Selecting for Innovation

Gustavo Manso*

June 18, 2026

PRELIMINARY AND INCOMPLETE - NOT FOR CIRCULATION

ABSTRACT

Firms that seek innovation must design incentives not only to induce exploration, but also to attract workers suited to exploratory tasks. This paper develops a screening model in which exploratory and exploitative workers self-select into different incentive schemes. Contracts that attract explorers backload rewards, tolerate early failure, provide job security, require long-term commitment, and use feedback to preserve the value of experimentation while deterring exploiters from mimicking. By distinguishing motivation from selection, the model provides a framework for interpreting empirical evidence on innovation incentives and suggests new tests based on who pursues exploratory work under different incentive schemes.

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

Innovation incentives do more than change behavior inside the firm: they also change who wants to work there. This distinction is central in standard personnel economics. 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. Incentive schemes are therefore not only motivational devices. They are also selection devices.

This paper extends that logic to innovation, but with an important difference from the standard ability-sorting interpretation. In canonical sorting models, high-powered incentives attract workers with higher scalar ability. In innovation settings, the relevant heterogeneity is type-based rather than vertical: some workers are a better fit for exploratory tasks, while others are a better fit for routine execution. As a result, the contracts that select explorers need not be the steepest pay-for-performance contracts. They may be contracts that tolerate early failure, reward long-run success, provide job security, require commitment, and make feedback useful for course correction.

A large theoretical and empirical literature studies how such practices motivate exploration, typically interpreting them through the lens of moral hazard: as inducements for a given worker to experiment rather than exploit known methods. This paper studies the selection counterpart: the same practices may also change who wants to do exploratory work in the first place.

Building on Manso (2011), we embed the exploration–exploitation trade-off from multi-armed bandit problems in a two–period principal–agent model with hidden types. The agent may be an explorer pursuing a novel and uncertain approach or an exploiter applying a safe and well–understood method. The principal observes only successes and failures, not the agent's type or chosen action, and offers a menu of contracts. This stark model isolates the selection force, but the screening problem remains nontrivial: contracts that support experimentation can also provide insurance and rents to exploiters.

The model shows that the same contractual features that motivate exploration—tolerance for early failure, rewards tied to long-run success, commitment not to renegotiate after poor early performance, and informative feedback—also help select explorers. They make experimentation valuable to workers who benefit from learning whether a novel approach is superior, while making the contract less attractive to exploiters, who prefer prompt performance pay and lower exposure to delay. The contract that attracts exploiters has the mirror structure: rewards are more immediate, failure is penalized more sharply, and continuation is less valuable after poor early performance. Innovation contracts are therefore not merely motivational devices. They are also selection devices: they can select innovators.

The empirical implications also differ from the standard motivation-only interpretation. Existing tests typically ask whether changes in compensation, governance, failure tolerance, or feedback increase innovative output. These tests are joint tests of motivation and selection: the same policies may change both the behavior of incumbent workers and the composition of workers who enter, remain, exit, or select into exploratory roles. Observed changes in innovative output may therefore combine two margins: motivation and selection.

Recognizing this joint-test structure generates two additional empirical predictions. First, following Lazear (2000), a complete empirical account should decompose output effects into changes in behavior by incumbent workers and changes in workforce composition. If a firm adopts long-horizon, failure-tolerant contracts, the model predicts not only shifts in patenting or experimentation but also changes in worker flows and contract take-up. Second, the selection logic predicts stronger effects when outside labor markets for exploratory workers are more competitive: firms facing more intense competition for exploration-suited workers stand to gain more from offering differentiated contracts.

The paper relates to three literatures. The first studies worker sorting through incentive pay. Early theories by Salop and Salop (1976), Harris and Raviv (1978), and Lazear (1986) show that firms can use compensation contracts to induce self-selection of workers. Empirical evidence from Lazear (2000), Banker et al. (2001), Eriksson and Villeval (2008), Cadsby et al. (2007), and Dohmen and Falk (2011) confirms that workers select into incentive schemes based on ability, risk attitudes, and other traits. This paper extends the sorting logic from productivity and effort to type-based sorting between exploratory and exploitative work.

The second literature studies incentives for innovation under moral hazard. Holmstrom (1989) argues that innovative activities have noisier performance measures and therefore require more tolerant incentive schemes. Aghion and Tirole (1994) emphasize the role of control rights in spurring innovation. Manso (2011) shows that the optimal contract to motivate innovation involves tolerance for early failure, rewards for long-term success, and managerial job security. This paper focuses on hidden worker types, showing that the same instruments also serve as screening devices.

The third literature studies organizational ambidexterity. This work emphasizes the tension between exploiting established capabilities and

exploring new possibilities, and asks whether firms should pursue exploration inside existing units or separate it into distinct structures (March, 1991; Tushman and III, 1996; O'Reilly and Tushman, 2013). The selection perspective sharpens this organizational distinction. When motivation is the main obstacle, exploration can be integrated into existing units through slack time, internal contests, or broad-based innovation incentives. When selection is the main obstacle, firms may need structurally separate environments that attract exploratory workers under tailored contracts and career paths.

The paper is organized as follows. Section II introduces the baseline principal–agent model of exploration and exploitation and describes the solution strategy and economic mechanism used throughout the analysis. Section III characterizes the optimal exploiter and explorer contracts and highlights the resulting selection effects. Section IV studies implementation with and without commitment on the principal's side, while Section V studies termination. Section VI studies how the provision of feedback reshapes incentives and selection. Section VII discusses empirical implications for compensation, governance, bankruptcy regimes, and intrapreneurship in large organizations. Section VIII concludes. All proofs are in the Appendix.

II The Model

This section sets up a two-period principal–agent model in which a firm offers a menu of limited-liability contracts before observing whether the worker is an explorer or an exploiter. Contracts condition wages on the history of successes and failures, while the worker's type and actions remain hidden. The direct one-principal screening problem is difficult to read because the two contracts are linked by self-selection: each contract must motivate its intended type and also discourage imitation by the other

type. The solution strategy is therefore to use an equivalent two-principal formulation, in which one auxiliary principal designs the explorer contract and another designs the exploiter contract, each subject to a ride-along constraint for the other type. This separates the menu problem into two cost-minimization problems and makes clear how the optimal contracts screen explorers from exploiters.

A Environment and Screening Problem

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 with action 2, the agent can learn from the unsuccessful attempt and pursue a less ambitious fallback version of the exploratory project. As such, I assume that the expected continuation success prob-

ability $E[p_2 \mid F, 2]$ after a first-period failure with action 2 is equal to p_1.

Failure does not literally give the agent access to action 1. Rather, the failed experiment reveals a conventionalized version of the exploratory idea with the same success probability as action 1. For example, work on a radical molecule may reveal a less ambitious but commercially viable fallback candidate, and experience with a frontier design may reveal a more conventional product design.

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 fallback-capability assumption. Without feedback, a failed exploratory attempt need not reveal enough information to make a useful fallback available, capturing the idea that the agent may remain locked into a less promising variant of action 2, with $E[p_2 \mid F, 2] < E[p_2]$.

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 after success and after failure), have success probabilities strictly between zero and one.

Taking a productive action in a period entails a private cost $c > 0$, while shirking is costless. Thus $c_0 = 0$ and $c_1 = c_2 = c$. The principal and the agent are risk neutral and share a unit discount factor. Actions are unobservable.

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

$$\begin{aligned} R(\langle {}_i{}^j_k \rangle) = \{E[p_i]S + (1-E[p_i])F\} + E[p_i]\{E[p_j \mid S, i]S + (1-E[p_j \mid S, i])F\} \\ + (1 - E[p_i])\{E[p_k \mid F, i]S + (1 - E[p_k \mid F, i])F\}, \end{aligned}$$

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

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 in the productive opportunities they can access. Exploiters can use the

conventional action 1 at cost c, but cannot explore. Explorers can use the novel action 2 at cost c, but cannot exploit. This hard-type assumption is a stark benchmark, not a claim that workers literally face infinite costs of other tasks. It shuts down the within-agent choice between exploration and exploitation in order to isolate the selection margin: the principal is not asking how to make a given worker switch from exploitation to exploration, but how to design a menu that attracts workers whose comparative advantage is exploratory while deterring workers whose comparative advantage is exploitation from mimicking them. Thus the model complements the standard moral-hazard analysis of innovation contracts by focusing on workforce composition rather than on effort direction.

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.

I focus on the economically relevant region in which exploratory projects generate greater total surplus than exploitative projects. Because exploration and exploitation have the same effort cost, this means that the exploratory action plan strictly dominates the conventional one in expected output:

$$R(\langle {}_{2}{}^{2}_{2} \rangle) - C(\langle {}_{2}{}^{2}_{2} \rangle) \;>\; R(\langle {}_{1}{}^{1}_{1} \rangle) - C(\langle {}_{1}{}^{1}_{1} \rangle).$$

The screening problem is therefore not whether exploration is socially valuable, but whether the principal can select the exploratory type while preventing exploiters from mimicking the exploration contract.

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

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

The next subsection describes the solution strategy used to characterize the optimal separating menu.

B Solution Strategy

Solving the screening problem directly is not very transparent, because every change in one contract affects the other type's misreporting incentives. The full problem can nevertheless be decomposed into two type-specific contract-design problems, one for explorers and one for exploiters.

I use this decomposition only as an analytical device. Appendix A states the auxiliary two-principal formulation and the associated equivalence result.

The main text uses the implication of that result: a separating menu can be studied by characterizing each type's contract subject to three constraints. The contract must implement the intended action plan for its own type, deter the other type from riding along on that contract, and deliver the intended type's effective participation utility. I denote these effective required utilities by $\overline{U}_{\mathrm{E}}$ and $\overline{U}_{\mathrm{X}}$. They incorporate the outside option and, when relevant, the utility a type could obtain by mimicking the other contract. The appendix keeps track of the auxiliary utility caps used to prove this decomposition; the main text needs only $\overline{U}_{\mathrm{E}}$ and $\overline{U}_{\mathrm{X}}$, which index the rent that must be delivered to each type.

The contract characterizations use a likelihood-ratio logic. For any target plan π, comparison plan π', and outcome history h, define

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

Histories with high likelihood ratios are relatively more likely under the intended plan than under the relevant deviation or misreporting plan. They are therefore the least costly places to put incentives and, when participation binds, the least leaky places to deliver rents.

The construction has three conceptual parts. First, choose the least-cost baseline contract that implements the target action plan for the intended type, ignoring participation and cross-type imitation. Second, check whether the baseline can be modified to deter imitation by the other type while preserving the intended type's incentives. Third, when the intended type's participation constraint binds, deliver the required rent at the highest-likelihood-ratio histories that preserve incentives and separation.

Appendix B proves that the three-step construction achieves the optimal contract. Appendices C and D show that the baseline contracts in the main text and extensions cannot be profitably poisoned: any such adjustment either violates the intended type's incentive constraints or raises cost without improving separation. The main text therefore states the active contract form directly: the least-cost baseline contract plus rent delivery.

