Feedback Design in Dynamic Moral Hazard

Jeffrey C. Ely † George Georgiadis ‡ Luis Rayo §

May 17, 2023

Abstract

We study the joint design of dynamic incentives and performance feedback for an environment with a coarse (all-or-nothing) measure of performance. 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 comes at a cost because an ignorant agent is 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. Whenever the agent’s hazard rate of success falls sufficiently over time, the principal should eventually give up on them, as occurs in practice with up-or-out promotion policies.

*We are grateful to Sandeep Baliga, Matteo Camboni, Yeon-Koo Che, Theo Durandard, Alex Frankel, Ben Friedrich, Michael Gibbs, 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

The early stages of many careers serve as a trial period where employees seek a high-value reward (e.g., a promotion) while working hard to demonstrate their productivity; e.g., Landers, Rebitzer and Taylor (1996) and Barlevy and Neal (2019). A key design component of such jobs is the performance feedback offered to the employee, 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 environment where the underlying monitoring technology is coarse. In our baseline model, a principal wishes to maximize the length of time over which an agent exerts effort, and must discourage pauses throughout that time. The coarseness in the performance measure—which we model as an all-or-nothing signal observed only by the principal—means that hiding information, while costly, can be a profitable way to expand 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 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.

This model can be interpreted as a (stylized) representation of a professional services career—or a tenure-track career in academia—where an associate works with the pri-
mary goal of being promoted to partner, and the firm’s superiors potentially keep secret the exact moment when they have concluded that a promotion threshold (the all-or-nothing performance measure in our simple model) has been met. In our model, as long as effort is sufficiently valuable, the silent portion of the contract has positive length, and provided the hazard rate declines sufficiently over time, the associate is granted finite time to secure promotion. In addition, early promotion is accompanied by greater rewards. (We expand on this discussion in Section 4.)

While our model assumes that no more than one success is possible, we consider an extension where the players secure continuation payoffs that may stem from a continued relationship where the agent works on new tasks with their own monitoring technology (and with additional successes potentially arriving then). We also consider an alternative specification of the objective where the principal maximizes a weighted sum of effort and agent rents. Our main findings are robust to all these specifications, with the optimal arrangement being a two-phase one with all the hiding of information occurring early on.¹

**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.² These models, in contrast to ours, do not permit the principal to strategically provide feedback to the agent.

¹We also consider an extension where the reward schedule is exogenous and show by example that a silent phase can be beneficial in that case as well.

²Keller, Rady and Cripps (2005) and Bonatti and Horner (2011) analyze the equilibria of such good-news Poisson experimentation models.
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 myopic.3

Ely and Szydlowski (2020), Orlov, Skrzypacz and Zryumov (2020), 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. Because the agent’s decision is irreversible, their strategy depends both on the information received to date and on the information they expect to receive if they continue, which gives rise to rich dynamics. These models assume the state is binary; whereas in Ball (2022) it follows a Brownian motion.

Finally, in Varas, Marinovic and Skrzypacz (2020) and Hörner and Lambert (2021) the evolution of the state is influenced by the receiver’s actions. Smolin (2021) and Kaya (2022) are among the first to endogenize the game’s payoff structure, albeit parametrically. There the principal offers the agent a share of output, which coincides with the state. In our model, in contrast, the principal enjoys full flexibility when choosing the agent’s rewards (together with the feedback policy).4,5

---

3Hörner and Skrzypacz (2016) discuss conditions under which the optimal policy is not static.
4In 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.
5Also 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.
2 Model

At each instant $t$ and up to an endogenous terminal date $T$, an agent privately chooses whether to spend effort toward producing a binary signal, which we call a “success.” $T$ can be finite or infinite. 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.\footnote{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.}

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 $F : [0, \infty] \to [0, 1]$ of the cumulative time $e$ that the agent has spent working. $F(\cdot)$ is weakly increasing and satisfies $F(0) = 0$ and $F(\infty) \leq 1$. A useful equivalent representation is to imagine that a hidden random effort requirement $z \in [0, \infty]$ 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 $f$ (except possibly at $e = \infty$) 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 make use of the following function

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

whose significance will soon be clear. The problem of maximizing expected effort minus expected rewards will be well-behaved when $\Phi(t)$ is strictly increasing, which we will
assume.\textsuperscript{7}

The principal selects the terminal time $T$ and offers the agent a monetary reward $R(t)$ if success arrives at $t \in [0,T]$.$^8$ 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 any. By standard arguments it is sufficient to consider direct feedback policies in which the principal recommends to the agent whether to work or wait in each period. 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, it is without loss for the principal to recommend 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 $q(s|t)$ denoting the probability that the agent is still asked to work at date $s$ conditional on having succeeded at $t \leq s$. Since quitting is permanent (and no success can occur after quitting) this function is non-increasing in $s$. If the agent has not yet succeeded, the principal simply recommends that they keep working (until $T$).\textsuperscript{9}

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

\textsuperscript{7}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.

