{E}} - U_{\mathrm{E}}^{t} > 0$ and (ii) the least–leaky restoration relies on w_F (so incentive feasibility requires simultaneously increasing w_{SS}), the additional wage cost needed to satisfy the PC can outweigh the baseline incentive saving $\min\{E[p_2]\,\underline{w}, c_2\}$ computed above. Hence, once the explorer's participation constraint binds, termination *can* increase the cost of implementing exploration.

This is the sense in which termination, despite relaxing IC at the baseline, may raise overall implementation cost when participation is binding, as claimed in item (i).

Part (ii). Item (ii) is a comparative statement about how much termination reduces continuation value for the *explorer* versus a *misreporting exploiter* who accepts the explorer contract. The key object is again the value of the failure branch under the no–termination contract.

Let superscript t indicate termination after a first–period failure. Fix any implementation of exploration on the explorer contract, and compare it to the same wage schedule with termination added after F.

For a type $\theta \in \{\mathrm{E}, \mathrm{X}\}$ and induced period–1 action profile π_θ, define the continuation surplus after failure as the type's expected second–period utility (net of the second–period cost) conditional on a first–period failure under the no–termination contract:

$$S_\theta(F \mid \pi_\theta) \;\equiv\; \mathbb{E}\big[u(w_{F\cdot}) - c(\cdot) \,\big|\, F, \pi_\theta\big].$$

Termination removes this continuation value entirely. Hence the loss from introducing termination for type θ can be written as

$$\Delta_\theta \;\equiv\; U_\theta^0 - U_\theta^t \;=\; \Pr(F \mid \pi_\theta)\, S_\theta(F \mid \pi_\theta).$$

By limited liability and the option to shirk in period 2, $S_\theta(F \mid \pi_\theta) \geq \underline{w}$ for both types.

Typical region. In the canonical implementations considered in the paper (and in particular whenever the misreporting exploiter exerts in period 1 and, following a first–period failure, either shirks in period 2 or exerts with a binding second–period IC), the misreporting exploiter earns no continuation rent above the floor after failure:

$$S_{\mathrm{X}}(F \mid \pi_{\mathrm{X}}) = \underline{w}, \qquad \Rightarrow \qquad \Delta_{\mathrm{X}} = (1 - p_1)\underline{w}.$$

For the explorer, period–1 exploration yields success probability $E[p_2]$, so $\Pr(F \mid \pi_{\mathrm{E}}) = 1 - E[p_2]$, and $S_{\mathrm{E}}(F \mid \pi_{\mathrm{E}}) \geq \underline{w}$. Thus

$$\Delta_{\mathrm{E}} = (1 - E[p_2])\, S_{\mathrm{E}}(F \mid \pi_{\mathrm{E}}) \;\geq\; (1 - E[p_2])\, \underline{w}.$$

Under the maintained ordering $p_1 > E[p_2]$, we have $(1 - E[p_2]) > (1 - p_1)$, and therefore

$$\Delta_{\mathrm{E}} \;\geq\; (1 - E[p_2])\underline{w} \;>\; (1 - p_1)\underline{w} \;=\; \Delta_{\mathrm{X}}.$$

So, in this typical region, termination cuts more into the explorer's continuation value than into the misreporting exploiter's.

Why the corollary says "typically." If instead the explorer contract places substantial rents behind failure in the no–termination regime (e.g. because w_F and/or post–failure continuation pay must rise to satisfy the explorer's participation constraint), then $S_{\mathrm{X}}(F \mid \pi_{\mathrm{X}})$ need not equal $\underline{w}$ and can be larger. In that case the comparison becomes

$$\Delta_{\mathrm{E}} > \Delta_{\mathrm{X}} \quad \Longleftrightarrow \quad (1 - E[p_2])\, S_{\mathrm{E}}(F \mid \pi_{\mathrm{E}}) \;>\; (1 - p_1)\, S_{\mathrm{X}}(F \mid \pi_{\mathrm{X}}),$$

which may fail if the misreporting exploiter's continuation surplus after failure is sufficiently high. This is precisely why item (ii) is stated as a typical (region–dependent) selection effect rather than an unconditional inequality.