When several contracts implement the same action plan at the same cost, I use a simple tie-breaking convention. I first favor contracts that remain robust if shirking has a very small chance of success. If this still leaves more than one optimal contract, I select the limiting contract that pays the agent earlier. This refinement does not change the economic tradeoffs; it only picks a unique representative among equivalent least-cost contracts.

III Optimal Screening Contracts: The Exploiter and Explorer Contracts

This section characterizes the two contracts in the optimal separating menu. Following Manso (2011), I fix the target action plans and solve for the least-cost incentive schemes that implement them under limited liability and hidden actions. The exploiter contract implements repeated use of the known action, $a_{\mathrm{X}} = \langle {}_{1}{}^{1}_{1} \rangle$, and provides the benchmark pay-for-performance scheme. The explorer contract implements repeated experimentation, $a_{\mathrm{E}} = \langle {}_{2}{}^{2}_{2} \rangle$, and shows which features are distinctive to contracts that attract and motivate exploratory workers. Comparing the two contracts isolates the screening role of backloaded rewards, failure tolerance, and the allocation of rents.

A Exploiter Contract

We first characterize the contract that implements exploitation, $a_X = \langle {}_1{}^1_1 \rangle$. This contract is the benchmark against which the explorer contract's distinctive features will be identified. 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. We begin with the least-cost baseline contract and then characterize how surplus is delivered to the exploiter when his participation constraint binds.

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 $\alpha \equiv c/p_1$ 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 $\langle {}_1{}^1_1 \rangle$ then pays

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

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.

The next step is rent delivery. When the exploiter's participation constraint binds and his reservation utility exceeds the baseline level, $\overline{U}_X > 2\underline{w}$, the principal must 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. Because success is more likely for an exploiter than for an explorer, the principal delivers rent through $\tilde{w}_S$ only, keeping $\tilde{w}_{SS}$ and $\tilde{w}_{FS}$ at their incentive minimal levels.

PROPOSITION 1 (Exploiter Contract)*: The optimal exploiter contract has two cases.*

If $\overline{U}_X \leq 2\underline{w}$*, the baseline contract is optimal. It implements exploitation* $a_X = \langle {}_1{}^1_1 \rangle$ *by paying a constant bonus for each success and zero after failure:*

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

If $\overline{U}_X > 2\underline{w}$*, the principal uses the baseline contract and delivers the required extra surplus* $\Delta = \overline{U}_X - 2\underline{w} > 0$ *by increasing the first-period success payment* $\tilde{w}_S$ *only, keeping* $\tilde{w}_{SS}$ *and* $\tilde{w}_{FS}$ *at their incentive–minimal levels.*

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

$$\tilde{w}_S = \tilde{w}_{SS} = \tilde{w}_{FS} = \alpha = \frac{c}{p_1} = \frac{5}{7}, \qquad \tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0,$$

so that under the target plan $\langle {}_1{}^1_1 \rangle$ the exploiter's expected utility equals $2\underline{w} = 0.8$.

In the numerical solution, the optimal contract holds $\tilde{w}_{SS} = \tilde{w}_{FS} = \alpha = 5/7$ 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 bonus schedule: $\tilde{w}_S$ has a kink at $2\underline{w} = 0.8$, 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

[1] The benchmark parameters are $p_0 = 0$, $E[p_2] = 0.60$, $E[p_2 \mid S, 2] = 0.95$, $p_1 = 0.70$, $c = 0.50$, and a wage floor $\underline{w} = 0.40$.

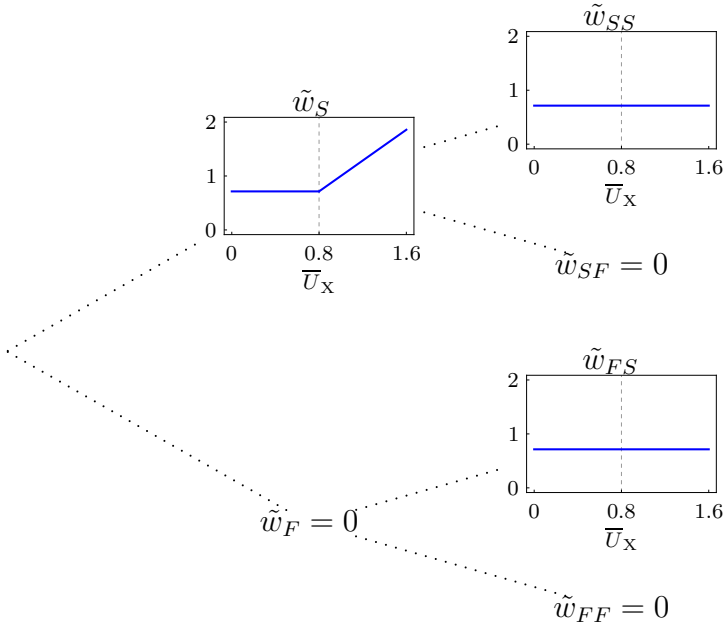


Figure 1. Optimal contract that implements exploitation as a function of the minimal rent $\overline{U}_X$ under the benchmark parameters. Total pay resembles a repetition of standard pay–for–performance across periods; when participation binds, extra surplus is delivered by raising only the first–period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

with predictable outcomes, the optimal incentive relies on immediate and proportional rewards for observable success. The resulting contract provides high–powered, short–term incentives that tie compensation tightly to current performance, ensuring continuous effort. 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 $a_E = \langle {}_2{}_2^2 \rangle$. This is the central contract in the paper because it shows how incentives for experimentation also screen for explorers. The principal faces a two–period moral–hazard problem under limited liability, with the added complication that an exploiter may misreport as an explorer. We begin with the least-cost baseline contract and then characterize how surplus is delivered to the explorer when his participation constraint binds.

In the baseline case where participation and ride–along constraints are slack, failure payoffs are set at the floor and effort is induced through success-contingent bonuses that separate by history. After a first-period failure, the continuation success probability is $E[p_2 \mid F, 2] = p_1$, so continuation effort requires a recovery bonus $\tilde{w}_{FS} = c/p_1$. After a success, the double-success bonus must also provide the first-period option value of reaching the successful branch, yielding

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

The explorer earns baseline utility $2\underline{w}$.

The first-period success payment $\tilde{w}_S$ does not enter the 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 the first-period incentive on $\tilde{w}_{SS}$ and keeps $\tilde{w}_S$ at the minimum.

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 recovery bonus $\tilde{w}_{FS}$ remains fixed at c/p_1. Extra rents are instead deliv-

ered through histories that are most indicative of exploration. The principal first raises $\tilde{w}_{SS}$, a history reached only after continued exploratory success, until a misreporting exploiter is just indifferent between shirking and exerting. After that point, surplus is allocated between $\tilde{w}_{SS}$ and, when appropriate, $\tilde{w}_F$ according to the likelihood-ratio comparison of the explorer and exploiter plans.

PROPOSITION 2 (Explorer Contract)*: The optimal explorer contract has two cases.*

If $\overline{U}_E \leq 2\underline{w}$*, the baseline contract is optimal. It implements exploration* $a_E = \langle {}_2{}^2_2 \rangle$ *by paying*

$$\tilde{w}_S = \tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0, \qquad \tilde{w}_{FS} = \frac{c}{p_1}, \qquad \tilde{w}_{SS} = \frac{(1 + E[p_2])\, c}{E[p_2]\, E[p_2 \mid S, 2]}.$$

If $\overline{U}_E > 2\underline{w}$*, the principal uses the baseline contract and delivers the required extra surplus* $\Delta = \overline{U}_E - 2\underline{w} > 0$ *as follows. The principal first raises* $\tilde{w}_{SS}$ *until the misreporting exploiter is just indifferent between working and shirking. Once this first-period IC binds, any remaining surplus is delivered by increasing either:*

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

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

(ii) $\tilde{w}_{SS}$ *and* $\tilde{w}_F$ *jointly along the misreporting exploiter's first-period IC frontier, if failure is more diagnostic, that is, if*

$$\Lambda_{\langle {}_2{}^2_2 \rangle, \langle {}_1{}^1_1 \rangle}(SS) < \Lambda_{\langle {}_2{}^2_2 \rangle, \langle {}_1{}^1_1 \rangle}(F).$$

Along this frontier,

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

Figure 2 illustrates the optimal explorer contract under the benchmark parameters. When $\overline{U}_E \leq 2\underline{w}$, the baseline contract uses only the second–period bonuses $\tilde{w}_{FS}$ and $\tilde{w}_{SS}$ to induce exploration. When $\overline{U}_E > 2\underline{w}$, the principal must deliver additional surplus while preserving incentives and limiting leakage to a misreporting exploiter. The first increments of surplus are delivered through $\tilde{w}_{SS}$ alone: raising $\tilde{w}_{SS}$ relaxes the explorer's participation constraint while remaining relatively inexpensive in terms of exploiter rent. Once the exploiter incentive constraint becomes binding, however, the principal no longer raises $\tilde{w}_{SS}$ in isolation. Under the benchmark parameters, failure is more informative of exploration than SS, so further surplus delivery proceeds along the exploiter IC arc by raising $\tilde{w}_F$ and $\tilde{w}_{SS}$ together. The resulting wage paths exhibit kinks at the transition from the baseline contract to $\tilde{w}_{SS}$-only rent delivery, and again when the contract switches to joint adjustment along the IC arc.

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 likely to produce 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 continued success, the surplus–delivery step can even tilt pay toward F, rewarding short–term failure. This tolerance 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 perfor-

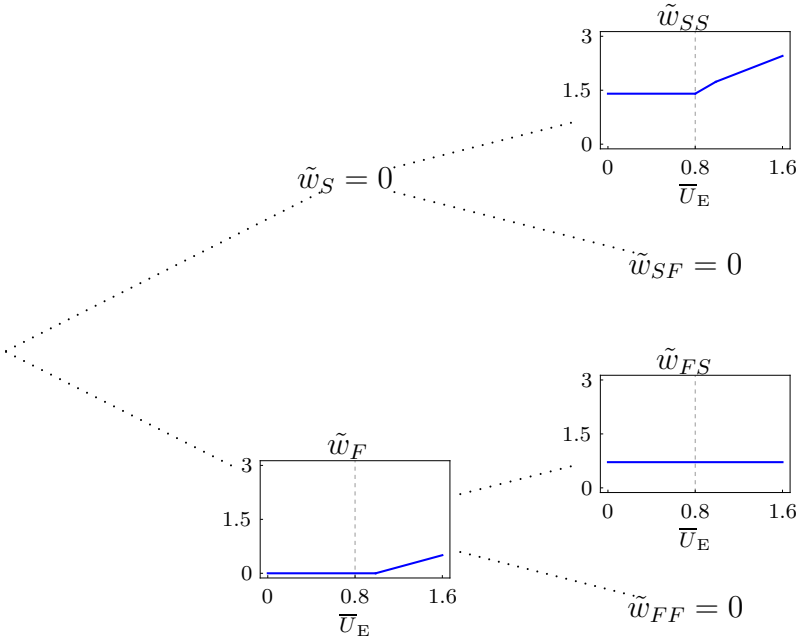


Figure 2. Optimal contract that implements exploration as a function of the minimal rent $\overline{U}_E$. The recovery bonus $\tilde{w}_{FS}$ remains fixed at its incentive-minimal level. When participation binds, extra surplus is delivered first by raising $\tilde{w}_{SS}$ and then by raising $\tilde{w}_{SS}$ and $\tilde{w}_F$ together.

mance. 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.

