Dynamic Monitoring Design

Yu Fu Wong*†

June 1, 2023

This paper introduces flexible endogenous monitoring into dynamic moral hazard. A principal can commit to not only an employment plan but also the monitoring technology to incentivize dynamic effort from an agent. Optimal monitoring follows a Poisson process that produces rare informative signals, and the optimal employment plan features decreasing turnover. To incentivize persistent effort, the Poisson monitoring takes the form of "bad news" that leads to immediate termination. Monitoring is non-stationary: the bad news becomes more precise and less frequent. When persistent effort is not required, optimal incentive provision features a trial period of non-stationary monitoring, and a combination of Poisson bad news that leads to termination and Poisson good news that leads to tenure.

Keywords: dynamic monitoring, moral hazard, information design
JEL classification: D82, D83, D86

*I am indebted to Qingmin Liu for his guidance and encouragement. I am grateful to Yeon-Koo Che, Navin Kartik, and Elliot Lipnowski for helpful discussions. I thank Alex Bloedel, Martin Cripps, Laura Doval, Erik Madsen, Jacopo Perego, Debraj Ray, Balasz Szentes, Weijie Zhong, as well as conference participants for comments. The online appendix can be found on tinyurl.com/dmdonlineappendix.

†Department of Economics, Columbia University. Contact: yufu.wong@columbia.edu.

1 Introduction

Moral hazard is central to economics. Over the past twenty years, extensive research has greatly expanded our understanding of dynamic moral hazard in employment relationships, organizational structures, and regulatory policies. Most of this research, however, concentrates on incentive provision under exogenous monitoring technologies. In many applications, the monitoring technology itself plays a vital role in designing incentives. This monitoring design presents two interrelated questions: "which signals to acquire" and "how to adapt to acquired signals." For instance, a school principal can monitor a teacher by verifying their punctuality or directly observing their teaching. During the observation, the principal can evaluate the teacher's performance by recognizing well-prepared examples or detecting obvious errors. Depending on the observations, the principal can then decide whether to suspend or terminate the teacher.

The goal of this paper is to incorporate endogenous monitoring into dynamic incentive provision *without* restrictions on the monitoring technology. We explore a dynamic moral hazard model where a principal monitors an agent's effort by acquiring informative signals and adapts the incentive scheme to the acquired signals.

In the model, the principal (she) is committed to monitoring the agent's (his) binary private effort subject to a monitoring cost, and adapting future employment to past signals. The wage is fixed. We model flexible public monitoring by a Blackwell experiment on the agent's effort. We assume that the monitoring cost function satisfies likelihood-ratio separability (Assumption 1) and compound reduction (Assumption 2), with the relative entropy function as an example.

To isolate incentive provision from the choice of effort recommendation, we first analyze the main model where the principal must incentivize effort during employment. We show that optimal monitoring takes the form of Poisson bad news that leads to immediate termination, i.e., the agent is never employed again (Theorem 1). The optimal incentive scheme has minimal history dependence: it depends only on the signal history through the action history, i.e., the length of employment. Over time, the Poisson bad news becomes more precise and less frequent. The agent's continuation value increases as termination becomes a more effective incentive instrument.

The frequency decreases so quickly that the relationship continues indefinitely with positive probability. Our results are consistent with Mincer and Jovanovic (1981) who document that the length of employment explains half of the variation in termination, and the hazard rate of termination (turnover) is decreasing.

The main economic force is the information exposure to the agent. To provide incentives, the principal can either acquire more precise signals or adapt future employment more sensitively to the acquired signals. On one hand, more precise signals are more costly. On the other hand, more sensitive adaptation implies greater volatility in the agent's continuation value. The public signal therefore exposes more information about the agent's updated continuation value in the sense of second-order stochastic dominance, enabling the agent to devise more elaborate deviations from the recommended effort. It imposes a shadow cost on the principal as she faces more stringent incentive compatibility constraints to prevent such deviations.

The tradeoff between monitoring and information exposure drives the non-stationary incentive scheme. Optimal incentive provision equalizes the marginal cost of monitoring and that of information exposure (Equation (3)). As the continuation value increases, the Poisson bad news, which leads to termination, exposes more information about the updated continuation value. Therefore, the principal uses more precise signals to match the increased exposure. With more precise signals and a more punishing threat of termination, the frequency needed to incentivize effort decreases.

Our analysis highlights two salient features of flexible endogenous monitoring, which substantially differ from exogenous monitoring. The first is that the principal can probabilistically mix informative and uninformative signals. By keeping the incentive scheme unchanged, the uninformative signal does not expose any payoff-relevant information to the agent or create additional incentive compatibility constraints. We show that, as time period shortens, the principal's value increases and the mixed monitoring converges to Poisson monitoring (Lemma 1). This contrasts with exogenous monitoring models where players must adapt to a higher number of less precise signals. Their values decrease in these models because the players observe more signals that are *informative* about the continuation values and create more incentive compatibility constraints, as demonstrated by Abreu, Milgrom, and Pearce (1991) in partnership games.

The second feature is that the principal can acquire more precise signals by aggregating less precise ones *across periods*, which results in the minimal history dependence. Because more precise signals necessitates greater responses, the principal aggregates multiple signals until one of them is precise enough to prompt an immediate reaction, in the form of termination (Lemma 2). Acquiring this signal at a Poisson rate becomes optimal in continuous time. Because the Poisson signal warrants an immediate change of action, the signal history corresponds one-to-one with the action history. The endogenous signal aggregation differs from exogenous monitoring models where the principal must accumulate less precise signals *over time* for them to be precise enough to warrant a reaction. Before termination, the incentive scheme depends sensitively on the history of accumulated signals in addition to the history of actions.

The comparative statics of the optimal incentive scheme also differs from exogenous monitoring models (Proposition 2). I show that, when the principal and the agent becomes twice as patient, optimal monitoring slows down by a factor of two, i.e., both the frequency and the precision evolve at half speed. This is because slower monitoring and incentive backloading perfectly account for the decreased value of current effort in comparison to future streams, and so the value of the principal and that of the agent remain the same. This result contrasts with the folk theorem in exogenous monitoring models, where the principal can attain her first-best at the patient limit because fixed monitoring over an expanding horizon essentially reveals the agent's effort.

In addition, I show that, when the relationship is more profitable for the principal, the optimal incentive scheme is kickstarted as if the agent had already been employed for some time. The augmented length of employment increases the precision and decreases the frequency of Poisson bad news and therefore prevents early terminations. This result follows from the minimal history dependence as the agent's continuation value corresponds one-to-one to the length of employment, unlike exogenous monitoring models.

To incorporate effort recommendation, we analyze an extension where the principal does not need to incentivize effort during employment. This extension introduces shirking as a new incentive instrument. We show that the optimal incentive scheme can take one of four forms, all of which feature Poisson bad news that leads to immediate termination and/or Poisson good news that leads to immediate tenure, i.e., permanent shirking (Theorem 2). The Poisson signals increases in precision and decreases in frequency during a trial period of deterministic length.