Part (iii). Under termination, the explorer loses the continuation value on the failure branch. Let

$$\Delta_{\mathrm{E}} \;\equiv\; U_{\mathrm{E}}^{0} - U_{\mathrm{E}}^{t}$$

denote this loss. If the explorer's participation constraint binds under termination, then restoring the explorer's pre–termination utility requires delivering (at least) Δ_{E} in additional expected utility through the remaining wage nodes.

Consider environments in which the *misreporting exploiter is active* on the explorer contract, in the sense that he exerts in period 1 and, following a first–period failure, either shirks in period 2 or exerts with a binding second–period incentive constraint. In both cases, his continuation surplus after failure under the no–termination explorer contract is exactly the floor, so the misreporting exploiter's own utility loss from introducing termination is

$$\Delta_{\mathrm{X}}^{\mathrm{loss}} \;\equiv\; U_{\mathrm{X}}^{0} - U_{\mathrm{X}}^{t} \;=\; (1-p_1)\,\underline{w}.$$

To restore Δ_{E} while keeping the explorer on the intended exploration plan, the principal adjusts second–period wages at SS and on the failure node F. A marginal change (dw_{SS}, dw_F) affects expected utilities as

$$dU_{\mathrm{E}} \;=\; E[p_2]\,E[p_2|S,2]\,dw_{SS} + (1-E[p_2])\,dw_F, \qquad dU_{\mathrm{X}} \;=\; p_1^2\,dw_{SS} + (1-p_1)\,dw_F.$$

Hence the induced change in the misreporting exploiter's utility per unit restored to the explorer is

$$\frac{dU_{\mathrm{X}}}{dU_{\mathrm{E}}} \;=\; \begin{cases} \Lambda_{SS} \equiv \dfrac{p_1^2}{E[p_2]\,E[p_2|S,2]}, & \text{if } dw_F = 0, \\ \Lambda_F \equiv \dfrac{1-p_1}{1-E[p_2]}, & \text{if } dw_{SS} = 0, \end{cases}$$

and for any nondegenerate mix it is a convex combination

$$\frac{dU_{\mathrm{X}}}{dU_{\mathrm{E}}} \;=\; \omega\,\Lambda_{SS} + (1-\omega)\,\Lambda_F, \qquad \omega \in (0,1).$$

If restoration could be done purely through the failure node ($dw_{SS} = 0$), the leakage to the misreporting exploiter would be $\Lambda_F\,\Delta_{\mathrm{E}}$. Using $\Delta_{\mathrm{E}} \geq (1 - E[p_2])\underline{w}$ (since the explorer can always guarantee the floor after F by shirking in period 2 under no termination), this implies

$$\Lambda_F\,\Delta_{\mathrm{E}} \;\geq\; \Lambda_F(1-E[p_2])\underline{w} \;=\; (1-p_1)\underline{w} \;=\; \Delta_{\mathrm{X}}^{\mathrm{loss}}.$$

However, pure–F restoration is not incentive–feasible: increasing only w_F relaxes the explorer's first–period incentive constraint and undermines the intended exploration behavior. Thus any incentive–feasible restoration must involve $dw_{SS} > 0$, i.e. a nondegenerate mix with $\omega > 0$. In the region where failure is more informative for screening than double success,

$$\Lambda_F < \Lambda_{SS},$$

we therefore have

$$\frac{dU_{\mathrm{X}}}{dU_{\mathrm{E}}} \;=\; \omega\,\Lambda_{SS} + (1-\omega)\,\Lambda_F \;>\; \Lambda_F.$$

Consequently, restoring the explorer fully by Δ_{E} raises the misreporting exploiter's utility by

$$\Delta_{\mathrm{X}} \;\equiv\; U_{\mathrm{X}}^{\mathrm{rest}} - U_{\mathrm{X}}^{t} \;=\; \Big[\omega\,\Lambda_{SS} + (1-\omega)\,\Lambda_{F}\Big]\Delta_{\mathrm{E}} \;>\; \Lambda_{F}\,\Delta_{\mathrm{E}} \;\geq\; \Delta_{\mathrm{X}}^{\mathrm{loss}}.$$

Equivalently,

$$U_{\mathrm{X}}^{t} + \Delta_{\mathrm{X}} \;>\; U_{\mathrm{X}}^{t} + \Delta_{\mathrm{X}}^{\mathrm{loss}} \;=\; U_{\mathrm{X}}^{0},$$

so in these environments *any* restoration of the explorer's lost utility after termination necessarily raises the misreporting exploiter's expected utility above his pre–termination level. This proves part (iii). □