\textsuperscript{8}As we show in the Appendix, it is without loss to restrict to deterministic terminal dates. Furthermore, 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.

\textsuperscript{9}The recommendation policy could in principle include a probability that the agent is asked to keep working if they have not yet succeeded. However, since the terminal date $T$ can be adjusted, it is without loss of generality to set this probability to 1, as we show in the appendix.
\[ p(s) = 1 - \int_0^s [1 - q(s|t)] f(t) dt, \]  

(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

\[ \int_0^T R(s)f(s) ds - c \times \int_0^T p(s) ds, \]

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; but as we show in Section 5, this formulation leads to similar results as a more general one that allows for such benefits.\(^{10}\)

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

\(^{10}\) We have also assumed for simplicity that players don’t discount time; our results remain qualitatively unchanged if they have a common discount rate. In either case, there is no loss in assuming that the prize is awarded at \( T \).
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

The expected payoff earned by the agent from \( t \) onward if they 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

\[
\tilde{U}(t, \Delta t) := \int_{t+\Delta t}^T R(s)f(s-\Delta t)ds - c \times \int_{t+\Delta t}^T p(s|\omega^\Delta t_t)ds
\]

where the integrals begin at \( t + \Delta t \) because the agent’s flow payoff during the pause is zero, and \( p(s|\omega^\Delta t_t) \) 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 \( U(t) \geq \tilde{U}(t, \Delta t) \), or equivalently, upon subtracting \( U(t + \Delta t) \) from both sides of this inequality,

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

Dividing through by \( \Delta t \) 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 $t$,

$$R(t)f(t) - cp(t) \geq cQ(t)f(t) + \int_t^T \left[R(s) - cQ(s)\right]f'(s)ds - c[F(T) - F(t)],$$

(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 $\lim_{\Delta t \to 0} \frac{U(t, \Delta t) - U(t + \Delta t)}{\Delta t}$, 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.\(^{11}\)

One may also interpret (IC) using the language of mechanism design. Suppose the agent’s “true type” is their accumulated effort, so that at every instant along the equilibrium path there is a new incarnation of the agent. To ensure that the agent does not marginally misrepresent this type downward at time $t$ (by pausing momentarily), the change in equilibrium payoff $-U'(t)$ must be bounded below, per the Envelope Theorem, by (minus) the direct derivative of her aggregate future flow payoffs with respect to

\(^{11}\)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 $F(T) - F(t)$. 

9
their true type, which is precisely the right-hand-side of (IC).

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 $q(⋅|⋅)$, there exists a unique reward schedule $R(⋅)$ that satisfies the incentive constraint with equality at every date $t$. It is given by

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

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 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

---

12 Upon substituting for $p(s)$ and changing the order of integration, the agent’s continuation payoff is

$$\int_t^T [R(s) - cQ(s)]f(s)ds - c \int_t^T [1 - F(s)]ds - c \int_t^T \int_t^T q(s|u)f(u)du ds,$$

where the first term represents expected net rewards due to a future success, the second term denotes future effort costs in the event that the agent has not yet succeeded, and the last term represents expected future effort costs owing to a success at time $t$ or earlier. Here, the “true type” (actual accumulated effort) enters only through $f$, $F$, and the upper limit of the inside integral in the last term, which shrinks when the agent pauses.
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. 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 \( q(s|t) = p(t) = 1 \). As a result, the minimal implementing reward schedule is

\[
R_{\text{silence}} = \frac{c}{f(T)}.
\]

This flat (time-invariant) schedule is the closest the principal can get to the (increasing) zero-rent reward schedule \( R(t) = c/f(t) \) without provoking 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 \( q(s|t) = 0 \) and \( p(t) = 1 - F(t) \). We term this design the \textit{pronto} policy. The corresponding minimal reward schedule is

\[
R_{\text{pronto}} = \frac{c}{\lambda(T)}.
\]

which is flat for a similar reason as before: the ideal schedule (in this case the zero-rent schedule net of the rebate) is increasing, and so a flat schedule is the closest the principal
can get to the 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 inframarginal 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 \( q(s|t) = \mathbb{I}_{s \leq t+d} \) 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, \( R(T) = [1 - F(T-d)] / f(t) \) (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

$$\int_0^T p(t) \, dt - c \int_0^T \underline{p(t)(1 + \Phi(t)) - (1 - p(t))} \, dt,$$

