Feedback Design in Dynamic Moral Hazard*

Jeffrey C. Ely † George Georgiadis ‡ Luis Rayo §

February 15, 2024

Abstract

We study the joint design of dynamic incentives and performance feedback for an environment with a coarse (all-or-nothing) measure of performance, and show that hiding information from the agent can be an optimal way to motivate effort. Using a novel approach to incentive compatibility, we derive a two-phase solution that begins with a "silent phase" where the agent is given no feedback and is asked to work non-stop, and ends with a "full-transparency phase" where the agent stops working as soon as a performance threshold is met. Hiding information leads to greater effort, but an ignorant agent is also more expensive to motivate. The two-phase solution—where the agent's ignorance is fully frontloaded—stems from a "backward compounding effect" that raises the cost of hiding information as time passes.

*We are grateful to Sandeep Baliga, Matteo Camboni, Yeon-Koo Che, Theo Durandard, Matteo Escudé, Drew Fudenberg, Alex Frankel, Ben Friedrich, Michael Gibbs, Johannes Hörner, Sofia Moroni, Zvika Neeman, Malesh Pai, Alessandro Pavan, Michael Powell, Anna Sanktjohanser, Roland Strauss, Juuso Toikka, Alex Wolitzky, as well as to participants at various seminars and conferences for helpful comments.

†Charles E. and Emma H. Morrison Professor of Economics, Northwestern University. jeff@jeffely.com

‡Associate Professor of Strategy, Northwestern Kellogg. g-georgiadis@kellogg.northwestern.edu

§Erwin P. Nemmers Professor of Strategy, Northwestern Kellogg. luis.rayo@kellogg.northwestern.edu

1 Introduction

A key design component of many jobs, such as those in professional services firms, is the performance feedback offered to employees, as it allows them to adjust their behaviors and learn what their future rewards might look like. While some experts argue that a policy of full transparency is best—that is, keeping employees fully apprised of their performance—such practice is far form uniform as employers may see a strategic gain from hiding information or postponing its release; e.g., Maister (1993).

Here we study the optimal joint design of performance feedback and monetary rewards in a dynamic-agency environment where the underlying monitoring technology is coarse. We argue that because of this coarseness, concealing information from the agent about their performance can be a profitable way to increase their effort, despite the fact that doing so requires paying them greater rewards.

In our baseline model, the principal wishes to maximize the agent's aggregate effort net of monetary rewards, and must discourage strategic pauses by the agent. The only performance measure available to the principal is an all-or-nothing signal of effort. The resulting problem is challenging because the (dynamic) feedback policy can in principle be highly complex, and interact in complicated ways with the chosen monetary rewards.

To solve it, we show that it suffices to restrict attention to policies that discourage instantaneous effort pauses by the agent—akin to a first-order approach. This allows us to show that the optimal contract consists of two phases, with the agent first kept fully in the dark and then kept fully apprised of their performance. This bang-bang solution arises from a desire to keep the agent working beyond the time when the performance measure records a success in combination with a "backward-compounding effect," which implies that hiding information is more costly if that hiding occurs farther in the future, and hence it is optimal to frontload the agent's ignorance.

Our finding that concealing information early on can be profitable for the principal may speak to the early stages of professional services jobs, where the firm's superiors sometimes keep performance information hidden from their associates to artificially extend the initial trial period of their careers. In tandem with that lack of transparency, associates are granted a finite time to secure promotion and may earn larger rewards the sooner they are promoted. These features have analogs in our model as well. (We expand on this discussion in Section 4.)

We also consider various extensions where a two-stage contract remains optimal. These concern, respectively, the possibility that players secure additional payoffs from a continued relationship, that the agent is able to directly learn about a success, and that the output technology features a "learning-by-doing" component. In all these cases, backward compounding again leads to a bang-bang solution that frontloads the agent's ignorance.

Finally, we consider an extension where the agent is able to succeed an unlimited number of times, rather than only once, but where the principal observes those successes only if they engage in costly monitoring. Here we show that provided monitoring is sufficiently costly, withholding information by means of repeated silent phases—while monitoring the agent less—is better than full disclosure. This finding points to the potential benefits of adopting a worse (but cheaper) effort signal, and coupling it with less information for the agent.

Related literature. We contribute first and foremost to the literature on dynamic agency models under moral hazard; see Sannikov (2008), and for an overview, Georgiadis (2022). Canonical models assume that incremental output at each instant (or each period) can take many values, whereas our work is closer to models where it is binary. In Mason and Välimäki (2015), for example, a principal designs a contract to motivate a Poisson "breakthrough" and in Green and Taylor (2016) two breakthroughs are required. Halac, Kartik and Liu (2016) consider a setting where players learn about the feasibility of a breakthrough—which, in our language, is equivalent to a declining hazard rate—and the agent is privately informed about their ability.[1] These models, in contrast to ours, do not

[1] Keller, Rady and Cripps (2005) and Bonatti and Horner (2011) analyze the equilibria of such good-news

permit the principal to strategically provide feedback to the agent.

We also contribute to the literature on information design. Rayo and Segal (2010) and Kamenica and Gentzkow (2011) study the optimal provision of information in static environments, and Ely (2017) and Renault, Solan and Vieille (2017) extend these analyses to dynamic settings. The latter two papers consider a game between a receiver (e.g., an investor) and a sender (e.g., an advisor) where the sender observes a payoff-relevant state variable that evolves exogenously, and chooses a message policy to entice the receiver to take an action. Here the optimal policy is effectively a static one, treating the receiver as if they were myopic.[2]

Ely and Szydlowski (2020), Orlov, Skrzypacz and Zryumov (2020), Ball (2022), Kaya (2022), and Smolin (2021) consider games between a principal (sender) and an agent (receiver) in which the agent decides when to stop supplying an action, while the principal monitors the evolution of a payoff-relevant state (which is independent of the action) and transmits messages to entice the agent not to stop. In all of these papers the agent's action is observable (so there is no moral hazard) and the agent's monetary rewards are exogenously given or severely restricted.[3] These papers therefore exclusively focus on the design of feedback whereas we study the interaction between feedback and monetary rewards and the optimal joint design of these, and do so in a moral-hazard environment where the agent can secretly manipulate that feedback.[4]

In a career-concerns framework, Hörner and Lambert (2021) study the design of information provided to an external market to motivate the agent. Similarly, in Ostrizek (2022) all information is public and the focus is on the tradeoff between precise information, which enables higher-powered monetary incentives, and coarse information, which

Poisson experimentation models.

[2]Hörner and Skrzypacz (2016) discuss conditions under which the optimal policy is not static.

[3]In Smolin (2021) the reward is a function of the feedback provided but this function is exogenous and not part of the design. In Kaya (2022) the principal pays a share of profit that is fixed and not responsive to the agent's actions.

[4]In a setting with multiple agents, Ely et al. (2022) also permits the principal to flexibly choose the agents' reward schedule. However, characterizing the optimal contest hinges on having sufficiently many agents so that full rent-extraction is possible.

slows down learning and reduces future agency rents.[5]

2 Baseline Model

We begin with the simplest version of the model, which we then extend in various directions. At each instant t and up to an exogenous deadline $\overline{T}$, an agent privately chooses whether to spend effort toward producing a binary signal, which we call a "success." We take $\overline{T}$ to be large but finite. Success, which occurs at most once, is observed only by a principal and arrives stochastically as a function of the agent's accumulated effort. The principal, who enjoys commitment power, designs monetary rewards together with a real-time feedback policy on the basis of that signal, with the goal of maximizing the agent's total effort net of monetary payments. Both players are risk neutral and the agent is cash-constrained.[6]

At each moment the agent either works or waits—that is, effort is binary—and when working, incurs a flow cost $c < 1$. The agent's probability of success at or prior to any point in time is given by the function $F : [0, \overline{T}] \to [0, 1]$ of the cumulative time e that the agent has spent working. $F(\cdot)$ is weakly increasing and satisfies $F(0) = 0$ and $F(\overline{T}) \leq 1$. A useful equivalent representation is to imagine that a hidden random *effort requirement* $z \in [0, \overline{T}]$ is drawn according to the c.d.f. F and success occurs as soon as the agent's accumulated effort e reaches z.

We will assume that F has a differentiable density f and that both $f(e)$ and the hazard rate $\lambda(e) := f(e)/(1 - F(e))$ are weakly decreasing in e. Some of our results will also

[5] Also related is the contracting literature where the principal acquires costly information about the agent's effort(s); see, for example, Georgiadis and Szentes (2020) and Orlov (2022). In these models, information matters only indirectly via its impact on the players' payments.