The first two forms are up-or-out schemes where the agent is either terminated or tenured by the end of the trial period. The first form uses Poisson bad news in the trial period, and the agent's continuation value increases so much that he attains tenure at the end absent arrivals. Symmetrically, the second form uses Poisson good news in the trial period, and the continuation value decreases so much that the agent gets terminated at the end absent arrivals. The last two forms feature stationary two-sided Poisson monitoring (bad news that leads to termination and good news that leads to tenure) after the trial period. The third form uses Poisson bad news during the trial period, while the fourth form uses Poisson good news.

Related literature

Our paper incorporates flexible endogenous monitoring into dynamic moral hazard. Pioneering work including Rubinstein (1979) and Rogerson (1985) formulates the problem as repeated games with stationary exogenous monitoring. DeMarzo and Sannikov (2006) and Sannikov (2008) introduce the martingale representation approach in continuous-time models. A central insight from this literature is that the incentive scheme depends not only on the history of actions but also on the history of signals because players need to accumulate imprecise signals over time. With endogenous monitoring, however, we show that the optimally precise signal leads immediate termination and therefore the incentive scheme features minimal history dependence.

The role of monitoring in dynamic moral hazard has been discussed as early as in Abreu, Milgrom, and Pearce (1991), who show in a partnership game that more frequent observations shrink the set of equilibrium values due to information exposure. Sannikov and Skrzypacz (2010) show that Brownian monitoring and Poisson monitoring provide incentives in different ways in a continuous-time partnership game. Fudenberg and Levine (2007, 2009) and Sadzik and Stacchetti (2015) examine how the details of monitoring technology in discrete time affect the incentive provision at the continuous-time limit. A key result in this literature is the folk theorem that the first-best can be achieved at the patient limit under exogenous monitoring. We show, however, that the theorem does not hold under endogenous monitoring because optimal monitoring becomes less informative in response.

Recent developments in dynamic incentive provision have incorporated restricted forms of endogenous monitoring.[1] Liu (2011) endogenizes the number of past observations, as do Marinovic, Skrzypacz, and Varas (2018) and Varas, Marinovic, and Skrzypacz (2020) the timing of observations. Halac and Prat (2016) incorporate the frequency of Poisson good news, and likewise Piskorski and Westerfield (2016) and Orlov (2022) the frequency of Poisson bad news. Fahim, Gervais, and Krishna (2021) and Zeng (2022) include the precision of Brownian monitoring. In particular, Dai, Wang, and Yang (2021) endogenize the direction of conclusive Poisson news with fixed frequency and find that, when the agent's continuation value is low, *good news* monitoring is optimal because the conclusive bad news is not frequent enough to incentivize effort. By contrast, by endogenizing both the frequency and precision, we find that frequent but inconclusive *bad news* is optimal. This highlights how restrictions in monitoring impact qualitative predictions.

We adapt the belief-based approach in information design to the moral hazard model in order to overcome the problem of degenerate beliefs. Kamenica and Gentzkow (2011) introduce the Bayesian persuasion problem and belief-based approach. With this approach, Ely (2017), Ely and Szydlowski (2020), Hébert and Zhong (2022), and Koh and Sanguanmoo (2022) study dynamic persuasion. See Bergemann and Morris (2019) for a survey. The approach is also applied to rational inattention under posterior separable attention costs (Caplin and Dean, 2015). Ravid (2020) studies a game with rationally inattentive players and finds that unreasonable equilibria arise from the possibly degenerate belief about endogenous actions. To overcome this problem, we model monitoring by the distribution of likelihood ratio and the monitoring cost by an "experimental cost" (Denti, Marinacci, and Rustichini, 2022). We adapt properties of attention costs to the monitoring cost: namely, posterior separability to likelihood-ratio separability (Assumption 1) and sequential learning

[1] Georgiadis and Szentes (2020) and Li and Yang (2020) study optimal static incentive provision with flexible endogenous monitoring.

proofness (Bloedel and Zhong, 2020) to compound reduction (Assumption 2).[2]

The flexibility of dynamic monitoring in our model relates most closely to Zhong (2022) and Georgiadis-Harris (2021). They study flexible dynamic information acquisition before a one-off decision, and show the optimality of Poisson signals that lead to an immediate (change of) decision. The optimality of Poisson monitoring in our game-theoretic model results from pooling information exposed to the agent, which is not present in their decision models.

Finally, our model provides a moral hazard theory for the negative empirical relationship between the hazard rate of termination and the length of employment, complementing the existing search theory (Burdett, 1978; Jovanovic, 1984) and experience theory (Jovanovic, 1979).[3] See Gibbons and Waldman (1999) for a survey.

2 Dynamic monitoring model

We model the dynamic incentive provision problem in continuous time. The principal commits to an incentive scheme that consists of a monitoring technology to acquire public signals about the agent's private effort, and a contingent plan to adapt future employment to the signal history. The agent cannot commit and he chooses whether to exert effort when employed.

We introduce the timeline of the dynamic incentive scheme and then formulate the principal's design problem. Heuristically, the stage game at each instant $t \in [0, \infty)$ in continuous time proceeds as follows.

1. The principal publicly chooses whether to employ the agent for the instant, i.e., $h_t \in \{0, 1\}$. The stage game ends if she chooses not to employ, i.e., $h_t = 0$.

2. The principal publicly chooses a costly monitoring technology.

[2] Although the relative entropy attention cost vanishes as the prior approaches the degenerate belief that the agent exerts effort, it converges to the relative entropy monitoring cost when appropriately scaled.

[3] Exogenous monitoring models, such as Sannikov (2008), predict the hazard rate of termination to be zero initially and therefore do not decrease monotonically.

3. The agent privately chooses whether to exert costly effort, i.e., $a_t \in \{0, 1\}$.

4. The chosen monitoring technology generates a public signal about the current private effort.

2.1 Monitoring technology and monitoring cost

The principal chooses a monitoring technology that specifies how to monitor the agent's private effort based on past signals. We model the monitoring technology by a càdlàg martingale $\mathbf{\Gamma}$ with $\mathbf{\Gamma}_0 = 0$ that specifies the cumulative likelihood ratio $\mathbf{\Gamma}_t - \mathbf{\Gamma}_s$ during time interval $(s, t]$. For such a monitoring technology, we define the cumulative monitoring cost up to time t as a stochastic process

$$C_t(\mathbf{\Gamma}) := \limsup_{\Delta t \to 0} \sum_{m=1}^{\lceil t/\Delta t \rceil} C\left(1 + \mathbf{\Gamma}_{m\Delta t} - \mathbf{\Gamma}_{(m-1)\Delta t}\right) ,$$

where C is the monitoring cost function and $\lceil \cdot \rceil$ rounds up to the nearest integer. In this section, we shall elaborate on how continuous-time processes $\mathbf{\Gamma}$ and $C_t(\mathbf{\Gamma})$ model the monitoring technology and monitoring cost by introducing their discrete-time counterparts.