Proof of Proposition 7. To implement the exploiter's action plan $\left\langle {}_{1}^{1}\,{}_{1}^{1} \right\rangle$ at minimum cost, the principal optimally withholds feedback about first-period performance. Without feedback, the agent cannot condition his second-period action on the first-period outcome. This eliminates all second-period contingent incentive constraints and effectively "poisons" the contract for any misreporting explorer: an explorer-type agent who deviates under no-feedback cannot adjust after a failure and is stuck with a lower expected success probability in period 2 (since $E[p_2 \mid F, 2] < E[p_2]$ when $p_0 = 0$). As a result, the no-feedback policy makes deviating less attractive and reduces the cost of inducing effort in both periods.

Under no feedback, the contract must satisfy $w_S = w_F$ – otherwise the agent would learn the period-1 outcome from a differing first-period wage. We can therefore set $w_S = w_F = 0$ without loss of generality: any positive common first-period pay can be deferred to second-period payments, which only lowers the principal's cost while maintaining incentives. Similarly, any reward for total failure can be eliminated. Paying $w_{FF} > 0$ (after two failures) does nothing to encourage effort or truth-telling, so in a cost-minimizing contract we set $w_{FF} = 0$. Thus, we restrict attention to contracts with compensation only in outcomes that include at least one success. Let $\tilde{w}_{SS} > 0$ denote the bonus for two successes, and let

$\tilde{w}_{SF} = \tilde{w}_{FS} > 0$ denote the bonus for a single success (in either period order).

With no feedback, any feasible deviation strategy must choose the same second-period action after success or failure (since the agent is uninformed about the interim result). Thus we consider incentive compatibility constraints $\text{IC}_{\langle i \, j \rangle}$ for $i, j \in \{0, 1\}$, where (for notation) 0 denotes "no effort" (shirking) and 1 denotes "exert effort on the conventional action." It suffices to enforce the following three IC constraints, which cover the most profitable deviations:

- $\text{IC}_{\langle 0 \, 0 \rangle}$ (Shirk in both periods):

$$p_1^2 \, \tilde{w}_{SS} \; + \; 2\, p_1(1 - p_1) \, \tilde{w}_{SF} \; \geq 2c_1.$$

- $\text{IC}_{\langle 1 \, 0 \rangle}$ (Shirk in period 2):

$$p_1^2 \, \tilde{w}_{SS} \; + \; p_1(1 - 2p_1) \, \tilde{w}_{SF} \; \geq c_1.$$

- $\text{IC}_{\langle 0 \, 1 \rangle}$ (Shirk in period 1):

$$p_1^2 \, \tilde{w}_{SS} \; + \; p_1(1 - 2p_1) \, \tilde{w}_{SF} \; \geq c_1.$$

The principal's goal is to satisfy all IC constraints at the lowest possible cost. The cost-neutral frontier is characterized by the locus of $(\tilde{w}_{SS}, \, \tilde{w}_{SF})$ combinations that yield the same total expected wage outlay for the intended plan a_X. Setting $\text{IC}_{\langle 0 \, 0 \rangle}$ to bind with equality (to minimize slack) gives one such relationship:

$$p_1^2 \, \tilde{w}_{SS} \; + \; 2\, p_1(1 - p_1) \, \tilde{w}_{SF} \; = \; 2c_1,$$

which indeed fixes the expected wage $E[W] = 2c_1$. Along this cost-neutral line, increasing the single-success reward $\tilde{w}_{SF}$ allows the principal to reduce the double-success bonus $\tilde{w}_{SS}$ (and vice versa) while keeping $E[W]$ constant. However, raising $\tilde{w}_{SF}$ tightens the one-period deviation

constraints $\mathrm{IC}_{\langle {}^{1}_{0}{}^{0}\rangle}$ and $\mathrm{IC}_{\langle {}^{0}_{1}{}^{1}\rangle}$ (which demand a sufficiently large $\tilde{w}_{SS}$ in comparison). We therefore solve for the **maximum** $\tilde{w}_{SF}$ that preserves those ICs. Substituting $\tilde{w}_{SS}$ from the cost-neutral equation into $\mathrm{IC}_{\langle {}^{1}_{0}{}^{0}\rangle}$, we get a linear condition:

$$p_1^2\Big(\frac{2c_1 - 2p_1(1-p_1)\,\tilde{w}_{SF}}{p_1^2}\Big) \;+\; p_1(1-2p_1)\,\tilde{w}_{SF} \;\geq\; c_1,$$

which simplifies to $2c_1 - p_1\,\tilde{w}_{SF} \geq c_1$. Rearranging, the single-success bonus must satisfy $\tilde{w}_{SF} \leq \frac{c_1}{p_1}$. In other words, $\frac{c_1}{p_1}$ is the largest bonus for one success that still keeps $\mathrm{IC}_{\langle {}^{1}_{0}{}^{0}\rangle}$ (and $\mathrm{IC}_{\langle {}^{0}_{1}{}^{1}\rangle}$) satisfied. At this point both one-period IC constraints bind exactly. Thus, the **least-cost no-feedback contract** is obtained by setting

$$\tilde{w}^*_{SF} \;=\; \tilde{w}^*_{FS} \;=\; \frac{c_1}{p_1}, \qquad \tilde{w}^*_{SS} \;=\; \frac{2c_1}{p_1},$$

with $\tilde{w}_S = \tilde{w}_F = \tilde{w}_{FF} = 0$ as established earlier. The principal's expected payment is

$$E[W] \;=\; p_1^2\,\tilde{w}^*_{SS} + 2\,p_1(1-p_1)\,\tilde{w}^*_{SF} \;=\; 2c_1,$$

which indeed equals the total effort cost for two periods.

Notably, this implementation cost $2c_1$ is the same as under the optimal feedback contract that motivates exploitation. In fact, $\tilde{w}^*_{SF} = c_1/p_1$ coincides with the one-period bonus α_1 that would induce effort with a standard pay-for-performance scheme, and $\tilde{w}^*_{SS} = 2\alpha_1$. In essence, the no-feedback contract replicates a "repeat α_1" bonus in each period (paying α_1 per success).

Under no feedback, the baseline exploiter contract pays bonuses $\tilde{w}_{SS} = 2c_1/p_1$ and $\tilde{w}_{SF} = \tilde{w}_{FS} = c_1/p_1$, with all other payments at the floor. These values are the minimal ones that satisfy the exploiter's incentive constraints, so the ride–along constraint cannot be relaxed by lowering any of them. Moreover, no feedback already reduces the misreporting ex-

plorer's continuation value by preventing action adjustment after failure. If the ride–along constraint binds, it must therefore be enforced on the explorer's side of the menu.

At the baseline contract, one–shot deviations $\langle {}_{2}{}^{0}_{0} \rangle$ and $\langle {}_{0}{}^{2}_{2} \rangle$ are equivalent by the martingale property of beliefs and yield expected utility exactly equal to their cost. Both are thus weakly dominated by full shirking $\langle {}_{0}{}^{0}_{0} \rangle$ or full exertion $\langle {}_{2}{}^{2}_{2} \rangle$. The former gives utility $2\underline{w}$, which is strictly below the explorer's contract utility, so the misreporting explorer must exert.

Let $a_X = \langle {}_{1}{}^{1}_{1} \rangle$ and $a_E = \langle {}_{2}{}^{2}_{2} \rangle$. The relevant likelihood ratios are:

$$\Lambda_{a_X,a_E}(SS) = \frac{p_1^2}{E[p_2]\,E[p_2 \mid S,2]}, \qquad \Lambda_{a_X,a_E}(SF) = \Lambda_{a_X,a_E}(FS) = \frac{p_1(1-p_1)}{E[p_2](1-E[p_2 \mid S,2])}.$$

Since $\Lambda_{a_X,a_E}(SF) > \Lambda_{a_X,a_E}(SS)$, then SF and FS are less leaky and should be used for surplus delivery.

Increases in $\tilde{w}_{SF} = \tilde{w}_{FS}$ will affect $\mathrm{IC}_{\langle {}_{1}{}^{0}_{1} \rangle}$. Hence they must be accompanied by increases in $\tilde{w}_{SS}$ along the exploiter IC arc:

$$p_1^2\, d\tilde{w}_{SS} + p_1(1-2p_1)\, d\tilde{w}_{SF} = 0.$$

This adjustment preserves the exploiter's incentives while delivering surplus through the least leaky outcomes.