[6] While assuming a single success is stylized, our goal is to capture the notion that in reality performance measures may be coarse owing to monitoring costs. More on this in Section 5.

make use of the following function

$$\Phi(t) := F(t)\,\frac{d}{dt}\frac{1}{f(t)}, \tag{1}$$

whose significance will soon be clear. The problem of maximizing expected effort minus expected rewards will be well-behaved when Φ is weakly increasing, which we will assume. These assumptions are satisfied if, for example, F arises from good-news Poisson experimentation (as studied for instance by Halac, Kartik and Liu, 2016) where the project can be "good" or "bad" (or the agent's ability "high" or "low"), unknown to both players, and a success arrives with constant hazard rate only if the project is good and the agent is working. A special case is that of a constant hazard rate.

The principal offers the agent a monetary reward $R(t)$ if success arrives at t, and no reward (or punishment) when it never arrives.[7] So that the reward does not reveal any undesired information, we assume (without loss) that it is paid at $\overline{T}$. In addition, the principal designs a feedback policy that specifies, for each point in time, a probability distribution over messages as a function of past messages and the time of past success, if it happened. By standard arguments it is sufficient to consider direct feedback policies in which the principal recommends to the agent whether to work or wait at each instant. We shall be sure to discuss how the resulting effort recommendation policies can be implemented using a feedback policy about the agent's success.

Because pauses in effort have no impact on the players' payoffs and merely waste time, it is without loss for the principal to recommend (possibly stochastically) that the agent either continue working and await further recommendations or permanently quit, rather than temporarily pause. Such a recommendation policy can be represented by a function

[7] Restricting to rewards that condition on time of success alone is without loss because, owing to the agent's risk-neutrality, a reward schedule $R(t,x)$ that conditions on a second random variable x, such as past feedback, can be replaced by a reward function equal to $\mathbb{E}_x R(t,x)$ (where the expectation is taken at time 0) without altering the ex-ante incentive constraint, and doing so prevents adding further incentive constraints in the future. Furthermore, rewarding the agent in the absence of a success merely hinders incentives to work, and imposing a punishment if a success never arrives is infeasible due to the agent being cash constrained.

$q(s|t)$ denoting the probability that the agent is still asked to work at date s conditional on having succeeded at $t \leq s$, together with a function $r(s)$ denoting the probability that the agent is still asked to work at date s conditional on *not* having succeeded before then. Note that since a recommendation to quit is a permanent one, both recommendation functions must be non-increasing in s (and so that the players' expected payoffs are well defined, we assume that they are integrable in all arguments).

To simplify notation, we will assume throughout the main text that $r(s) = 1$ up to some (endogenous) deterministic time T, and equals zero thereafter; that is, an unsuccessful agent is never asked to stop before T. (We still allow for a general r in all relevant proofs in the appendix.) Doing so is without loss because the principal's objective turns out to be linear in r, and hence a bang-bang policy will indeed be optimal.

Especially useful for the analysis is the total (ex-ante) probability $p(s)$ that the agent is asked to work at least until date s. This is given by

$$p(s) = 1 - \int_0^s [1 - q(s|t)] f(t) dt, \tag{2}$$

where the integral is the probability that the agent is asked to quit by s. Note that regardless of the recommendation policy, we have $p(s) \geq 1 - F(s)$. Also useful is the function $Q(t) := \int_t^T q(u|t) du$, which measures the expected *future* work for an agent who succeeds at t and obeys all recommendations. This future work occurs insofar as the agent is not informed about their success at t.

The agent's expected payoff from obeying the recommendations is

$$\underbrace{\int_0^T R(s) f(s) ds}_{\text{expected reward}} - c \times \underbrace{\int_0^T p(s) ds}_{\text{expected effort}},$$

where $R(s)f(s)$ is the ex-ante expected reward at time s measured in flow terms. The principal chooses a reward schedule, a recommendation policy, and a terminal date T to maximize

$$\int_0^T p(s)ds - \int_0^T R(s)f(s)ds$$

subject to the incentive compatibility constraint that the agent always finds it optimal to obey all recommendations. Here we have assumed that success is only a signal of effort that delivers no direct benefits to the players, which we shall relax later on.[8]

3 Incentive Compatibility

There is a variety of ways the agent can deviate from recommendations, including pausing and restarting at any time. Fortunately, as we will show, the optimal policy can be derived by focusing on a family of "local" incentive constraints. Here we will derive necessary conditions for a policy to dissuade the agent from brief (instantaneous) pauses, which we can use to identify a candidate optimal policy. We will then verify that there exist no profitable global deviations.

3.1 Instantaneous Pauses

Because we have assumed that the hazard rate of success (weakly) falls over time, the principal would like to promise, other things equal, greater rewards as time goes by. But this creates a challenge: the agent will be tempted to pause temporarily so they can succeed later, where rewards are greater, rather than now.[9] This leads us to conjecture that temporary pauses, rather than permanent ones, will be the hardest ones to deter. A natural place to start, then, is making sure that at least the shortest of such pauses—i.e., instantaneous ones—are deterred.

Fix a terminal date T. The expected payoff earned by the agent from t onward if they

[8] We have also assumed for simplicity that players don't discount time; our results remain qualitatively unchanged if they have a common discount rate.

[9] Such temptation to pause is known as the "dynamic agency" effect; see, for example, Halac, Kartik and Liu (2016). What is novel about this effect in our model is that it is impacted by the principal's feedback.

obey all recommendations, computed from the standpoint of time 0, is

$$U(t) := \int_t^T R(s) f(s) ds - c \times \int_t^T p(s) ds. \tag{3}$$

Now suppose the agent obeys all recommendations before t and after $t + \Delta t$, but shirks during the interval in-between. Such a deviation changes the arrival rate of success and therefore changes the distribution of recommendations after $t + \Delta t$ as well. The agent's continuation payoff at t, considering that pause, is

$$\widetilde{U}(t, \Delta t) := \int_{t+\Delta t}^T R(s) f(s - \Delta t) ds - c \times \int_{t+\Delta t}^T p(s|\omega_t^{\Delta t}) ds,$$

where the integrals begin at $t + \Delta t$ because the agent's flow payoff during the pause is zero, and $p(s|\omega_t^{\Delta t})$ denotes the total probability that the agent continues to spend effort through $s \geq t + \Delta t$ following this deviation (which is derived in the appendix). Incentive compatibility requires that $U(t) \geq \widetilde{U}(t, \Delta t)$, or equivalently, upon subtracting $U(t + \Delta t)$ from both sides of this inequality,

$$U(t) - U(t + \Delta t) \geq \int_{t+\Delta t}^T R(s)[f(s - \Delta t) - f(s)] ds - c \times \int_{t+\Delta t}^T [p(s|\omega_t^{\Delta t}) - p(s)] ds.$$

Dividing through by Δt and taking the limit as it converges to zero allows us to establish a *local* incentive compatibility constraint for instantaneous pauses.

Proposition 1. *The recommendation policy $q(\cdot|\cdot)$ and reward schedule $R(\cdot)$ are locally incentive compatible if, for all t,*

$$R(t) f(t) - cp(t) \geq cQ(t) f(t) + \int_t^T [R(s) - cQ(s)] |f'(s)| ds - c[F(T) - F(t)], \tag{IC}$$

where $p(t)$ is given in (2), and $Q(t) := \int_t^T q(u|t) du$ is the expected future work for an agent who succeeds at t.

The left-hand side of (IC) measures the agent's on-path expected flow rents at t; that is, their expected rewards minus instantaneous costs. These flow rents correspond to $-U'(t)$. The right-hand side is $\lim_{\Delta t \to 0}[\widetilde{U}(t, \Delta t) - U(t + \Delta t)]/\Delta t$, which represents the marginal impact of a pause at t on future rents. This pause has three effects (corresponding to each of the three terms on the right): First, it eliminates the agent's expected future effort cost $cQ(t)$ in the event they succeed at t. Second, it lowers the agent's accumulated effort and therefore raises the density of success $f(s)$ going forward, which in turn raises the chance of earning each of the net future rewards $R(s) - cQ(s)$. And, finally, it raises the chance that at each future date the agent has not yet succeeded and must therefore keep working.[10]

3.2 Minimal Rewards

Proposition 2 shows that for any given recommendation policy, there is a unique least-expensive way to deter all instantaneous pauses. This is achieved by meeting (IC) with equality at all times, starting at time T and working backwards.

Proposition 2. *Given any recommendation policy $q(\cdot|\cdot)$, there exists a unique reward schedule $R(\cdot)$ that satisfies the local incentive constraint with equality at every t. It is given by*

$$R(t) = c\left[\frac{p(t)}{f(t)} - \int_t^T \frac{f'(s)}{f(s)^2} p(s) ds - \int_t^T 1 - q(s|t) ds\right]. \tag{4}$$

Moreover, this reward schedule is pointwise smaller than any other implementing reward schedule.