In discrete time, the principal monitors the agent's current effort by choosing a Blackwell experiment, which we operationalize as a distribution of the likelihood ratio subject to the Bayes rule. A Blackwell experiment specifies the distribution of signal $\mathbb{P}^a$ for each binary private effort $a \in \{0, 1\}$ of the current period. A signal is informative about private effort only through its likelihood ratio[4] $L := d\mathbb{P}^{a=0}/d\mathbb{P}^{a=1} \in (0, \infty)$. Moreover, a distribution of likelihood ratio corresponds to a Blackwell experiment if and only if it satisfies the Bayes rule, i.e., $\mathbb{P}^{a=1}[L] = 1$. Therefore, we represent the Blackwell experiment by likelihood ratio distribution $\mathbf{L} \in \Delta_1(0, \infty)$, the set of probability measures on $(0, \infty)$ with expectation one, and its signal by likelihood ratio $L \sim \mathbf{L}$.

The Blackwell experiment incurs a monitoring cost on the principal. The moni-

[4]We assume away perfectly informative signals $L = 0, \infty$, which would be infinitely costly.

toring cost function C maps each experiment $\mathbf{L}$ to its cost $C(\mathbf{L}) \in [0, \infty]$. We make two main assumptions on the non-parametric monitoring cost: *likelihood-ratio separability* and *compound reduction*. Our leading example is the relative entropy cost function $C(\mathbf{L}) = \mathbb{E}_{L \sim \mathbf{L}} [-\log(L) + L - 1]$.

The likelihood-ratio separability assumption states that the monitoring cost is linear in probability, i.e., the monitoring cost is a convex moment of the distribution of likelihood ratio.

Assumption 1 (Likelihood-ratio separability) *There exists a convex $\mathcal{C}^2$ function $c : (0, \infty) \to [0, \infty)$ such that $C(\boldsymbol{L}) = \mathbb{E}_{L \sim \mathbf{L}} [c(L)]$ for all $\mathbf{L} \in \Delta_1(0, \infty)$.*

The cost function c is convex so that the monitoring cost is monotonic in the Blackwell order; i.e., more precise monitoring is more costly. We assume the uninformative experiment is costless, i.e., $c(1) = 0$, and normalize[5] $c'(1) = 0$. Likelihood-ratio separability states that, for $\alpha \in (0, 1)$ and monitoring $\mathbf{L}_1, \mathbf{L}_2$, the monitoring cost of the probabilistic mixing equals the convex combination of the costs; that is, $C(\alpha \mathbf{L}_1 + (1 - \alpha)\mathbf{L}_2) = \alpha C(\mathbf{L}_1) + (1 - \alpha)C(\mathbf{L}_2)$. One interpretation of separability is that each signal L costs $c(L)$, and so the monitoring cost of an experiment equals the expected cost of the realized signal when the agent exerts effort.

The compound reduction assumption states that compound monitoring is not cheaper than reduced monitoring.

Assumption 2 (Compound reduction) *For all $\boldsymbol{L}_1 \in \Delta_1(0, \infty)$ with finite support, and $\boldsymbol{L}_2 : \mathrm{supp}(\boldsymbol{L}_1) \to \Delta_1(0, \infty)$, the monitoring costs satisfy*

$$C(\mathbf{L}_1) + \mathbb{E}_{L_1 \sim \mathbf{L}_1} [C(\mathbf{L}_2(L_1))] \geq C(\mathbf{L}_1 \times \mathbf{L}_2(\mathbf{L}_1)) ,$$

where $\mathbf{L}_2(\mathbf{L}_1)$ is the mixture distribution of $\mathbf{L}_1$ and $\mathbf{L}_2$.

The assumption concerns a hypothetical scenario where the principal monitors the same effort twice (Figure 1). Conditional on the first signal L_1, the principal inde-

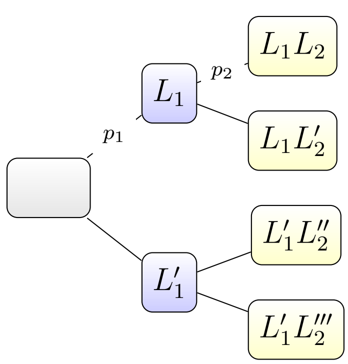
[5] We can normalize $c'(1)$ without changing the monitoring cost function C because all likelihood ratio distributions have expectation one.

pendently monitors the effort again to acquire the second signal L_2. For conditionally independent experiments, the likelihood ratios multiply to give the product likelihood ratio L_1L_2. The compound reduction assumption states that compound monitoring, which generates L_1 and then L_2, costs no less than reduced monitoring, which generates product likelihood ratio L_1L_2 directly. One interpretation is that the principal can monitor the same effort repeatedly, and so the monitoring cost is the reduced form of the least costly compound monitoring for a given distribution of product likelihood ratio. Note that compound monitoring costs the same as reduced monitoring under the relative entropy cost function; see Pomatto, Strack, and Tamuz (2023).

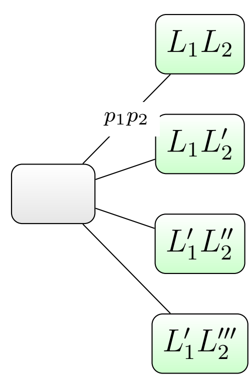
From the modeling standpoint, the compound reduction assumption helps us focus on how the agent's incentives drives the dynamics of the incentive scheme. It isolates economic agency friction from the technological monitoring friction, that the principal would lower monitoring costs by smoothing signal acquisition.

In addition to likelihood-ratio separability and compound reduction, we assume the Inada condition, $\lim_{L\to 0,\infty} c'(L)(L-1) - c(L) = \infty$, to guarantee the existence of optimal incentive scheme. See Mirrlees (1999) for an example of non-existence.

In discrete time, we represent the dynamic monitoring technology by the cumulative likelihood ratio and define the cumulative monitoring cost by partial sums. Let $\mathbf{L}_n$ denote the experiment in period n, which may depend on past signals. We write the cumulative likelihood ratio up to period n as a stochastic process, i.e., $\mathbf{\Gamma}_n := \sum_{m=1}^{n}(\mathbf{L}_m - 1)$. It is a discrete-time martingale because the likelihood ratio



(a) Compound monitoring.



(b) Reduced monitoring.

Figure 1: Compound monitoring and the corresponding reduced monitoring.

has expectation one. The cumulative monitoring cost is a stochastic process that can be written in terms of the martingale difference $\{\mathbf{\Gamma}_m - \mathbf{\Gamma}_{m-1} : m \leq n\}$. Formally,

$$\sum_{m=1}^{n} C(\mathbf{L}_m) = \sum_{m=1}^{n} C\left(1 + \mathbf{\Gamma}_m - \mathbf{\Gamma}_{m-1}\right) =: C_n(\mathbf{\Gamma}),$$

where $\mathbf{\Gamma}_0 := 0$. We write $C(\mathbf{L}) = \infty$ if $\mathbf{L} \notin \Delta_1(0,\infty)$. For any discrete-time martingale $\mathbf{\Gamma}$ such that $C_n(\mathbf{\Gamma})$ is almost surely finite for all n, there exists a sequence of experiments $\{\mathbf{L}_m\}$ adapted to past signals such that $\mathbf{\Gamma}_n := \sum_{m=1}^{n}(\mathbf{L}_m - 1)$.