The explorer–exploiter contract menu shows how compensation can screen explorers from exploiters when the principal can commit to the full state-contingent wage schedule, retain the agent after all interim histories, and provide the feedback needed for the agent to adapt. The next three sections relax these institutional assumptions. Commitment, termination, and feedback are not separate departures from the selection logic; they are levers that determine where rents can be promised, whether early failure is insured or punished, and whether exploration preserves its option value. Each extension therefore asks how the institution changes the relative attractiveness of exploratory and exploitative work.

IV Lack of Commitment

This section asks how the selection logic changes when the principal cannot commit to a long–term contract and is restricted to a sequence of one–period contracts. Commitment matters because the explorer-attracting contract uses future histories to deliver rents where they leak least to exploiters. If those future payments cannot be promised at the hiring stage, the principal loses precisely the instrument that makes exploration attractive without also attracting mimics. The effect is asymmetric: exploitation is naturally short-term, while exploration relies on credible long-horizon rewards and protection after early failure.

A Exploitation without commitment

The exploitation contract relies only on short–term incentives: the principal rewards current success and pays nothing after failure. As shown in Manso (2011), this baseline can be implemented by a sequence of one–period contracts. The same remains true when the exploiter's outside option is high. Proposition 1 delivers any additional surplus through the contemporaneous first-period success payment $\tilde{w}_S$, not through deferred payments such as $\tilde{w}_{SS}$ or $\tilde{w}_{FS}$. Hence both the baseline and the surplus-delivery region of the optimal exploiter contract can be replicated without long-term commitment. 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 optimal exploiter contract in Proposition 1 can be implemented via a sequence of short–term contracts. The baseline contract uses only current success bonuses, and when the participation constraint requires additional surplus, the principal raises only the first-period suc-*

cess payment $\tilde{w}_S$. Commitment therefore has no effect on the optimal exploiter contract.

In summary, commitment is irrelevant for the exploitation contract. The baseline contract is a sequence of one-period incentive contracts, and the surplus-delivery region also uses only the contemporaneous first-period success payment $\tilde{w}_S$. There is therefore no committed-contract advantage to preserve on the exploiter side. This contrasts with the exploration problem, where deferred rewards remain 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 back-loads rewards onto second-period histories. The key payment is $\tilde{w}_{SS}$: it is set above the continuation-IC minimum so that the explorer receives rents after the history that is most informative of successful experimentation. This is precisely the rent-placement instrument that selects explorers while limiting the value of mimicking for exploiters.

Without commitment, the principal cannot promise this above-minimum second-period success rent at the hiring stage. The payment $\tilde{w}_{SS}$ must fall to the level needed only to induce second-period exploration after first-period success. To restore first-period experimentation and deliver any participation surplus, the principal must shift incentives forward into the first-period wages $\tilde{w}_S$ and $\tilde{w}_F$. When the explorer's outside option exceeds the baseline utility $2\underline{w}$, the least-leaky no-commitment adjustment raises these two first-period wages in parallel, preserving the explorer's first-period incentive constraint. But this is still costly for selection: un-

like the committed rent on $\tilde{w}_{SS}$, first-period payments are also attractive to exploiters who mimic the explorer, so the ride-along payoff rises substantially relative to the commitment benchmark. The formal statement below records the resulting baseline contract and surplus-delivery rule.

PROPOSITION 4 (Exploration without commitment)*: Under no commitment, suppose the principal wishes to implement the explorer's target plan* $\langle {}_2{}^2_2 \rangle$*. If* $\overline{U}_{\mathrm{E}} \leq 2\underline{w}$*, the optimal contract is the baseline contract*

$$\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0, \qquad \tilde{w}_{FS} = \frac{c}{p_1}, \qquad \tilde{w}_{SS} = \frac{c}{E[p_2 \mid S, 2]}, \qquad \tilde{w}_S = \frac{c}{E[p_2]}.$$

If $\overline{U}_{\mathrm{E}} > 2\underline{w}$*, the principal delivers the additional surplus* $\Delta = \overline{U}_{\mathrm{E}} - 2\underline{w}$ *by increasing* $\tilde{w}_S$ *and* $\tilde{w}_F$ *in parallel. Under* $E[p_2] < p_1$*, a misreporting exploiter strictly prefers to exert rather than shirk in the first period, so there is no separate shirking case.*

COROLLARY 1 (No commitment raises exploiter ride-along rents)*: Under the optimal exploration contract, the ride-along payoff of a misreporting exploiter is weakly higher without commitment than with commitment, and strictly higher whenever the explorer participation constraint binds.*

The corollary follows because losing commitment forces surplus delivery onto first-period wages that an exerting exploiter also receives, rather than allowing the principal to concentrate rents on histories that are relatively more diagnostic of exploration.

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 placing rents on histories that are relatively unattractive to exploiters, especially first–period failure and double success, $(\tilde{w}_F, \tilde{w}_{SS})$, while keeping first–period success 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 $(\tilde{w}_S, \tilde{w}_F)$. Since $E[p_2] < p_1$, the misreporting exploiter strictly prefers to exert in the first period, so the optimal no-commitment surplus-delivery policy moves $\tilde{w}_S$ and $\tilde{w}_F$ in parallel from the start. This front-loading maintains incentives for exploration but makes the explorer menu more attractive to exploiters, weakening the selection margin relative to the commitment benchmark.

Figure 4 plots the ride–along payoff of a misreporting exploiter under commitment (solid) and no commitment (dashed) as a function of $\overline{U}_{\mathrm{E}}$. Because $E[p_2] < p_1$, the misreporter exerts already at the no-commitment baseline. As the explorer's outside option rises, the curves diverge: commitment allows the principal to place rents on histories that are informative of exploration, especially first-period failure and double success 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 is irrelevant for exploitation but central for exploration. The optimal exploiter contract, including the region in which additional participation surplus must be delivered, can be implemented period by period because all relevant payments are contemporaneous success rewards. The explorer contract is different: without commitment, the principal cannot use promised continuation rents to screen at the hiring stage and must instead front-load incentives on first-period wages. At the no-commitment baseline this front-loading restores the explorer's incentives but does not create a selection advantage; it also gives an exerting exploiter a positive ride-along value. Commitment matters for exploration because it preserves the ability to place rents on histories that are informative about experimentation,

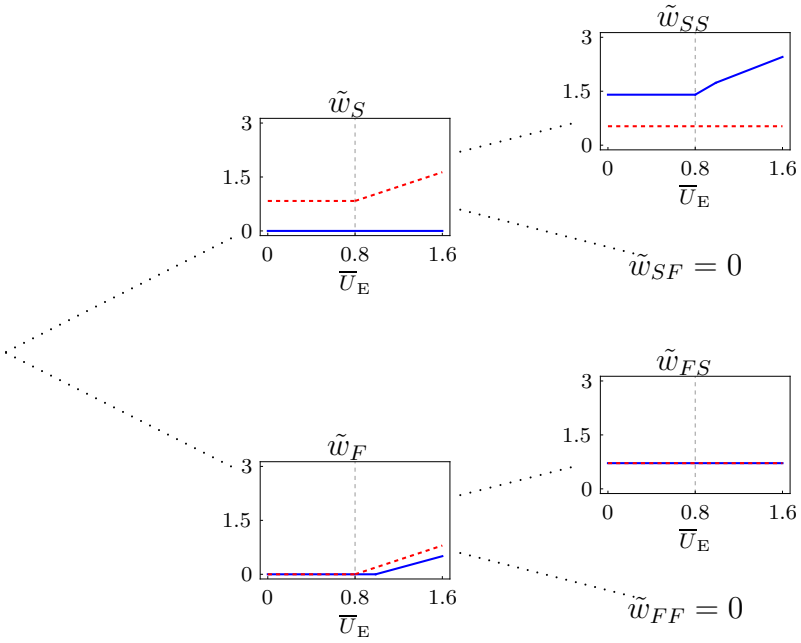


Figure 3. Optimal contract that implements exploration as a function of the minimal rent $\overline{U}_E$, with (solid blue curves) and without (dashed red curves) commitment. With commitment, extra surplus is delivered through history-contingent payments; without commitment, extra surplus is front-loaded on first-period wages.

thereby lowering the mimicking value relative to the no-commitment explorer menu.

V Termination

This section studies termination after first-period failure. Termination can be interpreted as firing a manager or discontinuing funding for a project following poor interim performance. I focus on its selection consequences: termination changes who wants a contract by removing continuation after early failure. This can make exploitation contracts less attractive to explorers, who fail early more often, but it can also make exploration contracts less attractive to precisely the workers who need protection while experimenting.

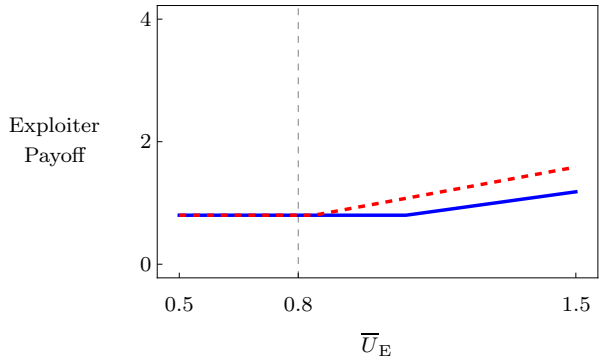


Figure 4. Exploiter ride–along payoff under the explorer contract with (solid blue curves) and without (dashed red curves) commitment as a function of the minimal rent $\overline{U}_E$. Commitment allows the principal to place rents on informative histories, keeping the mimicking payoff low; without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

I use the superscript t to denote termination after first-period failure: $\langle {}_i{}^j_t \rangle$ means that the worker takes action i in period 1, takes action j after first-period success, and is terminated after first-period failure. Thus $\langle {}_1{}^1_t \rangle$ and $\langle {}_2{}^2_t \rangle$ are the termination versions of the exploiter and explorer plans, respectively.