Intuitively, because $R(t)$ in the incentive constraint is affected only by future rewards rather than past ones, it is possible by working backward from T to meet this constraint with equality at all times. Furthermore, it is desirable to do so because raising the reward schedule above that level over some interval of time would force the principal to

[10] The probability that the agent has not yet succeeded at s is $1 - F(s)$; a pause today (which lowers accumulated effort from s to $s - \Delta s$) raises that probability by $f(s)$, which integrated over every future date is $F(T) - F(t)$.

raise rewards at all past times, so the agent does not pause, which needlessly inflates the principal's costs.

The first term on the right-hand side of (4) is the reward level that would give the agent zero flow rent at time t, measured from an ex-ante perspective. This terms grows with $p(t)$ because a higher work probability implies a larger ex-ante cost for the agent, and falls with $f(t)$ because a greater density of success means the reward $R(t)$ is more likely to materialize. If there were no future dates, this zero-rent reward is all the principal would need to offer. The second term represents a "backward compounding" effect arising from the fact that a greater future reward requires a greater present reward, as otherwise the agent would pause. This backward compounding is modulated by $-f'/f^2$ (the speed at which $1/f$ grows) because the faster this ratio grows (i.e., the faster the likelihood of success drops) the greater the future rewards have to be.[11] The last term is an "information rebate" that the principal gets if they inform the agent sometime in the future about a success at t: the lower the $q(s|t)$, the lower the effort the agent is asked to exert in the future *after* succeeding at t, and hence the lower the promised reward $R(t)$ needs to be.

To illustrate, consider two polar opposite policies: keeping the agent completely in the dark and keeping them fully informed. If the principal offers zero feedback, while asking the agent to work with probability one until T, we have $q(s|t) \equiv p(t) \equiv 1$. As a result, the minimal implementing reward schedule is

$$R^{silence} = \frac{c}{f(T)}.$$

Intuitively, to minimize backward compounding, the principal wishes to grant zero rents at the terminal time; hence, this flat schedule is the closest the principal can get to the entire zero-rent schedule $R(t) = c/f(t)$ (which is weakly increasing) without provoking

[11]The novelty of this backward compounding effect is that it incorporates the feedback policy; that is, it tells us how we need to adjust rewards so that the dynamic agency consideration (i.e., the temptation to delay effort) does not lead the agent to pause given how information is being revealed.

any pauses. Because the agent receives no feedback, there is no rebate for the principal.

If instead the principal keeps the agent fully apprised—in which case the agent stops working as soon as they succeed—we have $q(s|t) \equiv 0$ and $p(t) = 1 - F(t)$. We term this design the *pronto* policy. The corresponding minimal reward schedule is

$$R^{pronto} = \frac{c}{\lambda(T)},$$

which is flat for a similar reason as before: the ideal schedule (in this case the zero-rent schedule net of the rebate) is also increasing; thus, given that the principal wishes to grant zero rents at the terminal time, a flat schedule is the closest the principal can get to that ideal without causing the agent to pause.

Embedded in both these policies is a dynamic version of the classic rent/efficiency trade-off: the principal exactly internalizes the cost of effort at the margin (in this case, at the terminal date) but because $R(T)$ grows with T, they must also pay greater infra-marginal rents if they seek to expand the overall gains from trade. What differs between the two policies is, on the one hand, the information rebate and, on the other, the maximum expected effort that can be asked of the agent. Because the pronto policy maximizes the rebate, it minimizes the principal's expected cost per unit of effort; but since the agent quits as soon as they succeed, it also creates an upper bound on the agent's expected total effort. Silence, in contrast, allows the agent to work for an unbounded length of time, but since there is no rebate, the cost for the principal could be very large.

In between these two examples, there are vastly many ways for the principal to offer less than immediate feedback. One example is a "delay mechanism" where the agent is informed of a success after a constant delay d. Here $q(s|t) = \mathbb{I}_{\{s \le t+d\}}$ while $p(t) = 1$ up to time d and equal to $1 - F(t-d)$ thereafter. Thus, the minimal reward schedule takes a more complex form. Namely, $R(T) = c[1 - F(T-d)]/f(t)$ (so that rents are zero at the very end) and

$$R'(t)/c = \begin{cases} 1 & \text{if } t \in [0, d) \\ 1 - f(t-d)/f(t) & \text{if } t \in [d, T-d] \\ -f(t-d)/f(t) & \text{if } t \in (T-d, T], \end{cases}$$

which means that the reward schedule first increases, then decreases and finally decreases at an even faster rate.

Such delay mechanisms have been found to be optimal in different settings (e.g., Ely, 2017), but as we shall see, are suboptimal in ours because they do not fully frontload ignorance. The optimal policy will instead be a simpler combination of the two polar opposites above.

4 Optimal policy

Here we find the optimal policy by making use of the minimal reward schedule in Proposition 2. Our first step is to use that schedule to express the principal's objective solely in terms of the work probability $p(t)$ and terminal date T.

Lemma 1. *The principal's payoff evaluated at the minimal implementing reward schedule is*

$$\int_0^T p(t)\,dt - c\int_0^T \underbrace{p(t)(1+\Phi(t)) - (1-p(t))}_{\textit{virtual effort}}\,dt, \tag{Obj}$$

where $\Phi \equiv F \times (1/f)'$ *satisfies* $\Phi(0) = 0$ *and by assumption is weakly increasing.*

The first term in (Obj) is total effort. The second term, whose integrand we term "virtual effort," is the total effort cost as experienced by the principal; that is, true cost plus information rents for the agent due to the backward compounding of rewards. Observe that at time zero, virtual effort is equal to true effort (with both equal to 1), as backward compounding is not a factor then. The function $\Phi(t)$ captures the compounding effect: a greater $(1/f(t))'$ calls for greater future rewards and hence greater past ones, whereas a larger $F(t)$ means these past rewards are paid more often. The term $1 - p(t)$ captures the information rebate: when the principal pays the agent with information (which lowers $p(t)$ below 1) there is less need for monetary rewards.

We shall find the optimal policy by first solving a substantially relaxed problem where in addition to ignoring non-local deviations, the principal selects $p(t)$ directly, subject only to an upper and lower bound, without worrying about the need to generate this function with a suitable recommendation policy q. In particular, we consider the program

$$\sup_{T\in[0,\overline{T}],\ p(\cdot)} \int_0^T p(t)\,[1-2c-c\Phi(t)]\,dt + cT \tag{P}$$

$$\text{s.t.}\ \ 1-F(t) \leq p(t) \leq 1 \text{ for all } t, \tag{Feas}$$

where the objective is equal to (Obj) upon rearranging terms, the constraint $1 - F(t) \leq p(t)$ captures the requirement that the agent does not quit before succeeding (or reaching T), and the constraint $p(t) \leq 1$ stems from the requirement that $p(t)$ is a valid probability.

This problem is easy to solve because for any T the objective is linear in p, and since the function Φ is weakly increasing, the expression in braces—whose sign determines whether $p(t)$ should be set as high or as low as is allowed— is either always negative or once it becomes negative, it remains so throughout. The optimal terminal date T is then obtained by substituting the optimal p in the objective and optimizing over a single variable. Proposition 3 describes the solution.

Proposition 3. *Let t^* be the earliest time when $1 - 2c - c\Phi(t) \leq 0$. Provided $\overline{T} \geq t^*$, the relaxed problem (P) is solved by setting*

$$p(t) = \begin{cases} 1 & \text{if } t \in [0, t^*] \\ 1 - F(t) & \text{if } t \in (t^*, T], \end{cases}$$

for some $T \in [t^*, \overline{T}]$.

Note that $t^* > 0$ if and only if $c < 1/2$. The above schedule p is uniquely implemented by the recommendation policy that never tells the agent to stop before t^*, and after that, tells them to stop as soon as they succeed. The corresponding reward schedule is then obtained by substituting p and the corresponding q into (4).

This "bang bang" solution will constitute the backbone of every optimal policy, as we show next.

Theorem 1. *Suppose $\overline{T} > t^*$. Every optimal policy consists of at most two phases:*

1. ***Silent phase**: $t \leq t^*$. Here the principal asks the agent to work with probability one regardless of their success, and remains silent throughout— that is, $q(t|s) \equiv 1$. If the agent succeeds at any time during this phase, they earn reward*

$$c/\lambda(T) + cF(t^*)/f(t^*),$$

   where $\lambda(T)$ is the hazard rate at the terminal time.