(Obj)

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

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

(P)

subject to $1 - F(t) \leq p(t) \leq 1$ for all $t$,

(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, 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 $\Phi$ is strictly 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 single-crosses zero from above. 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 $1 - 2c - c\Phi(t) \leq 0$\(^{13}\). 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 $T > t^*$.

This “bang bang” effort schedule will constitute the backbone of every optimal policy, as we show next.

**Theorem 1.** 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, $q(s|t) \equiv 1$. 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.

\(^{13}\)Note that $t^* > 0$ if and only if $c < 1/2$. 
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, \( q(s|t) \equiv 0 \). 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 \( c < 1/2 \).\(^{14}\)

Intuitively, the pronto policy is always used during some length of time 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 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.\(^{15}\) 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 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

\(^{14}\)The optimal recommendation policy \( q \) is obtained by noting from (2) that \( p(t) = 1 \) is uniquely achieved by never telling the agent to stop; whereas \( p(t) = 1 - F(t) \) is uniquely achieved by informing the agent as soon as they succeed. The optimal reward schedule is in turn obtained by substituting the optimal \( q \) and \( p \) into (4). Notice that \( q \) is monotone and hence constitutes a valid probability.

\(^{15}\)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.
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.\(^{16}\) 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.

The optimal terminal date may be infinite. 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 be finite, in contrast, in a Poisson good news experimentation setting (see footnote 7 for the definition) because the hazard rate asymptotes to zero. In that case, expanding phase 2 requires ever greater rewards, which then get compounded backward.

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 \( \lambda(t^*) \geq \lambda(T) \).\(^{17}\) This inequality ensures that the agent earns non-negative rents throughout the pronto phase, and hence the policy characterized in Theorem 1 is globally incentive compatible.

The main practical implication of our model is that hiding information early on may be an optimal way to motivate the agent, despite the costs of doing so. 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, thus allowing the firm to hold onto associates with poor promotion prospects for longer (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

\(^{16}\)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.

\(^{17}\)A hump-shape would naturally arise, for example, when 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.

\(^{18}\)Maister himself 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.
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.\footnote{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.}

5 Extensions

Here we consider three extensions, which taken together suggest that the principal’s desire to frontload ignorance is robust.

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 model by allowing each of them to gain an exogenous continuation payoff once that announcement takes place, with the principal receiving $\pi$ and the agent $v$, in addition to the endogenous reward $R(t)$.\footnote{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 principal’s feedback policy.}

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.\footnote{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.} 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 $\pi$ 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

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

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 $\pi$, 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 $v$. 

18
5.2 Pareto Problem

Here we consider the problem of maximizing a weighted sum of principal profits and agent rents. This would capture for example a scenario where the agent must be guaranteed a minimum level of rents in order to participate, in which case the agent’s Pareto weight would correspond to a Lagrange multiplier for a constraint that rents do not fall below a certain level.\(^{22}\)

Letting \( \mu \in [0, 1) \) denote the agent’s Pareto weight, the principal now solves\(^{23}\)

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

Since \( \mu < 1 \), the principal, as before, wishes to minimize rewards given the desired effort schedule, and thus adopts the minimum reward schedule in Proposition 2. Upon substituting for this schedule, and manipulating terms, the objective becomes

\[
\int_0^T p(t)[1 - (2 - \mu)c - (1 - \mu)c\Phi(t)]dt + (1 - \mu)cT, \quad (P')
\]

which differs from the baseline problem \((P)\) only in the modified coefficients for the effort cost \(c\).

Because the term in square braces in \((P')\) is either always negative or single-crosses zero from above, the solution again begins with a silent phase (of possibly zero length) and ends with a pronto phase. This is because the costs of backward compounding—which are minimized when the agent’s ignorance is frontloaded—are still present. What changes is the length of the phases. The silent phase will now have positive length as long as \( c < 1/(2 - \mu) \) (rather than \( c < 1/2 \)), and because the point of single crossing grows with \( \mu \), this phase becomes longer the greater the agent’s Pareto weight. Intuitively speaking, keeping the agent in the dark now has the added benefit of giving the agent

\(^{22}\)The agent could also potentially arrive with some initial cash, which without loss the principal would extract upfront (given their commitment power) but would then need to return in the form of rents.