A Exploiter Contract with Termination

Allowing the principal to terminate the agent after a first-period failure changes selection into the exploitation contract by removing the continuation value of early failure. Explorers reach that failure branch more often than exploiters, so termination makes the exploitation menu less attractive to exploratory types. The terminated contract keeps the surviving success branch of the baseline exploiter contract but shuts down the failure branch: the agent works in the second period only after first-period success, $\tilde{w}_{SS}$ remains at c/p_1, and the first-period success bonus $\tilde{w}_S$ can be lower because dismissal after failure already provides discipline. When the exploiter's participation constraint binds, the least-leaky way to re-

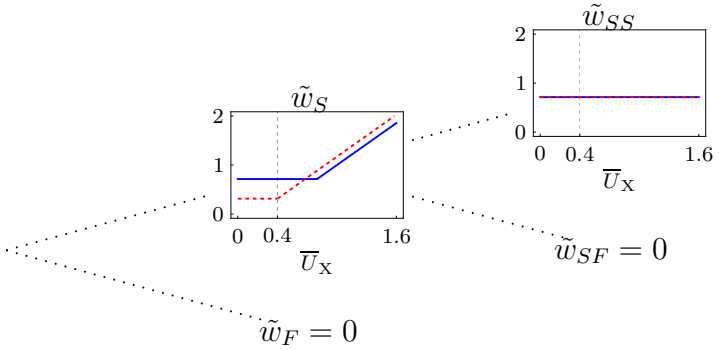


Figure 5. Optimal contract that implements exploitation as a function of the minimal rent $\overline{U}_X$, without termination (solid blue curves) and with termination (dashed red curves). With termination, the failure branch is shut down and incentives are concentrated on the surviving success branch.

store surplus is the first-period success payment $\tilde{w}_S$, a history reached more often by the target exploiter than by a mimicking explorer.

PROPOSITION 5 (Exploiter contract with termination)*: The optimal contract that implements exploitation with termination,* $a_{\mathrm{X}} = \langle {}_{1}{}^{1}_{t} \rangle$*, is as follows. If* $\overline{U}_X \leq \underline{w}$*, the optimal contract is the baseline contract*

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

If $\overline{U}_X > \underline{w}$*, the principal delivers the additional surplus* $\Delta = \overline{U}_X - \underline{w}$ *by increasing* $\tilde{w}_S$ *alone, leaving* $\tilde{w}_{SS}$ *unchanged.*

Figure 5 displays the optimal exploiter contract with and without termination in the relevant branches. For low participation requirements, termination lowers $\tilde{w}_S$ because the loss of continuation after failure substitutes for part of the monetary incentive needed to induce first-period effort. Once the exploiter's participation constraint binds, however, the principal restores the required surplus through $\tilde{w}_S$, so the terminated contract can require a higher first-period success payment than the no-termination benchmark. The continuation payment $\tilde{w}_{SS}$ stays at its baseline level, while the failure branch is shut down.

COROLLARY 2 (Termination and selection on the exploiter contract)*: Termination after early failure affects the selection of exploitation in two ways:*

(i) *Termination reduces the value of the exploiter contract more for a misreporting explorer than for an exploiter, making the contract less attractive to exploratory types.*

(ii) *If the exploiter's participation constraint binds, the principal can restore the required exploiter surplus while giving strictly less surplus back to the misreporting explorer.*

Taken together, Proposition 5 and Corollary 2 show how early termination changes selection into the exploiter contract. Termination removes the failure branch, which is reached more often by a mimicking explorer than by the target exploiter. This makes the exploiter menu less attractive to exploratory types. When the exploiter's participation constraint binds, the principal restores the exploiter's surplus through the first-period success payment $\tilde{w}_S$, a history reached more often by exploiters than by explorers. The same instrument that maintains the exploiter's participation therefore gives back less to the misreporting explorer. Termination thus 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 changes selection into the exploration contract by removing the very continuation value that makes experimentation attractive. Truthful explorers reach early failure relatively often, so terminating after that history makes the explorer menu less attractive to the type it is meant to select. At the baseline, however, termination also strengthens first-period incentives, so the contract can rely less on the double-success payment $\tilde{w}_{SS}$ than the

no-termination explorer contract. When the explorer's participation constraint binds, the principal must restore surplus on histories that preserve selection against a misreporting exploiter. The least-leaky direction depends on diagnosticity: surplus is placed on SS when double success is more diagnostic of exploration, while if failure is more diagnostic the principal shares surplus between F and SS along the misreporting exploiter's first-period IC frontier.

PROPOSITION 6 (Optimal explorer contract with termination)*: The optimal contract that implements exploration with termination,* $\langle {}_2{}^2_t \rangle$*, is as follows. If* $\overline{U}_{\mathrm{E}} \leq \underline{w}$*, the optimal contract is the baseline contract*

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

with a double success-contingent bonus

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

If $\overline{U}_{\mathrm{E}} > \underline{w}$*, the principal delivers the additional surplus* $\Delta = \overline{U}_{\mathrm{E}} - \underline{w}$ *in two steps. First, raise* $\tilde{w}_{SS}$ *until the misreporting exploiter is just indifferent between working and shirking in the first period. Once this exploiter IC binds, define*

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

(i) If $\Lambda_{SS} \geq \Lambda_F$*, raise* $\tilde{w}_{SS}$ *only.*

(ii) If $\Lambda_{SS} < \Lambda_F$*, share surplus between* F *and* SS *along the misreporting exploiter's first-period IC frontier,*

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

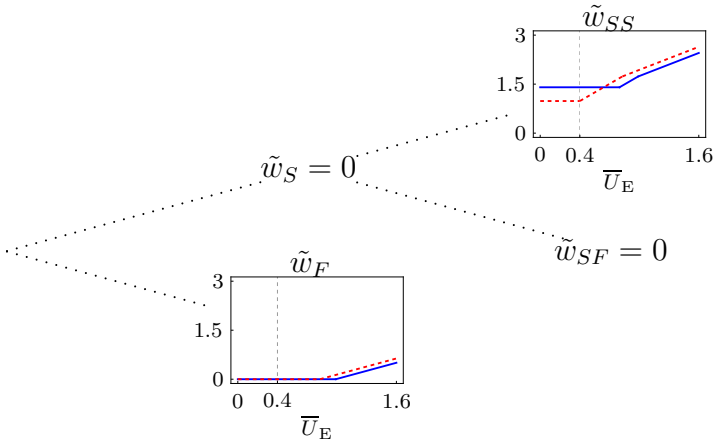


Figure 6. Optimal contract that implements exploration as a function of the minimal rent $\overline{U}_{\mathrm{E}}$, without termination (solid blue curves) and with termination after a first-period failure (dashed red curves). With termination, the failure branch is shut down and surplus is delivered on the surviving success branch and, when failure is diagnostic, on the first-period failure payment.

Figure 6 displays the optimal explorer contract with and without termination for the benchmark parameters. For low values of $\overline{U}_{\mathrm{E}}$, termination lowers $\tilde{w}_{SS}$ relative to the no-termination benchmark because dismissal after first-period failure strengthens first-period incentives. This ranking reverses once the explorer's participation constraint requires additional surplus. Under the benchmark parameters, failure is more diagnostic of exploration, so the terminated contract restores surplus through the failure payment $\tilde{w}_F$ together with the double-success payment $\tilde{w}_{SS}$ along the misreporting exploiter's IC frontier.

COROLLARY 3 (Termination and selection on the explorer contract)*: Termination after early failure affects the selection of exploration in two ways:*

(i) *Termination reduces the value of the explorer contract more for a truthful explorer than for an exerting misreporting exploiter.*

(ii) *If the explorer's lost continuation value is restored in the failure-diagnostic region, the restoration raises the exerting misreporting exploiter's payoff above its no-termination level.*

Taken together, Proposition 6 and Corollary 3 show how early termination changes selection into the explorer contract. Termination removes the failure branch, which is reached more often by a truthful explorer than by a misreporting exploiter. This makes the explorer menu less attractive to the very type it is meant to select. When the explorer's participation constraint binds, restoring the lost continuation value requires shifting compensation onto histories that a misreporting exploiter also reaches with high probability. In the failure-diagnostic region, restoration raises $\tilde{w}_F$ and must be paired with additional $\tilde{w}_{SS}$ to preserve the misreporting exploiter's IC; because the exploiter benefits from both the restored failure payment and the compensating success payment, his ride-along payoff increases relative to the contract with continuation. Termination therefore has a selection cost for exploration: it disciplines early failure, but after participation is restored it leaves the explorer contract more attractive to exploiters, weakening the contract's ability to select explorers.

VI Feedback

This section studies feedback as an information-design lever. The principal privately observes interim performance at the end of the first period but can choose whether to reveal it to the agent. Feedback affects the agent's ability to condition future actions on past outcomes and therefore changes both incentive provision and self-selection. Withholding feedback forces the agent to choose a single continuation strategy after success and failure. Providing feedback, in contrast, allows the principal to tailor incentives to each realized outcome and allows the agent to condition second-period actions on first-period output.

A key role of feedback in our environment is to preserve the fallback option in exploration. After a first-period failure with the novel action 2, revealing interim performance allows the agent to recognize the failure,

abandon the unsuccessful variant, and switch to the fallback option in period 2, restoring the success probability to p_1. If feedback is withheld, the agent cannot tell whether the first period succeeded or failed and therefore cannot condition on the fallback option; following a failure he is effectively stuck with a lower continuation success probability (i.e., $E[p_2 \mid F, 2] < E[p_2]$). Feedback is therefore especially valuable to explorers and much less valuable to exploiters, making it a selection instrument as well as an informational input.

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. 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.[2] 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

[2] 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.

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)*: The optimal contract that implements exploitation without feedback is as follows. If* $\overline{U}_X \leq 2\underline{w}$*, the optimal contract is the baseline contract*

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

The ride-along constraint is 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, withholding feedback strictly poisons *the exploitation contract for the explorer, making misreporting unattractive without affecting implementation costs for the target type.*

If $\overline{U}_X > 2\underline{w}$*, the principal delivers the additional surplus 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 changes 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 7 and 8 illustrate how withholding feedback affects the exploiter contract and the explorer's ride–along payoff. In Figure 7, the blue solid curves show the full–feedback contract from Section III, 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 $\tilde{w}_S = \tilde{w}_F$ and the principal replicates a simple pay–for–performance scheme that pays a constant bonus per success.

Figure 8 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 $\overline{U}_{\mathrm{X}}$. 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 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 second–period action after success and failure and thus forgo the option

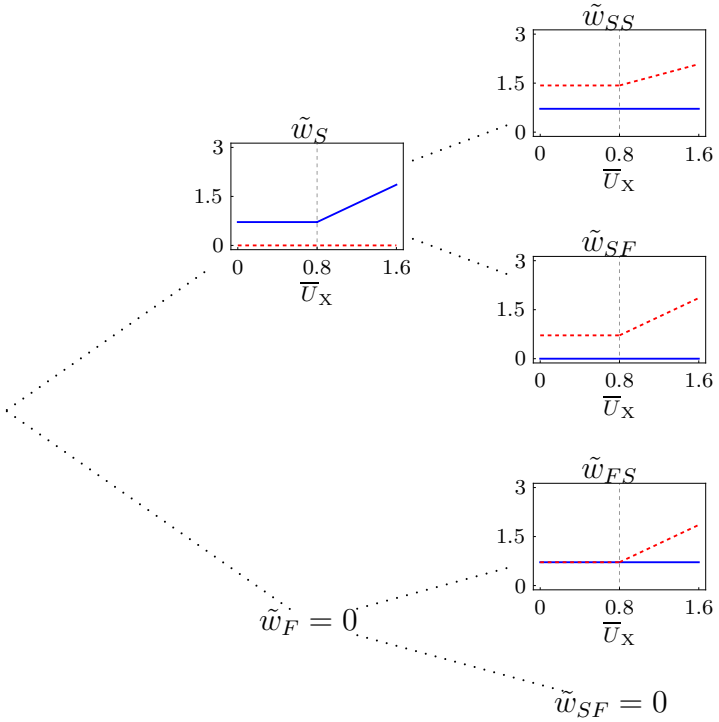


Figure 7. Optimal contract that implements exploitation as a function of the minimal rent $\overline{U}_X$, with feedback (solid blue curves) and without feedback (dashed red curves). Without feedback, the agent cannot condition on the first-period outcome, so the wage schedule collapses to a common pay-for-performance bonus across histories.

value of refreshing 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, after an actual failure he cannot exercise the fallback option and instead continues with the same instance of action 2, facing a lower success probability $E[p_2 \mid F, 2] < E[p_2]$.