2. ***Pronto phase**: $t \in (t^*, T]$. Here the principal asks the agent to quit as soon as they succeed while remaining otherwise silent—that is, $q(t|s) \equiv 0$. If the agent succeeds at any time during this phase, they earn reward*

$$c/\lambda(T).$$

Phase 2 always has positive length, whereas phase 1 has positive length if and only if $c < 1/2$.[12]

Intuitively, the pronto policy is always used for some duration because even though it allows the agent to promptly quit upon success, it minimizes the principal's cost per unit of effort as the agent is cheapest to motivate when kept fully informed. If the principal wishes greater effort than a pronto phase alone can achieve, the agent must at least sometimes be kept in the dark. Moreover, because ignorance necessitates greater rewards and

[12]If $\overline{T} < t^*$, the optimal contract consists only of a silent phase that lasts until $\overline{T}$ and pays reward $c/f(\overline{T})$.

these get compounded backward, it is best that such ignorance is maximally frontloaded; hence the initial silent phase. Provided effort is sufficiently valuable (specifically when $c < 1/2$) this silent phase is worth having.[13] Notice that this two-phase solution combines two of our earlier examples (silence and pronto) with the modification that due to backward compounding, the prize during the initial silent phase needs to grow to compensate for any rents earned by the agent during the pronto phase, which would otherwise lead the agent to pause during the initial phase.

As it turns out, the optimal policy need not be unique because the relaxed problem (P) may admit more than one optimal cutoff between phases and more than one optimal terminal date. Such multiplicity, however, is non-generic as it would not survive a slight perturbation of the function F.

The only remaining loose end is the possibility that the agent benefits from a global deviation; i.e., one involving pauses during more than one instant. Fortunately, the simple rewards in the theorem discourage all such deviations—and make it easy to check that this is the case. Observe that these rewards are non-increasing, and because both $f(t)$ and $\lambda(t)$ are weakly decreasing, they always grant the agent non-negative flow rents.[14] This makes a pause of any nature undesirable on two fronts: it causes the agent to miss out on a portion of such rents and shifts their success probability from the present to the future, where rewards are no greater.[15]

The optimal terminal date may be equal to $\overline{T}$ regardless of how large this exogenous deadline is. This occurs, for instance, when the hazard rate is constant as this allows the principal to extend phase 2 without giving up any rents. T would strictly smaller than the deadline, in contrast, if that deadline was sufficiently large and the agent faced a Poisson

[13] The $1/2$ appears because extending the length of the silent phase from dt to $2dt$ units of time means that a higher reward must be promised over the second such interval (owing to the agent's ignorance) and because of backward compounding, this higher reward must be promised over the first interval as well.

[14] The agent's expected flow rent is $R(t)f(t) - c$ during phase 1 because they have no information, and $R(t)\lambda(t) - c$ during phase 2 (conditional on not having already succeeded) as they are fully informed.

[15] Note also that since there is no further reward forthcoming after the agent is advised to stop working, it is always optimal to follow such a recommendation.

good news experimentation technology. In this case, the hazard rate asymptotes to zero, and hence expanding phase 2 requires expanding its reward without bound, which then gets compounded backward.

While in our model F is exogenous, we can ask what type of distribution would be most profitable for the principal. Note that if F was Poisson with an arbitrarily low hazard rate, the principal would be able to achieve profits arbitrarily close to first-best $(1-c)\overline{T}$ simply by using an always-pronto policy with time-invariant reward c/λ, as this ensures that the agent works until $\overline{T}$ with arbitrarily high probability while earning no rents. This requires promising a very high reward, but paying it with very low probability.[16]

While we have assumed a non-increasing hazard rate, Theorem 1 would still hold if the hazard rate $\lambda(\cdot)$ was hump-shaped provided that $\lambda(0) \geq \lambda(T)$.[17] This inequality ensures that the agent always earns non-negative rents throughout the pronto phase, and hence they have a dominant strategy to work throughout that phase. That they always find it optimal to work during the first phase as well follows from the assumption that f is non-increasing as this ensures that a global deviation not only causes the agent to lose flow rents, but also shifts their success probability from the present to the future, where rewards are no greater.

It is also worth noting how the model would differ if, by assumption, all pauses were permanent (i.e., the agent faced a stopping problem). In this simpler environment, the backward-compounding force would vanish and the principal would need only to ensure that the agent's continuation payoff is non-negative. Hence, the principal would readily achieve the first best by remaining silent throughout, asking the agent to always work with probability one, and promising them the zero-rent schedule $R(t) = c/f(t)$.

[16]Observe that due to backward compounding, it does not suffice that the hazard rate is low during the last part of the horizon and high early on, as in that case all rewards (not only later ones) would need to be high, and hence the agent would be able to secure a high early reward with high probability.

[17]A hump shape would naturally arise if, for example, the agent's productivity is either "high" or "low", players have a common prior, and the hazard rate of success conditional on high productivity increases with cumulative effort, whereas conditional on low productivity it is zero. Such a shape can be achieved while ensuring that the resulting distribution F is concave.

In closing, the main practical implication of our baseline model is that in the presence of a coarse performance signal, hiding performance information from the agent may be an optimal way to motivate them. Consistent with this prescription, Maister (1993) observes in his in-depth analysis of professional partnerships (where the prospect of promotion is the key motivator at the beginning of a career) that partners often withhold performance information from their associates in order to prolong their trial phase (see, pp. 170 and 173).[18] A formal empirical test could rely on both the timing of feedback and the prediction that promotion times will vary across associates—with bunching occurring for early promotions (at the end of the silent phase), and with those promoted sooner earning a greater prize, such as earning a larger raise or being assigned better opportunities post promotion.[19]

5 Extensions

Here we consider a series of extensions showing, on the one hand, that a two-stage contract remains optimal under several variations of the model, and, on the other hand, that hiding information from the agent can be profitable even when they can succeed an unlimited number of times.

5.1 Continuation Payoffs

A restrictive feature of our baseline model is that as soon as a success is announced, the relationship between the principal and agent effectively ends. Here we generalize the

[18]Maister argues (informally) that this practice can discourage strong performers, and therefore advises against it. Our model captures a version of that cost through the fact that hiding information requires raising the prize—but as we have shown, this practice can actually be beneficial if paired with the right rewards, and not abused.

[19]Other models can also account for a negative relationship between promotion time and reward, which has in fact been documented in some large firms; see Gibbs (1995) for a model and empirical evidence, and Ariga, Ohkusa and Brunello (1999) for evidence concerning a large Japanese manufacturer. What is unique to our model is the combination of that prediction with the information policy and the variation in promotion times.

model by allowing each of them to gain an exogenous continuation payoff once that announcement takes place, with the principal receiving π and the agent v (both time-invariant), in addition to the endogenous reward $R(t)$.[20]

These continuation payoffs could come from several places. For instance, the agent may be able to use the news of their success to secure an outside opportunity. Alternatively, the two parties may continue their relationship and each benefit from it, e.g., the principal may switch the agent to a new set of tasks with their own monitoring technology.[21] The principal's continuation payoff may also originate from success being intrinsically valuable.

This model is very similar to the baseline. To solve it, define the agent's reward $R(t)$ as including both the monetary reward and the continuation payoff v. Because delivering $R(t)$ to the agent costs only $R(t) - v$, the principal's payoff is now

$$\int_0^T p(s)ds - \int_0^T [R(s) - v]f(s)ds + \pi F(T),$$

where we have included the principal's payoff π conditional on the success arriving, which occurs with probability $F(T)$. Note that upon rearranging terms, this payoff is equal to the baseline payoff plus a new term $(v + \pi)F(T)$.

Since $R(t)$ represents the agent's total reward, Proposition 2 (which describes the least costly incentive compatible reward schedule) still applies. Hence, the principal's payoff can be expressed as

$$\int_0^T p(t)[1 - 2c - c\Phi(t)]dT + cT + (v + \pi)F(T),$$

[20]Since there is no discounting, the principal could equivalently receive their continuation payoff as soon as the success occurs. For the agent, however, it is important that they do not receive this payoff before the announcement as that would interfere with the feedback policy.

[21]For this model to accurately capture the possibility of multiple successes we need to assume that the agent is informed of a given success before they have a chance of attaining the next one. This could stem from the agent having to switch tasks in order to attempt that next success.

which resembles the original objective (P) but with the extra term $(v + \pi)F(T)$. Since a more distant terminal date T raises the probability that the continuation payoffs are obtained, this term gives the principal an incentive to extend the deadline.

Given that the work probability $p(t)$ affects only the first term in the objective—which is the same as in the baseline model—the optimal policy has the same two-phase structure as before, as characterized in Theorem 1; moreover, since that first term does not contain v or π, the duration of the silent phase, t^*, is unchanged. The only thing that changes is therefore that the principal stretches out the second phase.

Finally, notice that even though the total reward $R(t)$ must be positive to induce the agent to work, the monetary component of that reward could in principle be negative when v is large. That possibility is ruled out whenever $v \leq c/\lambda(T)$ as this guarantees that the reward for each of the two phases is at least v.