Any (continuous-time) monitoring technology $\mathbf{\Gamma}$, such that $C_t(\mathbf{\Gamma})$ is finite almost surely, is the limit of a sequence of discrete-time counterparts $\{\mathbf{L}_{\Delta t,m} : m \leq \lceil t/\Delta t \rceil\}$ in that $\sum_{m=1}^{\lceil s/\Delta t \rceil}(\mathbf{L}_{\Delta t,m} - 1) \to \mathbf{\Gamma}_s$ in distribution for all $s \leq t$.

With a slight abuse of notation, we refer to $\mathbf{L}$ also by a monitoring technology.

2.2 Payoffs from employment, effort, and monitoring

The principal and the agent derive flow payoffs from employment, effort, and monitoring. Moral hazard arises as the principal earns revenue from the agent's costly private effort.

The agent derives utility from being employed and incurs an effort cost. When employed ($h_t = 1$), he derives a constant flow utility $u > 0$, which can be interpreted as a fixed wage. He also incurs an effort cost $k > 0$ in exerting private effort ($a_t = 1$). The flow payoff of unemployment is normalized to zero. The agent's von Neumann–Morgenstern payoff is then

$$\int_0^\infty re^{-rt} h_t(u - ka_t)dt,$$

where $r > 0$ is the discount rate common to both the principal and the agent. We assume that the agent prefers being employed and exerting effort to not being employed, i.e., $u - k > 0$.

The principal derives revenue from the agent's private effort and incurs the monitoring cost. When employing the agent ($h_t = 1$), she earns flow revenue $\pi > 0$ if the agent exerts effort ($a_t = 1$) and zero otherwise.[6] The principal cannot directly observe her revenue but only infer it through monitoring. She incurs incremental monitoring cost $dC_t(\Gamma)$ for monitoring Γ over a small time interval Δt. The principal's flow payoff is normalized to zero when she does not employ the agent. Her von Neumann–Morgenstern payoff is then

$$\int_0^\infty e^{-rt} h_t \left(r\pi a_t dt - dC_t(\Gamma) \right) .$$

2.3 Dynamic monitoring problem

The principal commits to a dynamic incentive scheme to maximize her expected payoff subject to the agent's incentive compatibility.

A dynamic incentive scheme $\mathcal{M}$ consists of

- a filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$, which satisfies the usual conditions;
- a monitoring technology $\mathbf{\Gamma}$, which is a càdlàg martingale with $\mathbf{\Gamma}_0 = 0$;
- predictable employment decision h and effort recommendation a.

The filtered probability space can be larger than the natural filtration of the monitoring technology to accommodate public randomizations. The probability law $\mathbb{P}$ corresponds to the case where the agent always exerts effort. Because the signal realization depends on private effort, we denote by $\mathbb{P}^{a'}$ the law under predictable effort[7] a'.

The principal's problem is to choose a dynamic incentive scheme $\mathcal{M}$ to maximize her expected payoff

$$\mathbb{E}^a \left[\int_0^\infty e^{-rt} h_t \left(r\pi a_t dt - dC_t(\Gamma) \right) \right]$$

[6] Revenue π can be interpreted as the profit net of the fixed wage payment in the main model, where the principal must incentivize effort during employment.

[7] We define $\mathbb{P}^{a'}$ as the extension to the change of measure $\left.\frac{d\mathbb{P}^{a'}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t^{a'}$, where $Z_t^{a'}$ is the stochastic exponential of the martingale $\int (1 - a'_t) d\Gamma_t$, i.e., $dZ_t^{a'} = Z_{t-}^{a'}(1 - a'_t) d\Gamma_t$ and $Z_0^{a'} = 1$. The extension exists and is unique by the Girsanov theorem.

subject to the agent's incentive compatibility constraint

$$a \in \max_{a'} \mathbb{E}^{a'} \left[\int_0^\infty re^{-rt} h_t (u - ka'_t) dt \right].$$

We denote the principal's value of incentive scheme $\mathcal{M}$ by $V(\mathcal{M})$. We restrict the principal's choice to continuous-time incentive schemes that can be approximated by discrete-time incentive schemes in value. Formally, there exists a sequence of discrete-time incentive schemes $\{\mathcal{M}_{\Delta t}\}$ that converge to the continuous-time incentive scheme in value, i.e., $\lim_{\Delta t \to 0} V_{\Delta t}(\mathcal{M}_{\Delta t}) = V(\mathcal{M})$, where $V_{\Delta t}(\mathcal{M}_{\Delta t})$ is the principal's value in $\mathcal{M}_{\Delta t}$. See Appendix A.1 for the discrete-time monitoring problem.

Remark 1 The restriction of convergence from discrete time provides a solid foundation to the tractable and abstract continuous-time model. As Fudenberg and Levine (2007, 2009) and Sadzik and Stacchetti (2015) point out, under exogenous monitoring with Brownian motion, continuous-time incentive schemes need not be a good approximation of discrete-time incentive schemes with short time periods. Our restriction rules out pathological cases such as the "infinite switches" equilibrium of Keller, Rady, and Cripps (2005).[8] A sufficient condition for the restriction is that monitoring technology is a simple or compound Poisson process of bounded frequency.

3 Optimal dynamic incentive scheme

To isolate incentive provision from the choice of effort recommendation, we first study the main model where the principal must incentivize effort when employing the agent, i.e., $h_t = 1 \implies a_t = 1$. We defer the extension that incorporates that choice to Section 4.

3.1 Main result

The optimal dynamic incentive scheme uses Poisson monitoring. A monitoring technology is Poisson if the cumulative likelihood ratio $\boldsymbol{\Gamma}$ is a compensated Poisson

[8]For the discrete-time version of that model, see Hörner, Klein, and Rady (2022).

process, parameterized by jump size $\Delta\Gamma_t \in (-1, \infty)$ and bounded frequency λ_t in its natural filtration. The likelihood ratio of Poisson arrival is $L_t := \Delta\Gamma_t + 1 \in (0, \infty)$ by our normalization. It incurs a flow monitoring cost $dC_t(\Gamma) = \lambda_t c(L_t)dt$. We call such monitoring *Poisson bad news* if $L_t > 1$, because the arrival is more likely when the agent does not exert effort, and we call it *Poisson good news* if $L_t < 1$. We say the Poisson news is more precise if L_t is further away from 1, because the arrival becomes a stronger signal.

We now present the optimal incentive scheme and then elaborate on its properties.

Theorem 1 *In the optimal incentive scheme,*

- *the principal monitors for Poisson bad news that leads to immediate termination;*
- *conditional on no arrival, the Poisson bad news increases in precision, decreases in frequency, and eventually increases in cost;*
- *moreover, the frequency decreases so quickly that the agent is employed indefinitely with positive probability.*

The optimal incentive scheme has minimal history dependence: it depends on the signal history only through the action history. This is because optimal monitoring is just precise enough to warrant the agent's immediate termination upon Poisson arrival. Should the agent remain employed, the signal history, which is an interval of non-arrival, is uniquely characterized by the action history of sustained effort. Therefore, the precision and frequency of optimal Poisson monitoring, conditional on no arrival, are deterministic functions of the length of employment.

Conditional on no arrival, optimal monitoring is non-stationary in that the precision increases and frequency decreases in the length of employment (Figure 2). The dynamics follow from backloading the agent's payoffs and the tradeoff between costly monitoring and information exposure.