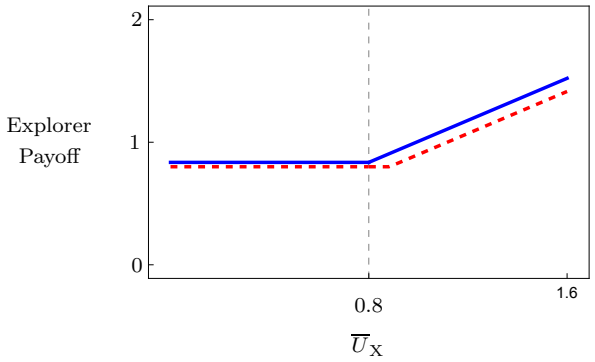


Figure 8. Explorer ride-along payoff under the exploiter contract as a function of the minimal rent $\overline{U}_X$, with feedback (solid blue line) and without feedback (dashed red line). Without feedback, a misreporting explorer cannot condition second-period behavior on early failure, so the exploiter menu is less attractive to exploratory types.

By contrast, if feedback is provided, the agent learns whether the original approach was successful and can use the fallback option after failure. The continuation success probability after failure is then restored to p_1. 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* $\langle 2_2^2 \rangle$*, or (ii) the feedback contract derived in Proposition 2 that implements adaptive continuation via* $\langle 2_2^2 \rangle$*. In both cases, the agent's expected utility and the principal's incentive cost are identical. However, with feedback the agent can use the fallback option after first-period failure and restore the continuation success probability to* p_1*, 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 after the initial experiment, he cannot use the fallback option after failure and is stuck with the lower continuation success probability $E[p_2 \mid F, 2]$. The principal can still induce $\langle {}_2{}^2_2 \rangle$ and deliver any required participation surplus at the same expected cost as under feedback, but the value of the project falls because the contract destroys the option value of fallback after an unsuccessful trajectory.

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 principal 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 preserving the fallback option after a failed experiment. The principal strictly prefers to provide feedback when implementing exploration. In equilibrium, feedback is used to preserve the fallback option in exploration and withheld to harden screening on the exploitation side.

VII Empirical implications

Manso (2011) motivates the empirical agenda for incentives and innovation with two natural tests. First, one can ask whether incentive practices used in tasks where innovation is more important resemble the practices that the theory identifies as exploration-friendly, while incentive practices used in tasks where innovation is less important resemble exploitation-friendly contracts. Second, one can ask whether adopting

exploration-friendly incentive practices actually leads to more innovation, and whether adopting exploitation-friendly practices leads to less innovation. These tests are central because they connect the contractual features emphasized by the theory—long horizons, tolerance for early failure, and feedback—to observed incentive systems, innovation outcomes, and the sorting of exploratory talent.

Those predictions also apply here. Contracts that insure early failure, reward long-run success, provide commitment, and make feedback useful for course correction should be used more when exploration is valuable, and their adoption should be associated with greater innovative output. The empirical predictions are thus fundamentally a joint test of motivation and selection. Innovation may rise because a fixed set of workers responds to a more exploration-friendly contract by experimenting more, or because the contract changes the composition of workers who enter, remain in, or choose exploratory work. The contribution of the present paper is to make the second channel a first-order empirical object. The same contract features that motivate exploration also screen for workers with a comparative advantage in exploratory work. Empirical tests should therefore ask not only whether innovation outcomes change after an incentive reform, but also whether the reform changes who works under the contract.

This distinction is analogous to the conceptual separation in Lazear (2000): in both settings, the observed effect of an incentive policy combines a within-worker effort response with a compositional change in the workers who operate under the policy. The analogy is conceptual rather than literal. In Lazear's setting, productivity is essentially scalar and selection operates through firm entry. Here, exploratory type is comparative—that is, it captures advantage in exploratory relative to exploitative tasks—and selection may operate through entry, retention, contract choice, assignment to exploratory roles, or exit after a policy reversal. An empirical design that observes only patents, citations, new

products, or other innovation outcomes estimates the sum of a *motivation effect*—incumbent workers exert more exploratory effort, holding type fixed—and a *selection effect*—workers with stronger exploratory ability or preferences sort toward environments that insure early failure and reward long-run success.

The selection view adds two empirical predictions that distinguish this paper from a pure motivation theory of incentives for innovation.

Prediction 1: exploration-friendly contracts change worker flows. When a firm, university, funding agency, or entrepreneurial ecosystem adopts incentives that are more tolerant of early failure, more long-horizon, more committed, or more feedback-rich, workers with a comparative advantage in exploration should be more likely to enter, remain in, choose, or be assigned to exploratory work. Conversely, when those features are weakened or rolled back, exploratory workers should be more likely to leave the organization or move out of exploratory roles. The relevant empirical margins include contract take-up, entry, retention, internal mobility, assignment to exploratory projects, and exit after reversal.

Contract take-up is the cleanest test when workers can choose among compensation schemes, career tracks, grants, or organizational units. Exploratory workers should be more likely to choose the option with long horizons, early-failure tolerance, continuation value, and feedback, provided the relevant equilibrium separates types across contracts. Entry and retention provide a second margin: workers whose pre-entry characteristics predict exploratory comparative advantage—such as prior exploratory projects, tolerance for repeated early failure, or relatively stronger performance in exploratory than exploitative tasks—should be more likely to join or stay when contracts protect the experimentation path. Exit after reversal is especially informative because it helps distinguish the participation margin from a pure effort response. Some workers may reduce

exploratory effort when a policy is withdrawn but remain in place; others may leave because the policy had satisfied their participation constraint. The latter response is a revealed-preference test: exit indicates that the withdrawn contract features were part of what made the exploratory role attractive relative to the worker's outside option.

Assignment and organizational sorting are informative but require more care. If workers volunteer for exploratory projects or choose internal tracks, assignment can reveal self-selection. If assignment is made by the organization, it instead tests whether firms, universities, or funding agencies screen for exploratory type and allocate workers accordingly. A sharp positive test is whether exploration-friendly policies change assignments for workers with pre-policy indicators of exploratory comparative advantage, holding fixed broad measures of ability or productivity. A corresponding falsification test is whether similar flows fail to appear for workers without such exploratory indicators, or after contract changes unrelated to exploration. Without such worker-flow or assignment evidence, an observed increase in innovative output remains a joint motivation-selection effect.

Prediction 2: competition for exploratory talent changes contract design. Holding fixed the technological value of exploration within the firm, the model predicts that a more competitive outside market for exploratory talent should shift contract design toward exploration-friendly features. When workers with a comparative advantage in exploration have better outside options, firms competing for those workers should not only raise compensation; they should also use failure tolerance, long-run rewards, commitment, job security, and feedback more intensively. These features are efficient screening instruments because they are valued relatively more by explorers than by exploiters.

This comparative static is distinct from the contract-design logic in Manso (2011). In that model, the degree of exploration-friendliness is driven by the relative cost and value of exploration versus exploitation for a given agent. Here, even holding fixed the firm's return to exploration, variation in the outside market for exploratory talent changes the optimal contract. The reason is not simply that firms must pay explorers more. When outside options improve especially for explorers, the participation constraint for exploratory types tightens relative to that of exploitative types. Uniform wage increases can satisfy participation but do little to improve type separation. By contrast, failure tolerance, long-run rewards, commitment, job security, and feedback are more valuable to explorers than to exploiters because explorers face longer experimentation paths and higher option value from continuing after early setbacks. These features therefore help relax the explorer's participation constraint while preserving screening. The implied test is cross-sectional or panel-based: after conditioning on the value of exploration, proxies for outside-option competitiveness—such as labor-market thickness, regional concentration of innovative employers, mobility rates among exploratory workers, or changes in enforceability of non-compete agreements—should predict greater use of failure tolerance, long-horizon rewards, commitment, and feedback.

The main identification challenge is that the competitiveness of the market for exploratory talent is likely correlated with the technological value of exploration itself. More informative tests would use variation in outside options that is less directly tied to the firm's own return to exploration, such as changes in labor mobility, non-compete enforcement, immigration constraints, or local shocks to alternative employers, ideally within industries or with controls for local exploration opportunities. These sources of variation are not perfect: for example, non-compete reforms may also affect knowledge spillovers and local learning, while shocks to alternative employers may change both outside options and the local

value of exploration. The prediction is therefore not merely that innovative regions use more exploration-friendly contracts; it is that, conditional on the demand for exploration, stronger outside options for explorers make exploration-friendly contract features more valuable as screening devices.

The existing empirical literature provides many settings in which this motivation-selection decomposition can be partially applied. Its main limitation for the present purpose is that most studies identify the effect of an incentive or governance change on innovation outcomes, but do not separately identify whether the effect operates through motivation of incumbent workers, selection of different workers, or both. A useful way to organize the evidence is by the primary contract feature that shifts. Stock options, long-term executive compensation, public listing status, takeover defenses, board independence, analyst coverage, and short-term market pressure affect the horizon over which innovative success is rewarded and short-run failure is disciplined (Laux, 2012; Baranchuk et al., 2014; Chang et al., 2015; Gonzalez-Uribe and Groen-Xu, 2017; Ladika and Sautner, 2020; Bernstein, 2015; Cong and Howell, 2021; Zhang et al., 2023; Atanassov, 2013; Chemmanur and Tian, 2018; Bhojraj et al., 2017; Balsmeier et al., 2017; He and Tian, 2013; Guo et al., 2019). Bankruptcy codes, creditor rights, social insurance, venture capital tolerance for failure, shareholder litigation risk, labor laws, job security, tenure, financial development, and wealth shocks affect the cost of failed experimentation or workers' exposure to downside risk (Acharya and Subramanian, 2009; Mann, 2018; Farè et al., 2024; Hombert et al., 2020; Gottlieb et al., 2022; Tian and Wang, 2014; Chemmanur et al., 2014; Barrot, 2017; Lin et al., 2019; Hassan et al., 2021; Mbanyele and Muchenje, 2022; Acharya et al., 2013, 2014; Bradley et al., 2017; Bena et al., 2021; Brogaard et al., 2018; Tripodi et al., 2025; Hsu et al., 2014; Bernstein et al., 2021). Performance evaluation and feedback systems can affect whether workers learn from early outcomes and redirect experimentation rather than simply repeat a known action (Gross, 2017; Chu et al., 2019; Howell, 2021).