5.2 More informed agent

In our baseline model the agent does not learn about a success unless the principal chooses to inform them. Here we instead allow the agent to learn about a success on their own, with some positive probability, independently of what principal chooses to disclose. As we shall see, the optimal contract will be a two phase policy that generalizes the baseline one by allowing for greater information for the agent during the initial phase.

Suppose in particular that the agent learns immediately of a success with some exogenous probability h even if the principal remains silent—or equivalently, that the principal must immediately inform them of a success with probability no lower that h, which means that $q(s|t) \leq 1 - h$. Because the minimum reward schedule for a given schedule q is identical to that in the baseline model, Lemma 1 (which characterizes the principal's objective) remains valid. Thus, we can obtain the optimal policy in similar way to before; namely, solve a modified relaxed problem that includes the new minimum disclosure requirement and then verify that the solution meets the remaining constraints.

The relaxed problem is now

$$\sup_{T\in[0,\overline{T}],\ p(\cdot)} \int_0^T p(t)\,[1-2c-c\Phi(t)]\,dt + cT \tag{P'}$$

$$\text{s.t. } 1-F(t) \le p(t) \le 1-hF(t) \text{ for all } t, \tag{Feas'}$$

where the new upper bound on $p(t)$ captures the fact that the agent cannot be asked to work at time t in the event that they have already succeeded by then and have learned about it, which given the minimum disclosure requirement, occurs with probability no smaller than $hF(t)$.

Since the objective is identical to that in the baseline model, assuming again that $\overline{T} > t^*$, this relaxed problem has the following bang-bang solution:

$$p(t) = \begin{cases} 1-hF(t) & \text{if } t\in[0,t^*] \\ 1-F(t) & \text{if } t\in(t^*,T] \end{cases}$$

Note that the cutoff time t^* is the same as in the baseline model. The only difference relative to the baseline is therefore the lower work probability (i.e., greater information for the agent) during the first phase.

Upon following the same steps in the proof of Theorem 1, and assuming $\overline{T} > t^*$, we obtain:

Proposition 4. *In the extended model with a more informed agent, the optimal contract is similar to the baseline one except that during the first phase (which as before lasts until t^*), the agent is asked to work as long as they don't learn on their own that they have succeeded, and is offered a reward $c/\lambda(T) + c(1-h)F(t^*)/f(t^*)$ for a success during that phase, which is lower than before.*

Here the principal stays as close as possible to the original bang-bang solution by initially keeping the agent as uninformed as possible, and then proceeding to a pronto phase. Because the agent is now more informed during the first phase, the principal can get away with a lower reward during that phase—but since the agent also quits sooner in expectation, this is a worse outcome for the principal.

5.3 Upfront effort investment

Here we consider a simple learning-by-doing scenario where the agent must exert some minimum amount S of cumulative effort before they are able to succeed, and after that they succeed according to a non-increasing density f as in the baseline model. Thus, the hazard rate starts at zero and jumps up as soon as the required effort investment is complete. As we shall see, the optimal contract for this scenario will be very similar to the baseline one, differing at most in the lengths of the two phases and their associated rewards.

Let time run from $-S$ to $\overline{T}$, suppose S is small enough that the principal can induce the agent to work while still earning positive profits, and focus on contracts where the agent is asked to work continuously between $-S$ and 0 (which is without loss because any effort pauses before the investment is complete would merely waste time). So that the agent is willing to make this investment, the principal must promise them an ex-ante utility $U(-S)$ no lower than 0, or, equivalently, a continuation utility $U(0)$ no smaller than the total investment cost $-cS$.[22]

The principal now maximizes the following weighted sum of profits and agent rents, where $\mu \geq 0$ denotes the appropriate multiplier for the new effort-investment constraint:

$$\int_0^T p(s)(1-\mu c)ds - \int_0^T R(s)(1-\mu)f(s)ds.$$

[22]The principal must also ensure that the agent does not withdraw effort during some subset of time before 0 and then makes up for the missing investment after that. But since the agent cannot earn any rewards before the investment is complete, such a deviation is at least weekly dominated by one where the agent works continuously until time zero and instead shirks, if they wish, after that, as this preserves the option to try to succeed as early as time 0.

Notice that $\mu < 1$, as otherwise this objective would call for making arbitrarily large transfers to the agent, which would in turn impose large losses on the principal. Thus, as in the baseline model, the principal wishes to pay the smallest possible reward given the desired effort schedule, and hence adopts the minimal reward schedule in Proposition 2.

Upon substituting for this minimal reward schedule, and manipulating terms, the objective becomes

$$\int_0^T p(t)[1-(2-\mu)c-(1-\mu)c\Phi(t)]dt + (1-\mu)cT,$$

which differs from the baseline objective in (P) only in the modified coefficients for the effort cost c. It follows from the properties of Φ that the solution again begins with a silent phase (of possibly zero length) and ends with a pronto phase. The silent phase, however, will now have positive length for a greater range of effort costs (i.e. $c < 1/(2-\mu)$ rather than $c < 1/2$) and will grow with μ because keeping the agent in the dark longer has the added benefit of giving them greater rents. The pronto phase, in contrast, may potentially grow or shrink (or remain unaffected) depending on the details of F.

Notice that, beyond the current application, this same contract would be optimal any time the principal wishes to maximize a weighted sum of profits and agent rents provided the agent's weight is smaller than the principal's. Doing so would, for example, allow the principal to satisfy a standard ex-ante participation constraint.

5.4 Multiple successes

In our baseline model, the imperfect effort signal (with only one success possible) was the sole friction preventing the principal from achieving the first best.[23] Here we eliminate that friction by allowing the agent to succeed an unlimited number of times, and instead introduce a monitoring cost. As we shall see, so long as that monitoring cost is high

[23] If effort was perfectly observable, the principal would be able to motivate the agent at flow cost c.

enough, withholding information from the agent—similarly to how the principal did so in the baseline model—is profitable.

Suppose, in particular, that each instant in which the agent works, a success arrives with Poisson rate $\lambda > 0$, regardless of any past effort or successes; i.e., according to p.d.f. $f(t) = \lambda e^{-\lambda t}$. The principal, however, gets to observe that success only if at that instant they are monitoring the agent, with that monitoring entailing a flow cost $m \geq 0$. We assume that $m < 1 - c$, so that effort is efficient even if it has to monitored. This version of the model echoes the baseline one in the sense that if the principal chooses not to pay the monitoring cost, an additional success is, in effect, impossible. We assume the principal can commit both to a reward and feedback policy, like in the baseline model, and also to a monitoring policy specifying at each instant whether or not to monitor the agent as a function of past messages and successes.[24] The agent, in turn, does not directly observe whether they are currently being monitored, but is aware of the monitoring policy the principal has committed to.

Notice that when $m = 0$ the principal can achieve the first best by always monitoring the agent, keeping them fully informed, and paying them a reward equal to c/λ each time they succeed. This policy induces the agent to always work while granting them zero rents, as it just barely compensates them for their effort cost c. Things are considerably more complicated when $m > 0$, as now the principal may not always wish to monitor, and yet the set of histories on which the principal can condition future monitoring (in addition to future messages and rewards) is immense. We shall be able to show, nonetheless, that a degree of silence will dominate full transparency provided m is large enough, specifically, greater than c.

Consider first a policy of full transparency, which we call *pronto*, where the agent is always informed of what the principal knows. Here the agent will work only if they are

[24] In practice, such commitment to a monitoring policy could be facilitated, for example, by the appointment of a committee whose specific task is to monitor employees and the fact that a firm that is monitoring less than it should will also end up paying the agent, on average, less than it should, and hence its reputation is likely to suffer.

being monitored, which given our assumptions the principal always finds it best to do. Indeed, by promising reward c/λ per success, the principal induces the agent to always work while granting them zero rents. Once we account for monitoring costs, this policy gives the principal a total payoff $(1 - c - m)\overline{T}$.[25]

Now consider an alternative contract with less transparency, which we call *cycles*, involving repeated silent intervals. Select an integer N, and split the agent's horizon $\overline{T}$ into equal intervals of length $\overline{T}/N$. During each such interval, monitor the agent *only* until they succeed once during the interval (or until the interval ends, whichever happens first) and remain silent until the end of the interval. At that time, inform the agent if they succeeded during the interval and, if so, pay them a reward $c/f(\overline{T}/N)$. As we have learned from the baseline model (Section 3.2), this reward is the minimum needed for the agent to work throughout a silent interval of the corresponding length, while knowing that effort is pointless after they have succeeded once, as any other success would be unobserved.[26]

Proposition 5. *In the extended model with unlimited successes, constant hazard rate, and flow monitoring cost $m > c$, the cycles contract dominates pronto whenever each interval is sufficiently short; i.e., N is sufficiently large.*

Intuitively, the advantage of *cycles* is that it saves on monitoring costs; its disadvantage is that due to the backward compounding of rewards caused by silence, it grants the agent rents. Indeed, recall from the baseline model that, at the margin, silence grants the agent c in rents, and these rents grow the longer the principal remains silent because rewards get compounded further. Thus, whenever m is greater than the marginal rent c, at least some amount of silence is desirable. When instead $m \leq c$, the *pronto* contract remains the

[25]The principal could potentially benefit from a random monitoring policy whereby a success is detected with probability less than one, but that is also less expensive. As long as there is a minimum probability of monitoring that the principal can credibly commit to, and its flow cost exceeds c, our conclusions below would be qualitatively unchanged.