The principal backloads the agent's payoffs so that his continuation value is increasing, because a more punishing termination eases incentive provision. For illustration, suppose that the optimal scheme monitors for stationary Poisson bad news that leads to immediate termination. Flexible monitoring allows us to consider a one-step deviation: a small and brief increase in frequency and decrease in precision that preserve incentive compatibility. To compensate for the additional risk of termination, the agent's continuation value increases in the absence of Poisson arrival. After the deviation, the principal reverts to stationary monitoring with decreased frequency, given the increased continuation value and thus threat of termination.

We argue that the deviation is profitable, which contradicts the optimality of stationary incentive schemes. The brief increase in the probability of termination is exactly compensated for by the decrease in the future, and therefore the principal receives the same expected revenue. The small deviation from the supposedly optimal scheme entails only a second-order increase in monitoring cost by the envelope argument. After the brief deviation, the agent's higher stakes allow the principal to incentivize effort with less frequent bad news, which reduces the monitoring cost. The reduction is of the first order because termination gives the minimum continaution value.

The increased continuation value implies a higher cost of information exposure.

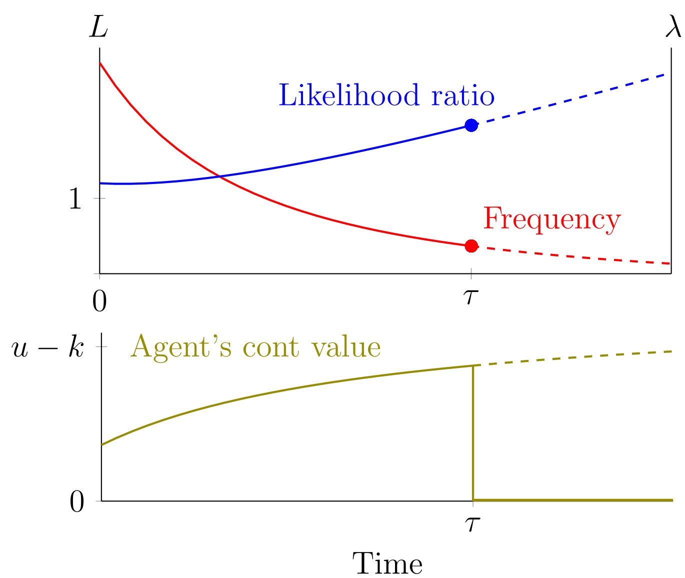


Figure 2: A realization of the optimal incentive scheme. A Poisson bad news arrives at τ.

When the agent's continuation value is higher, the public signal is more informative about the updated continuation value in the sense of second-order stochastic dominance, for fixed frequency λ. With arrival at frequency λ, the continuation value experiences a bigger drop to zero upon termination; with no arrival, the continuation value experiences a larger upward drift that is proportional to λ to compensate for the risk of termination. The improved information allows the agent to devise more elaborate deviations that are more costly for the principal to prevent.

The tradeoff between costly monitoring and information exposure results in the increasing precision, decreasing frequency, and eventually increasing monitoring cost. Because the principal can acquire more precise signals and adapt more sensitively to those signals, optimal incentive provision equalizes the marginal costs of incentive provision by monitoring and by sensitive adaptation, which leads to information exposure. The increasing exposure thus implies more precise monitoring over time. With more precise signals and more sensitive adaption, the frequency required to incentivize effort decreases. As the agent's continuation value approaches the maximum $(u-k)$, the threat of termination plateaus and thus the cost of the unboundedly precise Poisson monitoring eventually increases, by the Inada condition.[9]

The frequency of Poisson bad news decreases so quickly that the agent is employed and exerts effort indefinitely with positive probability. As the agent's continuation value approaches that of indefinite employment, i.e., $u-k$, the principal must use unboundedly precise but vanishingly infrequent bad news in order to keep the promised value to the agent. The precision increases quickly because the marginal cost of precision is bounded from above, i.e., $\lim_{L\to\infty} c'(L) < \infty$, by compound reduction. Therefore, the frequency needed to incentivize effort decreases quickly. The possibility of indefinite employment contrasts with exogenous monitoring models with bounded precision where the agent is eventually terminated.[10]

[9] One may argue that a decreasing frequency is the only way to realize the increasing continuation value. However, this argument is silent on the precision because of the increasing threat of termination.

[10] The sensitivity of the agent's continuation value to the likelihood ratio is bounded from below due to bounded precision. No shirking implies the existence of uniform finite time and positive probability such that the agent is terminated within that time time with at least that probability, regardless of his continuation value. The Borel–Cantelli lemma thus implies eventual termination.

Due to Spear and Srivastava (1987), the optimal incentive scheme admits a recursive formulation with the agent's continuation value W as the state variable (Figure 3). Mathematically, the continuation value W_t is the payoff the agent expects to attain after time t, i.e.,

$$W_t := \mathbb{E}_t \left[\int_t^\infty r e^{-r(s-t)} h_s \left(u - k a_s \right) ds \right].$$

With a slight abuse of notation, we denote by $V(W)$ the principal's value function, which is concave and attains zero upon termination, i.e., $V(0) = 0$. As W ranges from 0 to $u-k$, the likelihood ratio L of Poisson bad news increases from the uninformative 1 to the conclusive ∞, while the frequency λ decreases from ∞ to 0.

Remark 2 (Role of commitment) The principal's commitment to the incentive scheme is crucial for certain parameter values but not the others. When monitoring costs are high, the principal's value function can be negative for high continuation values due to very precise monitoring (Figure 3a). Commitment is essential because the principal could secure zero payoff by terminating the agent without bad news arrival. Conversely, when monitoring costs are low, the value function is always positive and therefore commitment is not necessary. The optimal incentive scheme can be suppoted as a Bayesian Nash equilibrium outcome by the trigger strategy: following any deviations by the principal, the principal never employs the agent again and the agent always shirks.

3.2 Overview of proof strategy

We highlight the main analytical challenges and outline our strategy to derive the optimal incentive scheme.

The dynamic monitoring problem can be decomposed into the choice of monitoring technology and the corresponding contingent plan, which present two analytical challenges. First, the infinite-dimensional monitoring technology prevents direct applications of the dynamic programming principle. In particular, we cannot establish a Hamilton–Jacobi–Bellman (HJB) equation by comparing a candidate HJB equation discrete-time Bellman equations because the agent's flow payoff contributes to incentive provision.[11] Second, for a fixed but a priori arbitrary monitoring technology, the evolution of the agent's continuation value is difficult to pin down because it depends on the cost of information exposure, which is in turn determined by future evolutions.

We derive the optimal incentive scheme in four steps by leveraging the discrete-time counterpart. First, the optimal discrete-time scheme does not use public randomization, which exposes information to the agent without providing incentives.

[11]By contrast, Zhong (2022) studies flexible dynamic information acquisition and manages to establish an HJB equation by comparing a jump-diffusion equation to discrete-time Bellman equations, because the decision maker in his model faces no incentive compatibility constraints and derives no flow payoffs.