\(^{23}\)If \( \mu \) were 1 or greater, the objective would call for granting infinite rents to the agent.
large rents. The pronto phase, in contrast, may potentially grow or shrink depending on the details of $\Phi$.

### 5.3 Exogenous Rewards

In this last extension we consider a scenario where the principal is able to design the feedback policy, but not the reward schedule $R$, which we take as exogenously given. For example, a division manager might have the authority to promote subordinates and provide feedback, but not to determine pay. Alternatively, the agent may work for an exogenous reputational prize, rather than a cash prize, that the principal does not control—and which arrives when a success is announced. The principal chooses $q(\cdot | \cdot)$ and $T$ to solve

$$
\sup \int_0^T p(t) dt \\
\text{s.t. (IC), } T \geq 0, \ q(\cdot | t) \in [0, 1] \text{ non-increasing for all } t,
$$

where $p(t)$ is given in (2).

Because the feedback policy—which can in principle be highly elaborate—will try to compensate for any imperfections in the reward schedule, this optimization problem is considerably more challenging than in the baseline case, and in fact does not appear to admit a general analytical solution. However, for any given terminal time $T$, this problem is a linear one (since both $p$ and $Q$ are linear in $q$), which facilitates numerical computation.

Figure 1 depicts a simple example involving a time invariant reward and good-news Poisson experimentation.\footnote{In this example, $c = 0.4$, $R(t) \equiv 5$, and $f$ assumes that with probability 0.9 the agent’s (or project’s) type is “high” and success arrives with unit hazard rate as long as the agent works. For comparison, under the fully optimal contract (as characterized in Theorem 1), the agent is paid $1.13 if they succeed during the silent phase, which lasts until $t^* = 0.4$, and paid $0.63 if a success arrives by $T = 1.64$.} The first panel depicts the work probability $p(t)$ (with the orange curve showing the full-transparency case as a benchmark), and the second panel the expected amount of time $Q(t)$ that the agent is asked to work after succeeding at...
Figure 1: The first two panels illustrate, respectively, the work probability, $p(t)$, and the expected amount of time that the agent is asked to work after succeeding at $t$, $Q(t)$, in an example where the agent is paid a fixed reward. The last panel depicts $q(s|t)$, the probability that the agent is asked to work at time $s$ conditional on having succeeded at $t$, for different values of $t$.

Like in the baseline model, this contract begins with a silent phase where the agent works with probability 1 (which here lasts until time 2.08); but once that phase ends, the principal begins to disclose information only gradually, thus holding onto an agent who succeeded during this initial phase for some extra length of time. (An agent who instead succeeds after the silent phase is immediately informed.) The principal would prefer to simply lower the reward and immediately inform the agent once the silent phase is over, but since that is not possible, the principal instead holds onto the agent for longer. Notice that since backward compounding makes ignorance more costly in the future, more and
more information is revealed as time passes, as reflected in a declining $Q(t)$.

The above contract is rather subtle as implementing the desired $Q(t)$ requires a complicated function $q(s|t)$ that lets the agent quit at an appropriate probabilistic rate (third panel). Fortunately, this complexity is avoided when the principal has control over the reward.

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. 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


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


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 in footnote 8 that restricting attention to deterministic terminal dates is without loss. To establish this, here, we prove our result 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 $T = \infty$, and (2) becomes

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

the agent’s expected payoff from obeying the recommendations is

$$\int_0^\infty r(s)R(s)f(s)ds - c \times \int_0^\infty p(s)ds,$$

and the principal’s objective is

$$\int_0^\infty p(s)ds - \int_0^\infty r(s)R(s)f(s)ds.$$

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 the continue to work through
\[ s \geq t + \Delta t \text{ equals} \]

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

Next, we characterize the following limit:

\[
\dot{p}(s|t) = \lim_{\Delta t \to 0} \frac{p(s|\omega^\Delta t) - p(s)}{\Delta t}
= \lim_{\Delta t \to 0} \left\{ r(s) \frac{F(s) - F(s - \Delta t)}{\Delta t}
- \frac{1}{\Delta t} \int_t^{t+\Delta t} r(u)f(u)q(s|u)du
- \int_{t+\Delta t}^s r(u) \frac{f(u) - f(u - \Delta t)}{\Delta t} q(s|u)du \right\},
\]