[26]Because the process generating successes is stationary and the promised reward at the end of each interval depends exclusively on what happened during that interval, the agent's incentive constraint is met so long as it is met within each interval considered in isolation.

better of the two.

While we have not solved for the fully optimal contract, this result points to the potential benefits of adopting a worse, but cheaper effort signal (in this case one that detects only a subset of successes) and coupling it with less information for the agent—as this can ensure that the agent keeps working despite the deteriorated signal.

6 Conclusion

We have argued that when a principal uses a coarse performance measure, hiding information from the agent is expensive but may nonetheless be an optimal way to motivate them. We have uncovered a novel factor, which we term "backward-compounding" of rewards, that makes hiding information more costly if it happens farther in the future. When a single success is possible, this results in an optimal two-phase contract, with all ignorance frontloaded, that starts with a fully silent phase and ends with a phase of full transparency.

The key challenge was to find an optimal feedback policy among the vast set of potential ones, which we have done by showing that it suffices, under certain general conditions, to focus on deterring instantaneous effort pauses. In future work, we shall attempt to apply the present methods to multi-agent (e.g., contest) settings, and to settings with more general monitoring technologies.

References

Ariga, Kenn, Yasushi Ohkusa, and Giorgio Brunello. 1999. "Fast track: is it in the genes? The promotion policy of a large Japanese firm." *Journal of Economic Behavior & Organization*, 38(4): 385–402.

Ball, Ian. 2022. "Dynamic information provision: Rewarding the past and guiding the future."

Bonatti, Alessandro, and Johannes Horner. 2011. "Collaborating." *American Economic Review*, 101(2): 632–663.

Ely, Jeffrey C. 2017. "Beeps." *American Economic Review*, 107(1): 31–53.

Ely, Jeffrey C, and Martin Szydlowski. 2020. "Moving the goalposts." *Journal of Political Economy*, 128(2): 468–506.

Ely, Jeffrey, George Georgiadis, Sina Khorasani, and Luis Rayo. 2022. "Optimal Feedback in Contests." *Review of Economic Studies*.

Georgiadis, George. 2022. "Contracting with moral hazard: A review of theory & empirics."

Georgiadis, George, and Balazs Szentes. 2020. "Optimal monitoring design." *Econometrica*, 88(5): 2075–2107.

Gibbs, Michael. 1995. "Incentive compensation in a corporate hierarchy." *Journal of Accounting and Economics*, 19(2-3): 247–277.

Green, Brett, and Curtis R. Taylor. 2016. "Breakthroughs, Deadlines, and Self-Reported Progress: Contracting for Multistage Projects." *American Economic Review*, 106(12): 3660–3699.

Halac, Marina, Navin Kartik, and Qingmin Liu. 2016. "Optimal Contracts for Experimentation." *The Review of Economic Studies*, 83(3): 1040–1091.

Hörner, Johannes, and Andrzej Skrzypacz. 2016. "Learning, experimentation and information design."

Hörner, Johannes, and Nicolas S Lambert. 2021. "Motivational ratings." *The Review of Economic Studies*, 88(4): 1892–1935.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review*, 101(6): 2590–2615.

Kaya, Ayça. 2022. "Paying with information." *Theoretical Economics*.

Keller, Godfrey, Sven Rady, and Martin Cripps. 2005. "Strategic Experimentation with Exponential Bandits." *Econometrica*, 73(1): 39–68.

Maister, David. 1993. "Quality work doesn't mean quality service." *Managing the professional service firm*, , (Part II): 69–78.

Mason, Robin, and Juuso Välimäki. 2015. "Getting it done: dynamic incentives to complete a project." *Journal of the European Economic Association*, 13(1): 62–97.

Orlov, Dmitry. 2022. "Frequent monitoring in dynamic contracts." *Journal of Economic Theory*, 206: 105550.

Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov. 2020. "Persuading the principal to wait." *Journal of Political Economy*, 128(7): 2542–2578.

Ostrizek, Franz. 2022. "Vague by Design: Performance Evaluation and Learning from Wages."

Rayo, Luis, and Ilya Segal. 2010. "Optimal information disclosure." *Journal of Political Economy*, 118(5): 949–987.

Renault, Jérôme, Eilon Solan, and Nicolas Vieille. 2017. “Optimal dynamic information provision.” *Games and Economic Behavior*, 104: 329–349.

Sannikov, Yuliy. 2008. “A continuous-time version of the principal-agent problem.” *The Review of Economic Studies*, 75(3): 957–984.

Smolin, Alex. 2021. “Dynamic evaluation design.” *American Economic Journal: Microeconomics*, 13(4): 300–331.

A Proofs

In Section 2, we cast the principal's problem as choosing a recommendation policy $q(\cdot|\cdot)$, a reward schedule $R(\cdot)$, and a terminal date T, while remarking that restricting attention to deterministic terminal dates is without loss. To establish this, we prove our results here under the assumption that a recommendation policy consists of two objects: $q(s|t)$ which denotes the probability that the agent is asked to work through date s conditional on having succeeded at $t \leq s$, and $r(t)$ which is a non-increasing function denoting the probability that the agent is advised to work through date t conditional on *not* having succeeded yet. In this case, it is without loss of generality to set $T = \overline{T}$, and (2) becomes

$$p(s) = r(s)[1 - F(s)] + \int_0^s r(u)f(u)q(s|u)du, \tag{5}$$

the agent's expected payoff from obeying the recommendations is

$$\int_0^{\overline{T}} r(s)R(s)f(s)ds - c \times \int_0^{\overline{T}} p(s)ds,$$

and the principal's objective is

$$\int_0^{\overline{T}} p(s)ds - \int_0^{\overline{T}} r(s)R(s)f(s)ds. \tag{6}$$

Note that choosing a (deterministic) terminal date T is equivalent to $r(\cdot)$ jumping from 1 to 0 at $t = T$; i.e., $r(t) \equiv \mathbb{I}_{t \leq T}$.

A.1 Proof of Proposition 1

Suppose the agent obeys all recommendations before t and after $t + \Delta t$, but shirks in the interval in-between. Then the total probability that they continue to work through

$s \geq t + \Delta t$ equals

$$p(s|\omega_t^{\Delta t}) = r(s)\left[1 - F(s - \Delta t))\right] + \int_0^t r(u)f(u)q(s|u)du + \int_{t+\Delta t}^s r(u)f(u - \Delta t)q(s|u)du.$$

Next, we characterize the following limit:

$$\begin{aligned}
\dot{p}(s|t) &= \lim_{\Delta t \to 0} \frac{p(s|\omega_t^{\Delta t}) - p(s)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \left\{ r(s) \frac{F(s) - F(s - \Delta t)}{\Delta t} \right. \\
&\qquad - \frac{1}{\Delta t} \int_t^{t+\Delta t} r(u)f(u)q(s|u)du \\
&\qquad \left. - \int_{t+\Delta t}^s r(u) \frac{f(u) - f(u - \Delta t)}{\Delta t} q(s|u)du \right\},
\end{aligned}$$

which represents the marginal change in the work probability $p(s)$ following an infinitesimal deviation at t. Since F is differentiable, the first term is $r(s)f(s)$. Applying L'Hopital's rule and the Lebesgue differentiation theorem, the limit of the second term (which exists almost everywhere) is simply the integrand evaluated at t, namely, $-r(t)f(t)q(s|t)$. As for the third term note that

$$\begin{aligned}
&- \lim_{\Delta t \to 0} \int_{t+\Delta t}^s r(u) \frac{f(u) - f(u - \Delta t)}{\Delta t} q(s|u)du \\
&= - \lim_{\Delta t \to 0} \int_t^s r(u) \frac{f(u) - f(u - \Delta t)}{\Delta t} q(s|u)du \\
&= - \int_t^s r(u) \lim_{\Delta t \to 0} \frac{f(u) - f(u - \Delta t)}{\Delta t} q(s|u)du = - \int_t^s r(u)f'(u)q(s|u)du,
\end{aligned}$$

where the third line follows from dominated convergence. Therefore, almost everywhere,

$$\dot{p}(s|t) = r(s)f(s) - r(t)f(t)q(s|t) - \int_t^s r(u)f'(u)q(s|u)du. \tag{7}$$