The funding of scientific research illustrates why empirical programs often bundle multiple contract features rather than shift a single dimension. The HHMI program studied by Azoulay et al. (2011) changed tolerance for failure, the horizon over which scientific success was rewarded, and the scope for feedback and redirection. Through the lens of the present model, such a program may affect scientific output both by motivating incumbent scientists to pursue riskier lines of inquiry and by attracting or retaining scientists whose comparative advantage is exploratory research. The selection perspective does not require assigning each empirical setting to a single contract feature. It asks, within each setting, whether the policy changes only the effort of a fixed set of workers or also the composition of workers who enter, remain in, select into the contract, or are assigned to exploratory work.

This distinction clarifies what is missing from many empirical tests. Existing work often makes substantial progress in isolating the causal effect of a policy on innovation outcomes using instruments, discontinuities, or other sources of quasi-experimental variation. But patents, citations, new products, or R&D choices usually measure the combined effect of motivation and selection. To separate the channels, empirical designs need outcome measures for worker flows, contract take-up, entry and retention, exit after policy reversals, or the composition of workers assigned to exploratory tasks, in addition to the usual innovation measures. They also need credible variation in the incentive environment, because worker flows may reflect firm quality, financial distress, career opportunities, or industry composition rather than selection induced by the contract itself. Proxies for exploratory comparative advantage could include stock measures of prior exploratory activity—such as pre-policy histories of exploratory projects, patent or publication breadth, and novelty of prior work—as well as behavioral indicators of persistence, such as willingness to continue after early failures, mobility across research areas, or pre-policy performance following setbacks. The most direct evidence

for the selection mechanism would show that exploration-friendly contracts attract and retain workers with stronger exploratory comparative advantage, and that this sorting is stronger when the outside market for such workers is more competitive.

VIII 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. The baseline explorer contract uses tolerance for early failure and backloaded rewards for long-term success to deliver rents in states that are especially valuable to explorers and less valuable to mimics. The institutional extensions show that this selection logic depends on the firm's ability to commit, its use of termination after failure, and its provision of feedback. Commitment makes long-horizon rent delivery credible; termination can discipline exploitation but often undermines exploration; and feedback preserves the option value that makes exploratory work attractive. 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, and places little weight on long-term commitment or feedback.

Finally, the model's implications suggest new avenues for empirical work. Much of the empirical literature on innovation incentives focuses on how contracts affect the behavior of a given workforce, but the analysis here suggests that incentive systems may matter just as much by changing who chooses to enter, remain in, or leave innovative environments. In the spirit of Lazear (2000), future research could ask whether failure-tolerant, long-horizon, feedback-rich organizations disproportion-

ately attract workers with stronger exploratory tendencies, and whether these workers exit when firms shorten evaluation horizons, increase termination risk, or weaken feedback and continuation opportunities. Empirically separating motivation from selection would therefore deepen our understanding of how incentive schemes shape innovation: not only by changing effort choices within the firm, but also by determining which types of workers are willing to pursue uncertain, long-term projects in the first place.

References

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

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

_ , _ , and _ , "Wrongful Discharge Laws and Innovation," *Review of Financial Studies*, 2014, *27* (1), 301–346.

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

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

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

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

Banker, Rajiv D., Seok-Young Lee, Gordon Potter, and Dhinu Srinivasan, "An empirical analysis of continuing improvements following the implementation of a performance-based compensation plan," *Journal of Accounting and Economics*, 2001, *30* (3), 315–350.

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

Barrot, Jean-Noël, "Investor Horizon and the Life Cycle of Innovative Firms: Evidence from Venture Capital," *Management Science*, 2017, *63* (9), 2773–3145.

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

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

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

Bhojraj, Sanjeev, Partha Sengupta, and Suning Zhang, "Takeover Defenses: Entrenchment and Efficiency," *Journal of Accounting and Economics*, 2017, *63* (1), 142–160.

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

Brogaard, Jonathan, Joseph Engelberg, and Edward Van Wesep, "Do Economists Swing for the Fences after Tenure?," *Journal of Economic Perspectives*, 2018, *32* (1), 179–94.

Cadsby, C. Bram, Fei Song, and Francis Tapon, "Sorting and incentive effects of pay-for-performance: an experimental investigation," *Academy of Management Journal*, 2007, *50* (2), 387–405.

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

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

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

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

Cong, Lin William and Sabrina T. Howell, "Policy Uncertainty and Innovation: Evidence from Initial Public Offerings Interventions in China," *Management Science*, 2021, *67* (11), 6629–7289.

Dohmen, Thomas and Armin Falk, "Performance pay and multidimensional sorting: productivity, preferences, and gender," *American Economic Review*, 2011, *101* (2), 556–590.

Eriksson, Tor and Marie-Claire Villeval, "Performance-pay, sorting and social motivation," *Journal of Economic Behavior & Organization*, 2008, *68* (2), 412–421.

Farè, Marie, Marcus Dejardin, and Eric Toulemonde, "Creditor Rights and Innovation: International Evidence," *Journal of Economic Behavior & Organization*, 2024, *216* (1), 356–375.

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

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

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

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

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

Hassan, M. Kabir, Reza Houston, and M. Sydul Karim, "Courting Innovation: The Effects of Litigation Risk on Corporate Innovation," *Journal of Corporate Finance*, 2021, *71*, 102098.

He, Jie and Xuan Tian, "The Dark Side of Analyst Coverage: The Case of Innovation," *Journal of Financial Economics*, 2013, *109* (4), 856–878.

Holmstrom, Bengt, "Agency costs and innovation," *Journal of Economic Behavior & Organization*, 1989, *12* (3), 305–327.

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

Howell, Sabrina, "Learning from Feedback: Evidence from New Ventures," *Review of Finance*, 2021, *25* (3), 595–627.

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

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

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

Lazear, Edward P., "Salaries and piece rates," *Journal of Business*, 1986, *59* (3), 405–431.

_ , "Performance Pay and Productivity," *American Economic Review*, 2000, *90* (5), 1346–1361.

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

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

Manso, Gustavo, "Motivating innovation," *Journal of Finance*, 2011, *66* (6), 1823–1860.

March, James G., "Exploration and exploitation in organizational learning," *Organization Science*, 1991, *2* (1), 71–87.

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

O'Reilly, Charles A. and Michael L. Tushman, "Organizational ambidexterity: past, present, and future," *Academy of Management Perspectives*, 2013, *27* (4), 324–338.

Salop, Joanne and Steven C. Salop, "Self-Selection and Turnover in the Labor Market," *Quarterly Journal of Economics*, 1976, *90* (4), 619–627.

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

Tripodi, Giorgio, Xiang Zheng, Yifan Qian, Dakota Murray, Benjamin F. Jones, Chaoqun Ni, and Dashun Wang, "Tenure, Security, and Scientific Innovation," Working Paper 30 2025.

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

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

A Two-Principal Formulation and Decoupling

A 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 principal's problem.

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 single-principal problem.*

The proof 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.

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. □

B Optimality of the Three-Step Construction

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.

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 relaxed minimization 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 optimal baseline contract that implements π for type θ without (RAC) and (PC). 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 reallocat-

ing 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 after Steps 2 the contract 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. □

C Proofs of Optimal Screening Contracts

A Proof of the Exploiter Contract

Proof of Proposition 1. Fix the target action plan $\langle {}_1{}^1_1 \rangle$. 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:

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

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

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

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

Solving these gives $\tilde{w}_S = \tilde{w}_{SS} = \tilde{w}_{FS} = \frac{c}{p_1}$ and $\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0$. Equivalently, the principal pays a constant bonus c/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 exerting explorer's rent without lowering cost, because success in period 1 is more diagnostic than success after success. (indeed $\Lambda_{\langle {}_1{}^1_1\rangle,\langle {}_2{}^2_2\rangle}(S) > \Lambda_{\langle {}_1{}^1_1\rangle,\langle {}_2{}^2_2\rangle}(SS)$). 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. Consider a misreporting explorer who accepts the baseline exploiter contract $\tilde{w}^{X,0}$. Under $\tilde{w}^{X,0}$, each success node pays the bonus c/p_1, each failure node pays zero above the floor, and the exploiter who follows $\langle {}_1{}^1_1\rangle$ obtains exactly $2\underline{w}$.

We first rule out shirking by the misreporting explorer. The relevant shirking margin is the first-period decision. To see this, consider the continuation behavior after each first-period history. If the explorer works in period 1 and succeeds, the continuation success probability is

$E[p_2 \mid S, 2] > p_1$, so the baseline success bonus c/p_1 makes second-period work strictly optimal. If the explorer works in period 1 and fails, the fallback exploratory technology has success probability $E[p_2 \mid F, 2] = p_1$, so the same bonus makes second-period work weakly optimal. Thus a misreporting explorer who works in period 1 also works in period 2 after either history.

By contrast, if the explorer shirks in period 1, then $p_0 = 0$ implies first-period failure for sure and no success-contingent bonus is paid. The first-period shirker reaches the failure continuation without having generated the productive exploratory history. At that node, shirking again yields the floor, while work only breaks even at the baseline bonus. Hence the shirking continuation can be taken to deliver the floor, and the first-period shirking path yields the two-period floor payoff $2\underline{w}$.

If instead the explorer works in period 1, the induced continuation plan is the non-shirking exploratory plan $\langle {}_{2}{}^{2}_{2} \rangle$. Her payoff under the baseline exploiter contract is

$$U_E(\tilde{w}^{X,0}, \langle {}_{2}{}^{2}_{2} \rangle) = W(\tilde{w}^{X,0}, \langle {}_{2}{}^{2}_{2} \rangle) - C(\langle {}_{2}{}^{2}_{2} \rangle).$$

The baseline exploiter contract pays the same bonus c/p_1 at each success node. Since effort costs are the same for exploration and exploitation, the maintained surplus-dominance assumption implies that the explorer's expected success-bonus receipts under $\tilde{w}^{X,0}$ exceed those of the exploiter by more than any difference in effort costs. Hence

$$W(\tilde{w}^{X,0}, \langle {}_{2}{}^{2}_{2} \rangle) - C(\langle {}_{2}{}^{2}_{2} \rangle) > W(\tilde{w}^{X,0}, \langle {}_{1}{}^{1}_{1} \rangle) - C(\langle {}_{1}{}^{1}_{1} \rangle) = 2\underline{w},$$

where the equality is the exploiter's baseline participation payoff. Thus first-period work yields strictly more than $2\underline{w}$, while first-period shirking yields only $2\underline{w}$. A misreporting explorer therefore exerts rather than shirks under the baseline exploiter contract.