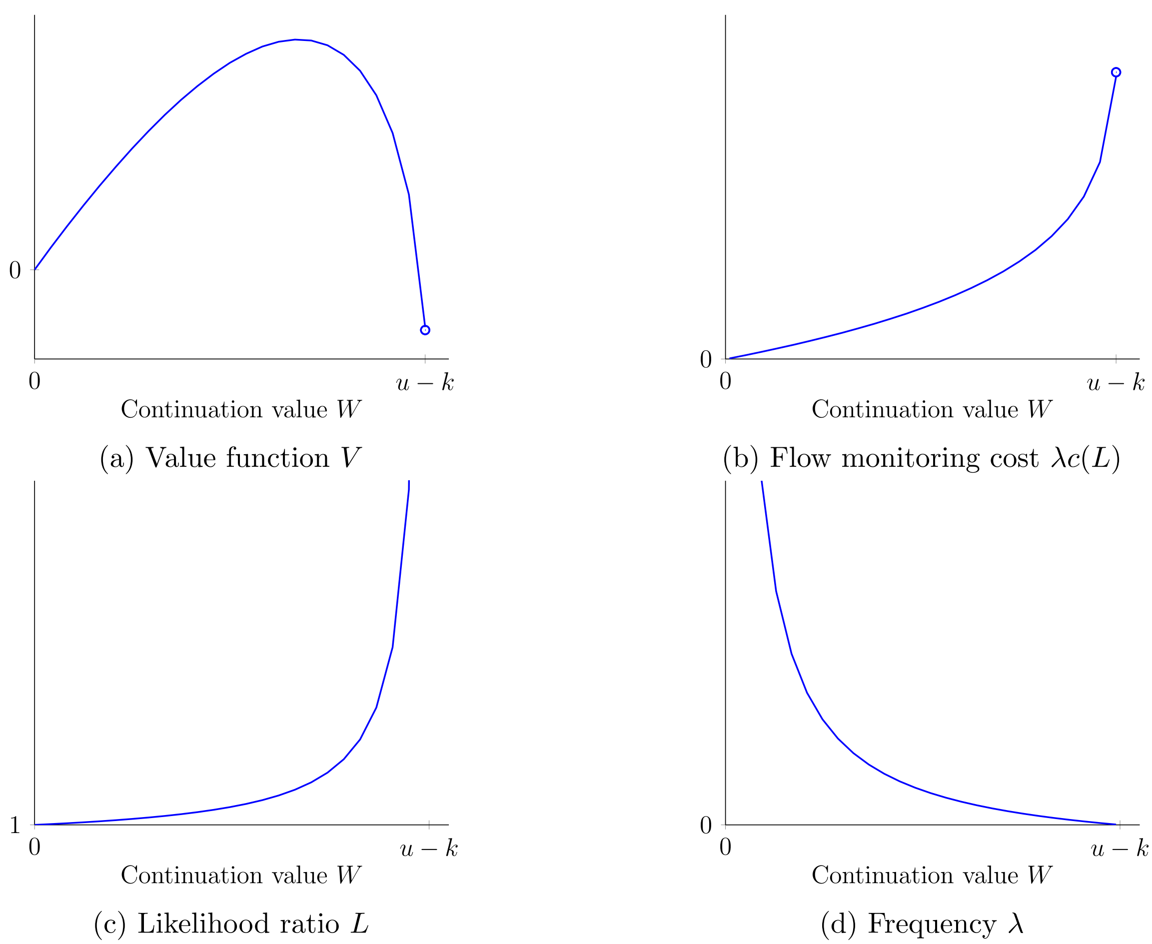


Figure 3: Value function and optimal monitoring.

Because temporary suspension is payoff-equivalent to randomized termination under common discounting, optimal signals lead to either termination or continued employment in the next period.

Second, Poisson monitoring is sufficient to maximize the principal's value. For any discrete-time scheme, we construct a continuous-time incentive scheme with compound Poisson monitoring by mixing informative and uninformative monitoring. The construction ensures that the agent is not exposed to additional information despite more frequent observations, enabling us to establish an HJB equation of Poisson monitoring for the value function.

Third, we show that the optimal discrete-time incentive scheme contains a signal that leads to immediate termination. If this were not the case, the principal could delay costly monitoring and limit information exposure by pooling the agent's information sets and aggregating signals across periods, which would reduce monitoring cost by compound reduction. By continuity, the immediate termination in discrete time implies that the principal's value function satisfies an HJB equation of immediate termination upon bad news arrival, and thus determines the evolution of the optimal incentive scheme.

Finally, we construct a candidate value function, verify its optimality, and derive the optimal incentive scheme by solving a first-order ordinary differential equation on the optimal likelihood ratio.

Remark 3 (Proofs that do not work) The binary Poisson signals may suggest simpler proof strategies, which unfortunately do not work. The two possibilities, termination and continued employment, do not necessitate binary signals via the revelation principle, which requires one signal for each of the continuum of continuation values. Relatedly, pooling signals that lead to continued employment would violate incentive compatibility, as the monitoring becomes less informative and the agent's continuation value becomes less sensitive to the signals.

In what follows, we shall elaborate on the sufficiency of Poisson monitoring and the existence of a signal that leads to immediate termination, because these two steps highlight the role of flexible monitoring and information exposure in dynamic incentive provision.

3.3 Sufficiency of Poisson monitoring

In the rich space of monitoring technologies, Poisson monitoring is sufficient to maximize the principal's value. As a result, the value function satisfies an HJB equation of Poisson monitoring.

Proposition 1 *The value function V is a viscosity solution to the HJB equation*

$$rv(W) = \sup_{\lambda,L,J} r\pi + r(W-u+k)v'(W) + \lambda\left(v(J) - v(W) - (J-W)v'(W)\right) - \lambda c(L)$$

subject to the instantaneous incentive compatibility constraint

$$\lambda(1-L)(J-W) = rk \tag{1}$$

for $W \in (0, u-k)$ and boundary condition $v(0) = 0$.

See Definition 3 in Appendix A.2 for the definition of viscosity solution.

Under Poisson monitoring, the incentive scheme specifies three control variables: monitoring frequency λ, likelihood ratio L, and the agent's continuation value J post-Poisson jump. The three must satisfy the agent's instantaneous incentive compatibility (IC), which is the continuous-time analog of the one-step deviation principle. Intuitively, Poisson bad news ($1-L<0$) should decrease the continuation value ($J<W$), and vice versa. Additionally, frequency λ must be high enough to overcome flow effort cost rk.

The HJB equation decomposes the principal's value into four components. The first is the flow revenue $r\pi$ from the agent's effort. The second reflects the principal's value due to the agent's expected continuation value, which grows at interest rate r and shrinks by flow payoff $r(u-k)$. The expected change translates to the principal's value by the marginal value $V'(W)$. The third term denotes the cost of information exposure. Mathematically, the cost equals the expected change in the principal's value due to the mean-preserving spread in the agent's continuation value. It either jumps from W to J with frequency λ or drifts in the opposite direction by $-\lambda(J-W)$. The fourth term is the cost of Poisson monitoring $\lambda c(L)$.