Next, recall that incentive compatibility requires that

$$U(t+\Delta t) - U(t) \leq \int_{t+\Delta t}^{\overline{T}} r(s)\left[f(s) - f(s-\Delta t)\right] R(s)ds + c \times \int_{t+\Delta t}^{\overline{T}} \left[p(s|\omega_t^{\Delta t}) - p(s)\right] ds.$$

Dividing both sides by Δt and taking the limit as $\Delta t \to 0$ we have

$$\begin{aligned} U'(t) &\leq \int_t^{\overline{T}} r(s)f'(s)R(s)ds + c \lim_{\Delta t \to 0} \int_{t+\Delta t}^{\overline{T}} \frac{p(s|\omega_t^{\Delta t}) - p(s)}{\Delta t} ds \\ \Rightarrow cp(t) - r(t)f(t)R(t) &\leq \int_t^{\overline{T}} r(s)f'(s)R(s)ds + c\int_t^{\overline{T}} \dot{p}(s|t)ds. \end{aligned} \tag{8}$$

Finally, for any T, using the definition $Q(t) := \int_t^T q(u|t)du$, letting $r(t) \equiv \mathbb{I}_{t \leq T}$, and rearranging terms yields (IC). □

A.2 Proof of Proposition 2

Fix a recommendation policy $\{q(\cdot|\cdot), r(\cdot)\}$, and hence $p(\cdot)$ and $\dot{p}(\cdot,\cdot)$, and let R be any (locally) incentive compatible reward schedule. From (8), this implies that it satisfies

$$r(t)R(t) \geq \underbrace{\frac{1}{f(t)}\left[c \cdot p(t) - c\int_t^{\overline{T}} \dot{p}(s|t)ds - \int_t^{\overline{T}} r(s)f'(s)R(s)ds\right]}_{=:Z^1(t)}.$$

Notice that the function $Z^1(t) \leq r(t)R(t)$ for all t. It follows from the fact that $f'(\cdot) \leq 0$ and (7) that $\dot{p}(s|t) \leq f(s) - \int_t^s f'(u)du = f(t)$, and hence

$$Z^1(t) \geq \frac{c}{f(t)}\left(p(t) - \int_t^{\overline{T}} f(t)ds\right) = c\left(\frac{p(t)}{f(t)} - (\overline{T} - t)\right) =: \beta(t)$$

for all t. Continuing in this manner, define for all $k \geq 2$, the function Z^k by

$$Z^k(t) = \frac{1}{f(t)}\left[c \cdot p(t) - c\int_t^{\overline{T}} \dot{p}(s|t)ds - \int_t^{\overline{T}} Z^{k-1}(s)f'(s)ds\right].$$

Because $f'(\cdot) \leq 0$ and $Z^1(t) \leq r(t)R(t)$ we have that $\beta(t) \leq Z^2(t) \leq Z^1(t)$ for all t. By induction we have that $\beta(t) \leq Z^k(t) \leq Z^{k-1}(t)$ for all t. We have thus constructed a pointwise decreasing sequence of functions bounded below by the function β. Let Z be the pointwise limit. By the dominated convergence theorem we have

$$Z(t) = \frac{1}{f(t)} \left[c \cdot p(t) - c \int_t^{\overline{T}} \dot{p}(s|t) ds - \int_t^{\overline{T}} Z(s) f'(s) ds \right].$$

Define a new reward schedule R^* by $r(t)R^*(t) = Z(t)$. Note that R^* satisfies (8) with equality, and moreover, it is weakly pointwise lower than the original R. We will next show that there is a unique R^* which satisfies the incentive constraint with equality and therefore that R^* is pointwise lower than any incentive compatible schedule (since R was arbitrary.)

It is not difficult to verify that (4) satisfies (IC) with equality for all t. Here we provide a detailed derivation. Let $G(t) := \int_t^{\overline{T}} r(s) f'(s) R^*(s) ds$ and $H(t) := c \cdot p(t) - c \int_t^{\overline{T}} \dot{p}(s|t) ds$. Notice that $G'(t) = -r(t) f'(t) R^*(t)$ almost everywhere.[27] We can then rewrite (8) as

$$G'(t) = \frac{f'(t)}{f(t)} \left[G(t) - H(t) \right].$$

This is a linear differential equation with boundary condition $\lim_{T \to \overline{T}} G(T) = 0$, and admits the following unique solution:

$$G(t) = f(t) \int_t^{\overline{T}} \frac{f'(s) H(s)}{f(s)^2} ds.$$

Note that

$$r(t)R(t) = -\frac{G'(t)}{f'(t)} = \frac{H(t)}{f(t)} - \int_t^{\overline{T}} \frac{f'(s) H(s)}{f(s)^2} ds. \tag{9}$$

[27] We can apply the Fundamental Theorem of Calculus at almost every t. The reward schedule $R^*(\cdot)$ is continuous almost everywhere because it satisfies the incentive constraint with equality and the right-hand side of the incentive constraint is continuous almost everywhere.

Letting

$$
\begin{aligned}
H_1(t) &= \int_t^{\overline{T}} r(s)f(s)ds, \\
H_2(t) &= -r(t)f(t)\int_t^{\overline{T}} q(s|t)ds, \text{ and} \\
H_3(t) &= -\int_t^{\overline{T}}\int_t^{s} r(u)f'(u)q(s|u)du\, ds,
\end{aligned}
$$

and using (7) we have $H(t) = cp(t) - c\left[H_1(t) + H_2(t) + H_3(t)\right]$. Notice that $H_3'(t) = -f'(t)H_2(t)/f(t)$, and by integrating by parts we have

$$
\frac{H_3(t)}{f(t)} - \int_t^{\overline{T}} \frac{f'(s)H_3(s)}{f(s)^2}ds = \frac{H_3(t)}{f(t)} + \int_t^{\overline{T}} \left(\frac{1}{f(s)}\right)' H_3(s)ds = \int_t^{\overline{T}} \frac{f'(s)}{f(s)^2}H_2(s)ds.
$$

Using integration by parts again, we have

$$
\frac{H_1(t)}{f(t)} - \int_t^{\overline{T}} \frac{f'(s)H_1(s)}{f(s)^2}ds = \int_t^{\overline{T}} r(s)ds.
$$

Then (9) can be rewritten as

$$
\begin{aligned}
r(t)R^*(t) &= c\,\frac{p(t) - H_1(t) - H_2(t) - H_3(t)}{f(t)} - c\int_t^{\overline{T}} \frac{f'(s)}{f(s)^2}\left[p(s) - H_1(s) - H_2(s) - H_3(s)\right] ds \\
&= c\left[\frac{p(t)}{f(t)} - \int_t^{\overline{T}} \frac{f'(s)}{f(s)^2}p(s)ds - \int_t^{\overline{T}} r(s)ds + r(t)\int_t^{\overline{T}} q(s|t)ds\right]. \qquad (10)
\end{aligned}
$$

For any T, substituting $r(t) \equiv \mathbb{I}_{t\leq T}$ and $q(s|t) = 0$ for all $s \geq T$ yields (4). □

A.3 Proof of Lemma 1

By substituting the minimal implementing reward schedule characterized in (10), we can rewrite the principal's objective, defined in (6), as

$$
\int_0^{\overline{T}} (1-c)p(t) + c\,f(t)\int_t^{\overline{T}} \frac{f'(s)}{f(s)^2}p(s)ds + c\,f(t)\int_t^{\overline{T}} r(s)ds - c\,r(t)f(t)\int_t^{\overline{T}} q(s|t)ds\,dt.
$$

To simplify the double integrals, we use that $\int_0^{\overline{T}} a(t) \int_t^{\overline{T}} b(s)ds\,dt = \int_0^{\overline{T}} b(t) \int_0^t a(s)ds\,dt$ for any integrable functions $a(t)$ and $b(t)$. In particular, we have

$$\int_0^{\overline{T}} f(t) \int_t^{\overline{T}} \frac{f'(s)}{f(s)^2} p(s)ds = \int_0^{\overline{T}} \frac{f'(t)F(t)}{f(t)^2} p(t)dt = -\int_0^{\overline{T}} p(t)\Phi(t)dt,$$

$$\int_0^{\overline{T}} f(t) \int_t^{\overline{T}} r(s)ds = \int_0^{\overline{T}} r(t)F(t)dt, \text{ and}$$

$$\int_0^{\overline{T}} r(t)f(t) \int_t^{\overline{T}} q(s|t)ds\,dt = \int_0^{\overline{T}} \int_0^t r(u)f(u)q(t|u)\,du\,dt = \int_0^{\overline{T}} p(t) - r(t) + r(t)F(t)\,dt,$$

where the last equality follows from (5). Using these identities, we can simplify the principal's objective as

$$\int_0^{\overline{T}} p(t)dt - c\int_0^{\overline{T}} p(t)\left(1 + \Phi(t)\right) - \left(r(t) - p(t)\right)\,dt,$$

and substituting $r(t) \equiv \mathbb{I}_{t \leq T}$ for any T yields (Obj). □

A.4 Proof of Proposition 3

When the principal chooses $r(\cdot)$ in lieu of a terminal date T, (P) can be rewritten as

$$\max_{p(\cdot),\, r(\cdot)} \int_0^{\overline{T}} p(t)\left[1 - 2c - c\Phi(t)\right]dt + c\int_0^{\overline{T}} r(t)dt$$
$$\text{s.t. } r(t)[1 - F(t)] \leq p(t) \leq 1 \text{ for all } t$$
$$0 \leq r(t) \leq 1 \text{ for all } t \text{ and non-increasing in } t.$$

Fixing an $r(\cdot)$, we can maximize the objective pointwise for each t, and noting that $\Phi(t)$ is weakly increasing by assumption, it follows that the bracketed expression, is either always negative, or once it becomes negative, it remains so throughout, and therefore the optimal p is the non-increasing function

$$p(t) = \begin{cases} 1 & \text{if } t \leq t^* \\ r(t)[1 - F(t))] & \text{if } t > t^*, \end{cases}$$

where t^* has been defined as the smallest t such that $1 - 2c - c\Phi(t) \leq 0$.