The relevant comparison for surplus delivery is consequently between the exploiter plan $\langle {}_1{}^1_1 \rangle$ and the non-shirking explorer plan $\langle {}_2{}^2_2 \rangle$. For that comparison, the first-period success node is the most diagnostic outcome: $\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(S) > \Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(SS)$ is equivalent to $E[p_2 \mid S, 2] > p_1$, and $\Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(S) > \Lambda_{\langle {}_1{}^1_1 \rangle, \langle {}_2{}^2_2 \rangle}(FS)$ is equivalent to $p_1 > E[p_2]$. Both inequalities hold by the exploratory ordering. Hence any required surplus Δ is optimally delivered by raising $\tilde{w}_S$ only, while keeping $\tilde{w}_{SS}$ and $\tilde{w}_{FS}$ at their incentive-minimal levels.

Combining the baseline contract construction, the impossibility of relaxing the ride–along constraint via the exploiter's wage schedule, and the likelihood-ratio ranking for surplus delivery completes the proof. □

B Proof of the Explorer Contract

Proof of Proposition 2. Fix the target plan $a_{\mathrm{E}} = \langle {}_2{}^2_2 \rangle$. For the baseline contract, failure payoffs do not help induce exploration and can be set to zero: $\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0$. The second-period ICs decouple cleanly.

(i) *After F:* to induce exploration in period 2, use $E[p_2 \mid F, 2] = p_1$:

$$p_1 \, \tilde{w}_{FS} \;\geq\; c \quad \Rightarrow \quad \tilde{w}_{FS} = \frac{c}{p_1}.$$

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

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

Period 1 exploration requires that choosing 2 in period 1 beats the relevant deviations. Relative to $\langle {}_0{}^2_2 \rangle$, the first-period exploration IC implies

$$E[p_2]E[p_2 \mid S, 2]\tilde{w}_{SS} - (1 - E[p_2])p_1\tilde{w}_{FS} \geq c - E[p_2]c - (1 - E[p_2])c,$$

or, using $\tilde{w}_{FS} = c/p_1$,

$$\tilde{w}_{SS} \geq \frac{(1+p_1)c}{p_1 E[p_2 \mid S, 2]}.$$

Relative to $\langle {}_{0}^{0}{}_{0}^{0} \rangle$, the expected bonus under $\langle {}_{2}^{2}{}_{2}^{2} \rangle$ must cover the two periods of effort cost:

$$E[p_2]E[p_2 \mid S, 2]\tilde{w}_{SS} + (1 - E[p_2])p_1\tilde{w}_{FS} \geq 2c.$$

Substituting $\tilde{w}_{FS} = c/p_1$ gives

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

Because $E[p_2] < p_1$, this all-shirking constraint is tighter than the $\langle {}_{0}^{2}{}_{2}^{2} \rangle$ constraint. The minimal bonuses are therefore

$$\tilde{w}_{FS}^{0} = \frac{c}{p_1}, \qquad \tilde{w}_{SS}^{0} = \frac{(1+E[p_2])c}{E[p_2]E[p_2 \mid S, 2]}.$$

Finally, $\tilde{w}_S$ does not tighten the binding IC at these primitives and can be kept at zero. This gives the stated baseline utility $2\underline{w}$.

A misreporting exploiter's payoff on the explorer contract is (weakly) increasing in $\tilde{w}_{SS}$ and $\tilde{w}_{FS}$. Both stand at their IC-minimal levels for $\langle {}_{2}^{2}{}_{2}^{2} \rangle$. In particular, $\tilde{w}_{FS} = c/p_1$ makes the exploiter just indifferent about exerting after a first-period failure, so raising $\tilde{w}_{FS}$ would create pure leakage. Any attempt to "poison" by lowering $\tilde{w}_{SS}$ and compensating with $\tilde{w}_S$ or $\tilde{w}_{FS}$ 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 $\Delta = \overline{U}_{\mathrm{E}} - 2\underline{w} > 0$. The recovery bonus remains fixed at $\tilde{w}_{FS} = c/p_1$. Under the maintained assumptions, the baseline explorer contract makes a misreporting exploiter shirk in the

first period. Indeed, the exploiter's first-period IC at the baseline is

$$p_1^2 \left(\tilde{w}_{SS}^0 - \tilde{w}_{FS}^0\right) \geq c,$$

and substituting the baseline bonuses gives

$$\frac{p_1^2(1 + E[p_2])}{E[p_2]E[p_2 \mid S, 2]} \geq 1 + p_1.$$

The surplus-dominance condition implies $E[p_2](1+E[p_2 \mid S,2]-p_1) > p_1$, and together with $p_1 < E[p_2 \mid S, 2]$ this inequality is strictly reversed. Thus the first step of surplus delivery must repair this shirking margin by raising $\tilde{w}_{SS}$ until the misreporting exploiter's first-period IC binds. By the LR principle comparing $\langle {}_2{}_{\frac{2}{2}} \rangle$ to $\langle {}_1{}_{\frac{1}{1}} \rangle$, the two informative surplus-delivery directions after this repair are SS and, if needed, F. Define *Step A (repair the misreporting exploiter's first-period IC).* Raise $\tilde{w}_{SS}$ until first-period exertion is just optimal for the misreporting exploiter. This is the relevant tightening margin: after a first-period failure, $\tilde{w}_{FS} = c/p_1$ already makes the exploiter indifferent between exerting and shirking, and raising $\tilde{w}_{FS}$ would only increase leakage.

Step B (deliver surplus without reopening shirking). Deliver any remaining Δ along the least-leaky direction that preserves the repaired exploiter IC. Compare likelihood ratios with the explorer plan in the numerator, $\Lambda_{SS} = \Lambda_{\langle {}_2{}_{\frac{2}{2}} \rangle, \langle {}_1{}_{\frac{1}{1}} \rangle}(SS)$ and $\Lambda_F = \Lambda_{\langle {}_2{}_{\frac{2}{2}} \rangle, \langle {}_1{}_{\frac{1}{1}} \rangle}(F)$. If $\Lambda_{SS} \geq \Lambda_F$, double success is more diagnostic of exploration, so raise $\tilde{w}_{SS}$ alone; this relaxes the misreporting exploiter's first-period IC. If $\Lambda_{SS} < \Lambda_F$, failure is more diagnostic and the principal increases $(\tilde{w}_F, \tilde{w}_{SS})$ along the misreporting exploiter's IC-indifference ray,

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

The explorer IC arc would instead require $d\tilde{w}_F = E[p_2 \mid S, 2] d\tilde{w}_{SS}$; since $E[p_2 \mid S, 2] > p_1$, moving in that direction after Step A would make

shirking strictly attractive for the misreporting exploiter. The exploiter IC ray is therefore the relevant frontier after the repair step.

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

D Proofs for Extensions

A Lack of Commitment

A.1 Exploitation without commitment

Proof of Proposition 3. 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 optimal baseline contract that implements $\langle {}_{1}^{1}{}_{1}^{1} \rangle$ 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, Proposition 1 shows that the optimal committed contract raises only the first-period success payment $\tilde{w}_S$. This transfer is paid immediately after the period–1 outcome and is therefore also implementable without commitment. The second-period bonuses $\tilde{w}_{SS}$ and $\tilde{w}_{FS}$ remain at their incentive-minimal levels c_1/p_1 and are not used for surplus delivery. Thus there is

no deferred surplus promise for the principal to renegotiate away in the exploiter contract. The sequence of short-term contracts replicates both the baseline contract and the surplus-delivery region of the committed optimum, so lack of commitment has no effect on the optimal exploiter contract. □

A.2 Exploration without commitment

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:

$$\text{after } F: \quad p_1\, \tilde{w}_{FS} = c \;\Rightarrow\; \tilde{w}^0_{FS} = \frac{c}{p_1}, \qquad \text{after } S: \quad E[p_2 \mid S, 2]\, \tilde{w}_{SS} = c \;\Rightarrow\; \tilde{w}^0_{SS} = \frac{c}{E[p}$$

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 $\underline{w}$ (the period–2 bonus just covers

c). Hence the period–1 exploration IC depends only on period–1 wages:

$$\underbrace{p_2(w_S - \underline{w}) + (1 - p_2)(w_F - \underline{w}) - c}_{\text{choose 2 in period 1}} \;\geq\; \underbrace{(w_F - \underline{w})}_{\text{choose 0 in period 1}}$$

$$\iff \quad p_2\,(w_S - w_F) \;\geq\; c.$$

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

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

Collecting terms yields the stated baseline schedule:

$$\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0, \qquad \tilde{w}_{FS}^0 = \frac{c}{p_1}, \qquad \tilde{w}_{SS}^0 = \frac{c}{E[p_2 \mid S, 2]}, \qquad \tilde{w}_S^0 = \frac{c}{E[p_2]},$$

which is implementable via a sequence of one–period contracts.

Suppose the participation constraint binds and set $\Delta = \overline{U}_{\mathrm{E}} - 2\underline{w} > 0$. A misreporting exploiter's continuation value is zero after either first-period history at the baseline continuation contracts. After F, he can work with success probability p_1 against $\tilde{w}_{FS} = c/p_1$, which exactly covers effort cost; shirking also yields zero. After S, working against $\tilde{w}_{SS} = c/E[p_2 \mid S, 2]$ yields $p_1 c/E[p_2 \mid S, 2] - c < 0$ because $p_1 < E[p_2 \mid S, 2]$, so he shirks and obtains zero. Thus first-period shirking yields

$$U_X(0) = \tilde{w}_F + 0 = 0,$$

whereas first-period exertion yields

$$U_X(1) = p_1 \tilde{w}_S + (1 - p_1)\tilde{w}_F - c = \frac{p_1 c}{E[p_2]} - c = \frac{c(p_1 - E[p_2])}{E[p_2]} > 0,$$

using $E[p_2] < p_1$. Hence the misreporting exploiter strictly prefers to exert; the shirking case is vacuous.

The principal therefore delivers any additional surplus by moving (w_S, w_F) in parallel, $dw_S = dw_F$, so that

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

keeps the explorer's first-period IC tight. This parallel move delivers additional surplus while preserving the renegotiation-proof period-2 implementation. 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 leak more rent to the exerting exploiter. □

B Termination

B.1 Exploiter contract with termination

Proof of Proposition 5. With termination, only the S-branch continues to period 2. The period–2 IC is the standard one–period effort condition:

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

Hence the continuation value above the floor after S is

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

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 $p_1(\tilde{w}_S + \tilde{V}_2) - c = p_1\tilde{w}_S - c$. Adding the floors,

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

Thus the least–cost choice is

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

At this choice the first-period IC binds, so the exerting agent receives the same payoff as a first-period shirker, namely the outside/minimum payoff $\underline{w}$. Thus the terminated baseline rent is $\underline{w}$. 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 $\underline{w} \geq \alpha$, no above–floor first–period bonus is needed.