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’Hôpital’s rule and the Fundamental Theorem of Calculus, the limit in the second term almost everywhere exists and equals \(-r(t)f(t)q(s|t)\).\(^{25}\) As for the third term note that

\[
- \lim_{\Delta t \to 0} \int_t^{t+\Delta t} r(u) \frac{f(u) - f(u - \Delta t)}{\Delta t} q(s|u)du
= - \lim_{\Delta t \to 0} \int_t^s r(u) \frac{f(u) - f(u - \Delta t)}{\Delta t} q(s|u)du
= - \int_t^s r(u) \lim_{\Delta t \to 0} \frac{f(u) - f(u - \Delta t)}{\Delta t} q(s|u)du = - \int_t^s r(u)f'(u)q(s|u)du,
\]

where the third line follows from dominated convergence. Therefore,

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

\(^{25}\)Because \( q(\cdot|t) \) and \( r(\cdot) \) are monotone, the integrand is continuous almost everywhere. Therefore by the Fundamental Theorem of Calculus the derivative of the integral with respect to \( \Delta t \) exists almost everywhere and equals the value of the integrand.
Next, recall that incentive compatibility requires that
\[
U(t + Δt) - U(t) \leq \int_{t+Δt}^{∞} r(s) [f(s) - f(s - Δt)] R(s) ds + c × \int_{t+Δt}^{∞} \left[ p(s|ω_{t}^{Δt}) - p(s) \right] ds.
\]
Dividing both sides by Δt and taking the limit as Δt → 0 we have
\[
U'(t) \leq \int_{t}^{∞} r(s)f'(s)R(s) ds + c \lim_{Δt→0} \int_{t+Δt}^{∞} \frac{p(s|ω_{t}^{Δt}) - p(s)}{Δt} ds
\]
\[
⇔ cp(t) - r(t)f(t)R(t) \leq \int_{t}^{∞} r(s)f'(s)R(s) ds + c \int_{t}^{∞} \dot{p}(s|t)ds. (8)
\]
Finally, for any T, using the definition \( Q(t) := \int_{t}^{T} q(u|t)du \), letting \( r(t) \equiv I_{t≤T} \), and rearranging terms yields (IC).

A.2 Proof of Proposition 2

Fix a recommendation policy \{q(·|·), r(·)\}, and hence \( p(·) \) and \( \dot{p}(·, ·) \), and let \( R \) be any (locally) incentive compatible reward schedule. From (8), this implies that it satisfies
\[
r(t)R(t) \geq \frac{1}{f(t)} \left[ c \cdot p(t) - c \int_{t}^{∞} \dot{p}(s|t)ds - \int_{t}^{∞} r(s)f'(s)R(s) ds \right].
\]
Notice that the function \( Z^{1}(t) ≤ r(t)R(t) \) for all \( t \). It follows from the fact that \( f'(·) ≤ 0 \) and (7) that \( \dot{p}(s|t) ≤ f(s) - \int_{t}^{s} f'(u)du = f(t) \), and hence
\[
Z^{1}(t) ≥ c \frac{f(t)}{f(t)} \left( p(t) - \int_{t}^{∞} f(t) ds \right) = c \left( \frac{p(t)}{f(t)} - \frac{F(∞) - F(t)}{f(t)} \right) =: β(t)
\]
for all \( t \). ²⁶ Continuing in this manner, define for all \( k ≥ 2 \), the function \( Z^{k} \) by
\[
Z^{k}(t) = \frac{1}{f(t)} \left[ c \cdot p(t) - c \int_{t}^{∞} \dot{p}(s|t)ds - \int_{t}^{∞} Z^{k-1}(s)f'(s)ds \right].
\]
²⁶Note that \( F(∞) ≤ 1 \). The inequality is strict if, with some probability, success cannot be achieved.
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 $\beta$. Let $Z$ be the pointwise limit. By the dominated convergence theorem we have

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

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 $G(t) := \int_t^\infty r(s)f'(s)R^*(s)ds$ and $H(t) := c \cdot p(t) - c \int_t^\infty \dot{p}(s|t) ds$. Notice that $G'(t) = -r(t)f'(t)R^*(t)$ almost everywhere.\(^{27}\) We can then rewrite (8) as