□

Proof of Proposition 8. With no feedback, the agent cannot distinguish whether the first-period outcome was a success or failure, so second-period actions must be the same regardless of history. The principal imposes this by setting

$$w_S = w_F, \quad w_{SF} = w_{FS},$$

and the agent must choose the same second-period action after either S or F.

To implement the plan $\langle {}_{2}{}^{2}_{2} \rangle$, the agent must exert effort in both periods regardless of the hidden interim result. Let $\tilde{w}_{SS}$ and $\tilde{w}_{SF} = \tilde{w}_{FS}$ denote the bonuses above the floor paid upon observing two successes or a mixed success, respectively.

Because $p_0 = 0$, the most diagnostic outcome against any deviation involving shirking is success. This implies that the likelihood ratios satisfy

$$\Lambda_{\langle {}_{2}{}^{2}_{2} \rangle, \langle {}_{0}{}^{0}_{0} \rangle}(SS) > \Lambda_{\langle {}_{2}{}^{2}_{2} \rangle, \langle {}_{0}{}^{0}_{0} \rangle}(SF), \quad \Lambda_{\langle {}_{2}{}^{2}_{2} \rangle, \langle {}_{0}{}^{2}_{2} \rangle}(SS) > \Lambda_{\langle {}_{2}{}^{2}_{2} \rangle, \langle {}_{0}{}^{2}_{2} \rangle}(SF),$$

$$\Lambda_{\langle {}_{2}{}^{2}_{2} \rangle, \langle {}_{2}{}^{0}_{0} \rangle}(SS) > \Lambda_{\langle {}_{2}{}^{2}_{2} \rangle, \langle {}_{2}{}^{0}_{0} \rangle}(SF).$$

Hence, by the likelihood–ratio rule, the least–cost contract rewards only the most diagnostic node, SS.

Baseline contract. Let $W \equiv \tilde{w}_{SS}$. Under plan $\langle {}_{2}{}^{2}_{2} \rangle$, expected utility is

$$U_{\mathrm{E}}^{\langle {}_{2}{}^{2}_{2} \rangle} = E[p_2]E[p_2 \mid S, 2]W - 2c_2.$$

The three IC constraints corresponding to deviations $\langle {}_{0}{}^{0}_{0} \rangle$, $\langle {}_{2}{}^{0}_{0} \rangle$, and $\langle {}_{0}{}^{2}_{2} \rangle$ require that

$$\begin{aligned} &\text{(i)} && E[p_2]E[p_2 \mid S, 2]W \geq 2c_2, \\ &\text{(ii)} && E[p_2]E[p_2 \mid S, 2]W \geq c_2, \\ &\text{(iii)} && E[p_2]E[p_2 \mid S, 2]W \geq c_2. \end{aligned}$$

The first inequality is the tightest, so the optimal bonus satisfies

$$\boxed{\tilde{w}_{SS} = \frac{2c_2}{E[p_2]E[p_2 \mid S, 2]}.}$$

All other bonuses are set to zero:

$$\tilde{w}_S = \tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FS} = \tilde{w}_{FF} = 0.$$

Now we will show that withholding feedback is a dominated strategy from the principal's perspective. Without feedback, the agent cannot

distinguish first-period outcomes, so the principal implements plan $\langle {}_2{}^2_2 \rangle$. The optimal contract rewards only the double–success node:

$$\tilde{w}^{\mathrm{NF}}_{SS} = \frac{2c_2}{E[p_2]\, E[p_2 \mid S, 2]}, \qquad \tilde{w}_x = 0 \text{ for all } x \neq SS.$$

This ensures the agent's expected reward covers effort costs $2c_2$:

$$E[p_2]\, E[p_2 \mid S, 2] \cdot \tilde{w}^{\mathrm{NF}}_{SS} = 2c_2.$$

With feedback, the principal can implement plan $\langle {}_2{}^2_2 \rangle$, resetting probability of success after failure to $E[p_2 \mid F, 2] = E[p_2]$, using the exact same contract as in Proposition 1.

In both cases, the agent's utility is $2\underline{w}$ and the expected implementation costs are the same to the principal $2c_2$. However, the revenues of the contract with feedback are higher since the agent is able to reset and obtain probablity of success after failure. The principal strictly prefers to implement exploration with feedback. □