The optimal control variables results from an incentive–cost analysis. While providing incentives in IC (1), they incur the costs of monitoring and information exposure, as shown in the HJB equation. Due to expected utility and likelihood-ratio separability, the incentives and costs are linear in frequency λ. Therefore, the optimal likelihood ratio and jump maximize the incentive–cost ratio

$$(L^*, J^*) \in \underset{\substack{L,J \\ \text{s.t. } (1-L)(J-W)>0}}{\arg\max} \frac{(1-L)(J-W)}{-\left(V(J)-V(W)-V'(W)(J-W)-c(L)\right)} . \tag{2}$$

The optimal frequency then follows from binding IC.

For given jump J, the optimal likelihood ratio L trades off between monitoring and information exposure costs, as shown in the first-order condition (FOC)

$$c'(L)(L-1)-c(L) = -\left(V(J)-V(W)-(J-W)V'(W)\right) . \tag{3}$$

The principal can acquire more precise signals and adapt the continuation value more sensitively to such signals. Therefore, at the optimum, the marginal monitoring cost equals the marginal cost of information exposure. If the continuation value is more volatile, i.e., it jumps and drifts further away from W, the public signal is more informative about the continuation value in the sense of second-order stochastic dominance. Optimal incentive provision then implies higher precision which, together with the increased sensitivity, implies lower frequency needed to incentivize effort.

The key idea of Proposition 1 is that continuous-time incentive schemes with compound Poisson monitoring can replicate discrete-time ones in value. Compound Poisson monitoring, which mixes informative and uninformative monitoring, does not expose additional information or create new incentive compatibility constraints.

Analogous to Poisson monitoring, compound Poisson monitoring is a monitoring technology where $\mathbf{\Gamma}$ is a compound Poisson process of bounded frequency.

Lemma 1 (Poisson replication) *For any discrete-time incentive scheme, there exists a compound Poisson incentive scheme that gives the principal strictly higher value.*

The sufficiency of compound Poisson monitoring implies that of Poisson monitoring, because a compound Poisson process is a convex combination of compensated Poisson processes. Due to the temporally separable expected utility and the likelihood-ratio separability, each Poisson component contributes linearly to the principal's objective and the agent's incentives. In continuous time, the compensation drift becomes linear, which is the main motivation for the continuous-time formulation. As a result, considering Poisson monitoring suffices to maximize the principal's value.

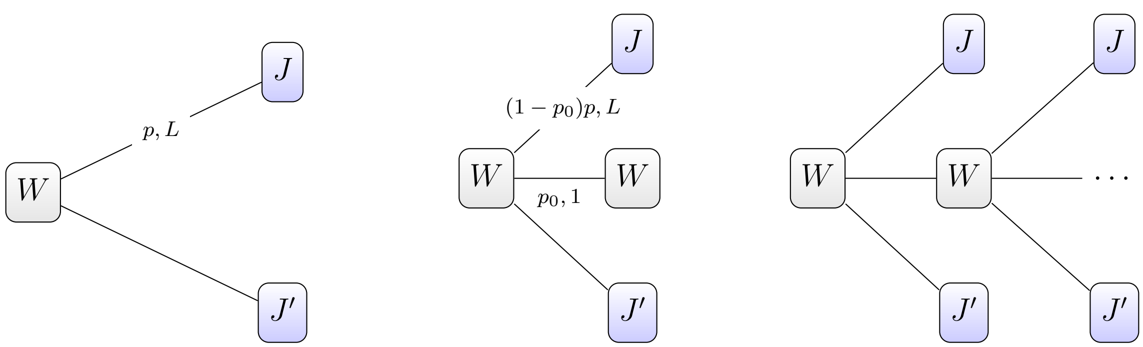
We show that the principal can increase her value by mixing informative and uninformative monitoring in shorter time periods. In a Δt-incentive scheme with continuation value W, the principal uses informative monitoring $\mathbf{L}$ and continuation value $\mathbf{J}$ to satisfy the Δt-IC constraint (Figure 4a); formally,

$$e^{-r\Delta t}\mathbb{E}_{\mathbf{L},\mathbf{J}}\left[(1-L)(J-W)\right] = (1-e^{-r\Delta t})k\,.$$

Analogous to the continuous-time IC (1), the covariation between the likelihood ratio and the continuation value equals the effort cost for the Δt-period. When the period is half as long, i.e., $\Delta t' := \Delta t/2$, the effort cost and revenue are also approximately halved due to temporally separable preferences.[12] The principal can halve the incentive by mixing the uninformative monitoring ($L \equiv 1, J \equiv W$) with probability $p_0 \approx 1/2$, and the informative monitoring $(\mathbf{L}, \mathbf{J})$ the rest of the time; formally,

$$e^{-r\Delta t'}\left(p_0 \times 0 + (1-p_0) \times \mathbb{E}_{\mathbf{L},\mathbf{J}}\left[(1-L)(J-W)\right]\right) = (1-e^{-r\Delta t'})k\,.$$

[12] They are slightly more than half because of the convex exponential discounting.



(a) Incentive scheme in $\Delta t > \Delta t'$.

(b) Mixing in the uninformative monitoring in $\Delta t'$.

(c) Iteratively mixing in the uninformative monitoring in $\Delta t'$.

Figure 4: Replicating an incentive scheme's value in shorter time periods.

The mixed incentive scheme replicate the expected revenue to the principal, who also shares the same discount rate. The monitoring cost decreases by slightly more than half because the principal adapts to the signals with a shorter delay $\Delta t'$, reducing the required frequency by more than half ($p_0 < 1/2$).

We show the sufficiency of compound Poisson monitoring by construction (Figure 4c). At initial continuation value W, the principal mixes informative and uninformative monitoring until the first informative signal arrives (Figure 4c). This process repeats itself at each new continuation value J. Because discrete time is countable, the iteration constructs a dynamic incentive scheme for shorter periods.[13] The probability of an informative signal becomes proportional to the effort cost and thus the period length, which is the defining feature of compound Poisson monitoring. Therefore, the construction converges to a compound Poisson incentive scheme.

Endogenous monitoring, which allows mixing with the uninformative monitoring, can increase the principal's value in shorter time periods, in contrast with exogenous monitoring models (Abreu, Milgrom, and Pearce, 1991). Despite more opportunities to deviate, the agent faces the same incentive scheme following any number of uninformative signals.

3.4 Signal leading to immediate termination

We shall show that the optimal employment plan terminates the agent immediately upon Poisson bad news arrival. The intuition is that the optimal discrete-time incentive scheme must react to some signal immediately in the form of termination, in order to justify the costs of monitoring and information exposure.

Lemma 2 (Immediate reaction) *Any optimal discrete-time incentive scheme contains a signal that leads to immediate termination.*

Even though periods beyond the next can adapt to the current signal, the optimal incentive scheme must react to some signal, through immediate termination.