Turning to the choice of $r(\cdot)$, substituting the above $p(\cdot)$ into the objective we have

$$\max_{0 \leq r(\cdot) \leq 1} \int_0^{t^*} [1 - 2c - c\Phi(t) + cr(t)]dt + \int_{t^*}^{\overline{T}} r(t) \{[1 - F(t)][1 - 2c - c\Phi(t)] + c\} \, dt$$

with the additional constraint that $r(\cdot)$ is non-increasing. Observe that for all $t \leq t^*$, the objective increases in $r(t)$ (at rate c). For $t > t^*$ the objective increases in $r(t)$ if and only if $[1 - F(t)][1 - c - c\Phi(t)] + c \geq 0$. Because the objective is linear in $r(\cdot)$, it is without loss to set $r(t) \in \{0, 1\}$, and hence there exists an optimal $r(t)$ with the following form:[28]

$$r(t) = \begin{cases} 1 & \text{if } t \leq T \\ 0 & \text{otherwise,} \end{cases}$$

for some $T \in (t^*, \overline{T}]$. This implies that $p(t) = 1$ for $t \leq t^*$, $p(t) = 1 - F(t)$ for $t \in (t^*, T]$ and $p(t) = 0$ for $t > T$, which is non-increasing as desired. Finally, note that the optimal $r(\cdot)$ is equivalent to choosing a deterministic terminal date T after which the relationship is dissolved. □

[28] We can obtain this solution as follows. Begin by ignoring the monotonicity constraint and find a relaxed optimal r with values contained in $\{0, 1\}$. Next, suppose that relaxed solution violates the monotonicity constraint and select the earliest adjacent intervals such that $r = 0$ over the first one (denoted A) and $r = 1$ over the second (denoted B). Observe that there is a constrained solution such that r is constant over $A \cup B$ (as otherwise we could weakly raise the objective by setting r constant and equal to whichever value the alternative conjectured constrained optimal had taken at time $\sup A$). Now set r equal to either 0 or 1 over $A \cup B$ and whenever possible equal to 1 (which is again without loss). Finally, continue repeating the same process until r is monotone throughout.

A.5 Proof of Theorem 1

We establish this theorem in 3 steps.

First, we note from (2) that the optimal $p(\cdot)$ characterized in Proposition 3 is implemented by the recommendation policy that sets $q(s|t) = 1$ for all $s \leq t^*$ and any t, and otherwise sets $q(s|t) = 0$. In other words, it comprises a silent phase $[0, t^*]$, during which the principal asks the agent to work with probability one regardless of their success, and a pronto phase $(t^*, T]$, during which the agent is advised to quit upon succeeding. Finally, if they have not succeeded by T they are advised to stop without reward. Note that $q(\cdot|t)$ is non-increasing for any t as required.

Second, we substitute the above recommendation policy into (4) to obtain the optimal reward schedule. For $t \in (t^*, T]$, we have

$$R(t) = c\left[\frac{1 - F(t)}{f(t)} - \int_t^T \frac{f'(s)}{f(s)^2}[1 - F(s)]ds - (T - t)\right] = \frac{c}{\lambda(T)},$$

where the last equality follows by integrating by parts and $\lambda(T) = f(T)/[1 - F(T)]$. Next, for $t \in [0, t^*]$ we have

$$R(t) = c\left[\frac{1}{f(t)} - \int_t^{t^*} \frac{f'(s)}{f(s)^2}ds - \int_{t^*}^T \frac{f'(s)}{f(s)^2}[1 - F(s)]ds - (T - t^*)\right] = c\left[\frac{F(t^*)}{f(t^*)} + \frac{1}{\lambda(T)}\right]$$

where the last equality again follows by integrating by parts and $\lambda(T) = f(T)/[1 - F(T)]$.

Finally, we verify that the above recommendation policy and reward schedule pair is globally incentive compatible. Following the recommendations during the second phase, when the agent is asked to work only if they have yet to succeed is a dominant strategy, because the prize is time-invariant and the agent earns rents (owing to the non-increasing hazard rate).[29] Turning to the first phase, if the agent shirks for Δ units of time and

[29]Note that this is the only place in the proof where we use that the hazard rate $\lambda(t)$ is non-increasing. In fact it suffices that $\lambda_T \leq \lambda_t$ for all $t \in [t^*, T]$ as we explain in Remark I following Theorem 1.

otherwise follows the recommendations, then their ex-ante payoff is

$$\frac{\widetilde{U}(0,\Delta)}{c} = \left[\frac{1}{\lambda(T)} + \frac{F(t^*)}{f(t^*)}\right] F(t^* - \Delta) + \frac{F(T-\Delta) - F(t^*-\Delta)}{\lambda(T)} - (T-\Delta) + \int_{t^*}^{T} F(t-\Delta)dt.$$

Using the concavity of F it is straightforward to show that $\widetilde{U}(0,\Delta)$ decreases in Δ, and so the agent prefers to follow the recommendations throughout the first phase as well. □

A.6 Proof of Proposition 5

Recall that the pronto contract gives the principal a payoff of $\Pi_{pronto} = (1-c-m)\overline{T}$. Now consider the cycles contract and dividing the horizon $\overline{T}$ into N equal-length intervals. The expected monitoring cost is

$$N \times m \times \left(\int_0^{\overline{T}/N} \lambda t e^{-\lambda t} dt + \frac{\overline{T}}{N} e^{-\lambda \overline{T}/N} \right) = \frac{N \times m}{\lambda}(1 - e^{-\lambda \overline{T}/N}),$$

and the expected compensation cost is

$$N \times c \times \frac{F(\overline{T}/N)}{f(\overline{T}/N)} = N \times c \times \frac{1 - e^{-\lambda \overline{T}/N}}{\lambda e^{-\lambda \overline{T}/N}} = \frac{N \times c}{\lambda}(e^{\lambda \overline{T}/N} - 1).$$

Thus, the principal's payoff is

$$\Pi_{cycles}(N) = \overline{T} - \frac{N \times c}{\lambda}(e^{\lambda \overline{T}/N} - 1) - \frac{N \times m}{\lambda}(1 - e^{-\lambda \overline{T}/N}),$$

and observe that $\lim_{N\to\infty} \Pi_{cycles}(N) = \Pi_{pronto}$. Treating N as real-valued, we can obtain

$$\Pi'_{cycles}(N) = -\frac{c}{\lambda}(e^{\lambda \overline{T}/N} - 1) + \frac{c\overline{T}}{N} e^{\lambda \overline{T}/N} - \frac{m}{\lambda}(1 - e^{-\lambda \overline{T}/N}) - \frac{m\overline{T}}{N} e^{-\lambda \overline{T}/N},$$

and it is straightforward to show that $\Pi'_{cycles}(N) < 0$ if and only if

$$\frac{m}{c} > \frac{1 - e^{\lambda\overline{T}/N} + (\lambda\overline{T}/N)e^{\lambda\overline{T}/N}}{1 - e^{-\lambda\overline{T}/N} - (\lambda\overline{T}/N)e^{-(\lambda\overline{T}/N)}}.$$

The right-hand side is approaches infinity as $N \to 0$, it decreases in N, and approaches 1 as $N \to \infty$. So provided that $m/c > 1$, there exists an N^* such that $\Pi_{cycles}(N)$ is decreasing for all $N \geq N^*$. Coupled with the fact that $\lim_{N\to\infty} \Pi_{cycles}(N) = \Pi_{pronto}$, it follows that the cycles contract dominates the pronto one for all $N \geq N^*$. □