If the exploiter's participation binds, let $\Delta > 0$ be the extra rent to deliver. Consider a misreporting explorer on the terminated baseline exploiter contract $\tilde{w}^{X,T,0}$. We first rule out shirking by the misreporting explorer. With termination, the relevant shirking margin is again the first-period decision. If the explorer works in period 1 and succeeds, the only surviving continuation branch is reached. On that branch, second-period work yields

$$E[p_2 \mid S, 2]\,\frac{c}{p_1} - c > 0,$$

because $E[p_2 \mid S, 2] > p_1$. Thus a misreporting explorer who works in period 1 and succeeds also works in period 2. If the explorer works in period 1 and fails, the relationship is terminated, so there is no failure-continuation shirking margin.

By contrast, if the explorer shirks in period 1, then $p_0 = 0$ implies first-period failure for sure. Termination follows, no success-contingent bonus is paid, and the shirking path yields only the first-period floor payoff. Hence the first-period shirking path is dominated by first-period work whenever the explorer's payoff from the induced non-shirking terminated plan $\langle {}_2{}^2_t \rangle$ exceeds the shirking payoff.

Under the maintained surplus-dominance assumption and equal effort costs, the non-shirking terminated exploratory plan is the relevant misre-

porting plan. Its payoff under the terminated baseline exploiter contract is

$$U_E(\tilde{w}^{X,T,0}, \langle {}_2{}^2_t \rangle) = W(\tilde{w}^{X,T,0}, \langle {}_2{}^2_t \rangle) - C(\langle {}_2{}^2_t \rangle),$$

which exceeds the payoff from first-period shirking and termination. Therefore the misreporting explorer exerts in period 1 and, conditional on success, exerts again in period 2. There is no SS shirking margin that requires raising $\tilde{w}_{SS}$ above c/p_1.

The relevant leakage comparison is consequently between the terminated exploiter plan $\langle {}_1{}^1_t \rangle$ and the non-shirking terminated explorer plan $\langle {}_2{}^2_t \rangle$. Increasing $\tilde{w}_S$ both delivers surplus to the target exploiter and weakly relaxes the misreporting explorer's first-period exertion constraint. Among the remaining nodes, S is also the most diagnostic surplus-delivery direction for the comparison between $\langle {}_1{}^1_t \rangle$ and $\langle {}_2{}^2_t \rangle$. Therefore any additional surplus is optimally delivered by increasing $\tilde{w}_S$ alone, keeping $\tilde{w}_{SS} = c/p_1$ fixed.

This establishes $\tilde{w}_{SS} = c/p_1$ and $\tilde{w}_S = \max\{c/p_1 - \underline{w}, 0\}$ at baseline, the irrelevance of an SS shirking repair under $E[p_2 \mid S, 2] > p_1$, and the surplus-delivery rule under termination. □

Proof of Corollary 2. We prove the two claims in turn.

Part (i). Termination removes the continuation following a first-period failure. Under the maintained ordering $E[p_2] < p_1$, the exploiter reaches this branch with probability $1 - p_1$, while a misreporting explorer who exerts in period 1 reaches it with probability $1 - E[p_2]$. Since $1 - E[p_2] > 1 - p_1$, termination removes the failure-branch continuation more often for the explorer. Under the current exploiter contract, no additional surplus is parked on the failure branch: the no-termination recovery payment is kept at its incentive-minimal level c/p_1, and the terminated contract sets the failure branch to zero. Thus the branch that termination removes is

hit with strictly higher probability by the misreporting explorer than by the exploiter. Termination therefore reduces the misreporting explorer's value by more than the exploiter's value, proving item (i).

Part (ii). Suppose the exploiter's participation constraint binds under termination and the principal must restore additional surplus $\Delta > 0$. Proposition 5 shows that the surplus is delivered by increasing the first-period success payment $\tilde{w}_S$ alone. Raising $\tilde{w}_S$ by Δ/p_1 gives the exploiter exactly Δ in expected utility. The relevant misreporting explorer exerts in period 1 and succeeds with probability $E[p_2]$, so the same wage increase raises her expected utility by

$$E[p_2]\frac{\Delta}{p_1} = \frac{E[p_2]}{p_1}\Delta < \Delta,$$

using $E[p_2] < p_1$. Thus the principal can restore the exploiter's surplus while giving strictly less back to the misreporting explorer, proving item (ii). □

B.2 Explorer contract with termination

Proof of Proposition 6. Because continuation happens only after S, any payment on the failure branch cannot help incentives and is set at zero: $\tilde{w}_F = \tilde{w}_{SF} = \tilde{w}_{FF} = 0$. Likewise, $\tilde{w}_S = 0$: shift rewards to $\tilde{w}_{SS}$ to reduce leakage to misreporting exploiter.

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

$$E[p_2 \mid S, 2]\ \tilde{w}_{SS} \ \geq\ c_2 \quad \Rightarrow \quad \tilde{w}_{SS} \ \geq\ \frac{c_2}{E[p_2 \mid S, 2]}. \qquad (\mathrm{IC}_{\langle {}_2{}^0_t \rangle})$$

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 $t=2$):

$$E[p_2]E[p_2 \mid S,2]\ \tilde{w}_{SS} \ \geq\ (1+E[p_2])c_2 - E[p_2]\underline{w}. \qquad (\mathrm{IC}_{\langle {}_0{}^2_t \rangle})$$

(ii) $\langle {}_0{}^0_t \rangle$ (shirk in both periods):

$$E[p_2]E[p_2 \mid S,2]\ \tilde{w}_{SS} \ \geq\ (1+E[p_2])c_2 - E[p_2]\underline{w}. \qquad (\mathrm{IC}_{\langle {}_0{}^0_t \rangle})$$

Collecting the constraints, the minimal success bonus is

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

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

Suppose the explorer's participation binds and define $\Delta = \overline{U}_{\mathrm{E}} - (1 + E[p_2])\underline{w} > 0$.

Step A (repair the misreporting exploiter's first-period IC). Increase $\tilde{w}_{SS}$ until the misreporting exploiter is just indifferent between working and shirking in period 1. With $p_0 = 0$, the SS node has infinite likelihood ratio against first-period shirking, so loading SS first is the least costly way to restore exertion while preserving $\langle {}_2{}^2_t \rangle$.

Step B (choose the least-leaky direction consistent with that repair). Compare the likelihood ratios of exploration relative to exploitation at the two informative nodes:

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

If $\Lambda_{SS} \geq \Lambda_F$ (double success more diagnostic of exploration), raise $\tilde{w}_{SS}$ alone; this relaxes the repaired exploiter IC. If instead $\Lambda_{SS} < \Lambda_F$ (failure more diagnostic), share surplus between F and SS along the misreporting exploiter's IC frontier,

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

Moving along the explorer IC would require $d\tilde{w}_F = E[p_2 \mid S, 2]d\tilde{w}_{SS}$; because $E[p_2 \mid S, 2] > p_1$, that direction would reduce the misreporting exploiter's IC slack after Step A and make shirking strictly attractive. The exploiter IC frontier is therefore the relevant post-repair arc. □

Proof of Corollary 3. We mirror the proof of Corollary 2, using the likelihood-ratio comparison from Proposition 6.

Part (i). Conditional on reaching F, termination removes the same continuation payoff $\underline{w}$ from a truthful explorer and from a misreporting exploiter. The truthful explorer reaches F with probability $1 - E[p_2]$, so his expected loss is $(1 - E[p_2])\underline{w}$. An exerting misreporting exploiter reaches F with probability $1 - p_1$, so his expected loss is $(1 - p_1)\underline{w}$. Under the maintained ordering $E[p_2] < p_1$,

$$(1 - E[p_2])\underline{w} > (1 - p_1)\underline{w}.$$

Termination therefore reduces the value of the explorer contract more for the truthful explorer than for the exerting misreporting exploiter. This proves part (i).

Part (ii). Now suppose the explorer's lost value $(1 - E[p_2])\underline{w}$ must be restored after termination and that the misreporting exploiter exerts. A unit of utility delivered to the explorer through SS leaks to the exploiter at rate

$$\frac{1}{\Lambda_{SS}} = \frac{p_1^2}{E[p_2]E[p_2|S, 2]},$$

where

$$\Lambda_{SS} = \frac{E[p_2]E[p_2|S, 2]}{p_1^2}.$$

A unit delivered through F leaks at rate

$$\frac{1}{\Lambda_F} = \frac{1-p_1}{1-E[p_2]},$$

where

$$\Lambda_F = \frac{1-E[p_2]}{1-p_1}.$$

These are the likelihood ratios in Proposition 6, now written with the explorer plan in the numerator.

In the failure-diagnostic region, $\Lambda_{SS} < \Lambda_F$, and restoring the explorer's rent uses F together with compensating increases in SS along the IC arcs identified in Proposition 6. Thus the average leakage rate from restored explorer utility to the exerting misreporting exploiter is

$$\bar{\Lambda} = (1-\alpha)\Lambda_F + \alpha\Lambda_{SS}$$

for some $\alpha > 0$, where α is the share of restored explorer utility delivered through the compensating increase in SS. Since $1/\Lambda_{SS} > 1/\Lambda_F$, it follows that

$$\bar{\Lambda} > 1/\Lambda_F.$$

Restoring the explorer's lost value therefore gives the exerting misreporting exploiter more than

$$\frac{1}{\Lambda_F}(1-E[p_2])\underline{w} = (1-p_1)\underline{w},$$

which is exactly the value he lost from termination. Thus the restoration raises the exerting misreporting exploiter's payoff above its no-termination level. This proves part (ii). □

C Feedback

C.1 Exploiter contract without feedback

Proof of Proposition 7. To implement the exploiter's action plan $\langle {}_1{}^1_1 \rangle$ 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 while maintaining incentives. Similarly, any reward for total failure can be eliminated. Paying $\tilde{w}_{FF} > 0$ (after two failures) does nothing to encourage effort or truth-telling, so in a cost-minimizing contract we set $\tilde{w}_{FF} = 0$. 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 $\mathrm{IC}_{\langle {}_j^i \rangle}$ for $i, j \in \{0, 1\}$. It suffices to enforce the following three IC constraints, which cover the most profitable deviations:

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

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

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

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

- $\mathrm{IC}_{\langle {}^{0}_{1} {}^{1}\rangle}$ (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 $\mathrm{IC}_{\langle {}^{0}_{0} {}^{0}\rangle}$ 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 explorer'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:

$$\Lambda_{a_X,a_E}(SS) = \frac{p_1^2}{E[p_2]\, E[p_2 \mid S, 2]},$$

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

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.

□

C.2 Explorer contract without feedback

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

$$\tilde{w}_S = \tilde{w}_F, \quad \tilde{w}_{SF} = \tilde{w}_{FS},$$

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

$$U_{\mathrm{E}}^{\langle {}_2{}^2_2\rangle} = E[p_2]E[p_2 \mid S, 2]W - 2c.$$

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

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

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

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

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:

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

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

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

With feedback, the principal can implement plan $\langle {}_2{}^2_2 \rangle$, resetting the probability of success after failure to $E[p_2 \mid F, 2] = p_1$, using the exact same contract as in Proposition 2.

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