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

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

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

Note that

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

\(^{27}\)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

\[ H_1(t) = \int_t^\infty r(s)f(s)ds, \]
\[ H_2(t) = -r(t)f(t) \int_t^\infty q(s|t)ds, \]
\[ H_3(t) = -\int_t^\infty \int_t^s r(u)f'(u)q(s|u)du ds, \]

and using (7) we have \( H(t) = cp(t) - c[H_1(t) + H_2(t) + H_3(t)]. \) Notice that \( H'_3(t) = -f'(t)H_2(t)/f(t), \) and by integrating by parts we have

\[ \frac{H_3(t)}{f(t)} - \int_t^\infty \frac{f'(s)H_3(s)}{f(s)^2}ds = \frac{H_3(t)}{f(t)} + \int_t^\infty \left( \frac{1}{f(s)} \right)' H_3(s)ds = \int_t^\infty \frac{f'(s)}{f(s)^2} H_2(s)ds. \]

Using integration by parts again, we have

\[ \frac{H_1(t)}{f(t)} - \int_t^\infty \frac{f'(s)H_1(s)}{f(s)^2}ds = \int_t^\infty r(s)ds. \]

Then (9) can be rewritten as

\[
\begin{align*}
    r(t)R^*(t) &= c \frac{p(t) - H_1(t) - H_2(t) - H_3(t)}{f(t)} - c \int_t^\infty \frac{f'(s)}{f(s)^2} [p(s) - H_1(s) - H_2(s) - H_3(s)] ds \\
    &= c \left[ \frac{p(t)}{f(t)} - \int_t^\infty \frac{f'(s)}{f(s)^2} p(s)ds - \int_t^\infty r(s)ds + r(t) \int_t^\infty q(s|t)ds \right].
\end{align*}
\]

For any \( T, \) substituting \( r(t) \equiv 1_{t \leq T} \) and \( q(s|t) = 0 \) for all \( s \geq T \) yields (4). \( \Box \)

### 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

\[
\int_0^\infty (1 - c)p(t) + cf(t) \int_t^\infty \frac{f'(s)}{f(s)^2} p(s)ds + cf(t) \int_t^\infty r(s)ds - cr(t)f(t) \int_t^\infty q(s|t)ds dt.
\]
To simplify the double integrals, we use that \[ \int_0^\infty a(t) \int_t^\infty b(s) \, ds \, dt = \int_0^\infty b(t) \int_0^t a(s) \, ds \, dt \]
for any integrable functions \( a(t) \) and \( b(t) \). In particular, we have

\[
\int_0^\infty f(t) \int_t^\infty \frac{f'(s)}{f(s)^2} p(s) \, ds = \int_0^\infty \frac{f'(t) F(t)}{f(t)^2} p(t) \, dt = - \int_0^\infty p(t) \Phi(t) \, dt,
\]

\[
\int_0^\infty f(t) \int_t^\infty r(s) \, ds = \int_0^\infty r(t) F(t) \, dt,
\]

\[
\int_0^\infty r(t) f(t) \int_t^\infty q(s|t) \, ds \, dt = \int_0^\infty \int_0^t r(u) f(u) q(t|u) \, du \, dt = \int_0^\infty p(t) - r(t) + r(t) F(t) \, dt,
\]

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

\[
\int_0^\infty p(t) \, dt - c \int_0^\infty p(t) (1 + \Phi(t)) - (r(t) - p(t)) \, dt,
\]

and substituting \( r(t) \equiv \mathbb{1}_{t \leq T} \) for any \( T \) yields (Obj). \( \square \)

### A.4 Proof of Proposition 3

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

\[
\max_{p(\cdot), r(\cdot)} \int_0^\infty p(t) \left[ 1 - 2c - c\Phi(t) \right] \, dt + c \int_0^\infty r(t) \, dt
\]

s.t. \( r(t) \left[ 1 - F(t) \right] \leq p(t) \leq 1 \) for all \( t \)

\[ 0 \leq r(t) \leq 1 \] for all \( t \) and non-increasing in \( t \).

Fixing an \( r(\cdot) \), we can maximize the objective pointwise for each \( t \), and noting that \( \Phi(t) \) is strictly increasing by assumption, it follows that the bracketed expression single-crosses
zero from above, 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 time 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

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

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 \( r \) must be non-increasing, there exists a \( T > t^* \) such that

\[
r(t) = \begin{cases} 
1 & \text{if } t \leq T \\
0 & \text{otherwise.}
\end{cases}
\]

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. \( \square \)

### 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

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

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

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

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).\(^{28}\) Turning to the first phase, if the agent shirks for \( \Delta \) units of time and otherwise follows the recommendations, then their ex-ante payoff is

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

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

---

\(^{28}\)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.