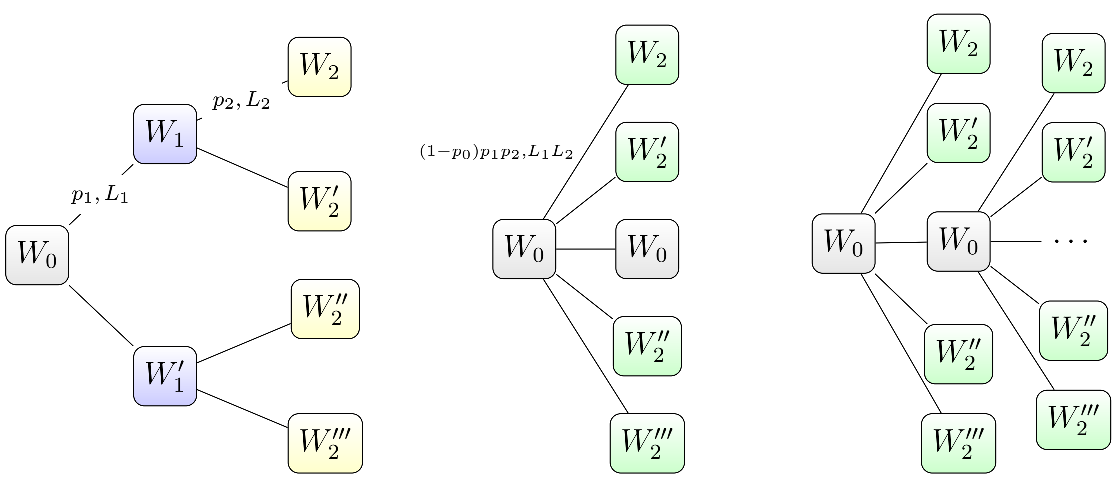
We prove Lemma 2 by contradiction. Suppose that, for some initial continuation

[13] We show that optimal monitoring in discrete time consists of finitely many signals.

value W_0, every signal L_1 leads to a corresponding W_1, at which the principal employs the agent for one more period and acquires a second signal L_2 that leads to W_2 (Figure 5a). We construct an alternative incentive scheme that yields a strictly higher value by delaying costly monitoring and information exposure (Figure 5b). The scheme mixes informative and uninformative monitoring: with probability p_0, it conducts uninformative monitoring that leads to the same continuation value W_0; with probability $1-p_0$, it conducts the informative, reduced monitoring that generates signal L_1L_2 that leads to W_2 (Figure 5b). Compound reduction is possible across the two periods because the effort choices are the same. Because the reduced monitoring can incentivize the agent for two periods, its probability is roughly $1 - p_0 \approx 1/2$.

The alternative scheme is incentive compatible and offers a strictly higher value to the principal by lowering the monitoring cost. Incentive compatibility results from the choice of p_0 and the pooling of the agent's information at the intermediate W_1's, which exposes the agent to less information.

The principal derives the same revenue in expectation because of the common discount rate, but the monitoring cost decreases for two reasons. First, the cost of reduced monitoring weakly decreases due to compound reduction. Second, the alternative scheme delays the first costly signal L_1 by one period in expectation. It



(a) Incentive scheme without immediate reactions.

(b) Compound reduction in one period.

(c) Compound reduction in multiple periods.

Figure 5: Immediate reaction and compound reduction in dynamic incentive schemes.

acquires the product signal L_1L_2 and reacts to it at W_2 in just one period, rather than two, as shown in Figure 5c. This delay in costly monitoring strictly decreases the discounted cost.

The optimal incentive scheme under endogenous monitoring has minimal history dependence. This is because an optimal signal in discrete time must be precise enough to warrant a change of action by Lemma 2, and acquiring such signal at a Poisson rate is optimal in continuous time. In contrast, with exogenous monitoring, the signals may not be precise enough to warrant a reaction and so the principal can only accumulate these signals over time, adapting the scheme to the signal history in addition to the action history.

3.5 Comparative statics

We present the comparative statics of the optimal incentive scheme in terms of the Poisson monitoring technology $\mathbf{\Gamma}$ keeping in mind that the optimal employment plan terminates the agent upon one bad news arrival. We say the optimal incentive scheme slows down by a factor of $\rho > 1$ if the new monitoring $\tilde{\mathbf{\Gamma}}_t$ shares the same law as $\mathbf{\Gamma}_{t/\rho}$; and we say it is kickstarted if $\tilde{\mathbf{\Gamma}}_t$ shares the same law as $\mathbf{\Gamma}_{t+T} - \mathbf{\Gamma}_T$ conditional on no arrival before $T > 0$. Moreover, we say the monitoring cost scales down if the new monitoring cost function $\tilde{c}$ equals γc for some $\gamma < 1$.

Proposition 2

1. *When the discount rate decreases from r to r', the optimal incentive scheme slows down by a factor of r/r'.*

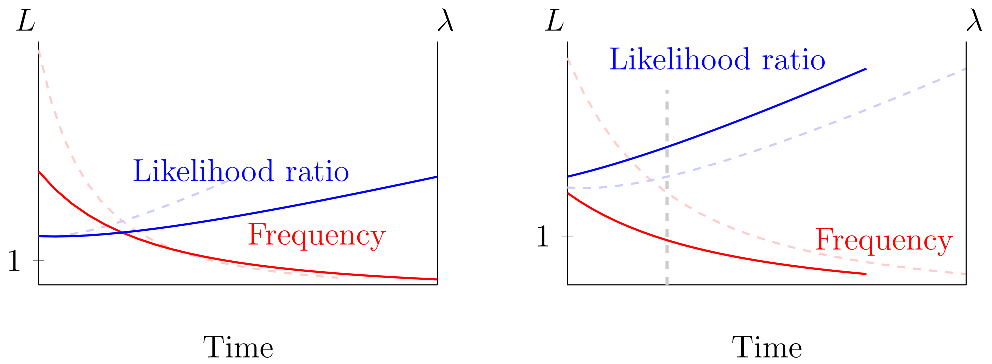
2. *When either the revenue increases or the monitoring cost scales down, the optimal incentive scheme is kickstarted.*

The first result is that, when both the principal and the agent become twice as patient, the optimal incentive scheme evolves at half speed (Figure 6a). This implies that the initial frequency of Poisson bad news is reduced by half, while the probability of indefinite employment and the values of the principal and the agent remain the same. The intuition is that both players find the current effort half as valuable when compared to the future stream. Incentivizing effort requires Poisson monitoring at half the frequency, which costs half as much by likelihood-ratio separability. The lower risk of termination thus implies that the agent's continuation value evolves more slowly. Mathematically, the HJB equation and incentive compatibility constraint in Proposition 1 are satisfied if the frequency λ is proportional to the discount rate while the value function V and the precision L depend only on the continuation value W.

The principal's value is bounded away from the first-best even at the patient limit, which contrasts the folk theorem in exogenous monitoring models. This is because, when the players have longer time horizons, the principal slows down costly monitoring by acquiring the same amount of information over the expanded horizon.

The second result is that, when the relationship becomes more profitable for the principal either through an increase in revenue or an equivalent decrease in monitoring cost, the optimal incentive scheme is kickstarted as if the agent had been employed for some time (Figure 6b). The scheme then evolves in the same way following the augmented length of employment. Given that the frequency of Poisson bad news decreases over the length of employment, the agent enjoys higher value and the probability of indefinite employment increases.

This result can be understood by decomposing the principal's problem into two optimization problems. The first minimizes the monitoring cost over all incentive schemes subject to the agent's initial continuation value, and the second maximizes the principal's value, which consists of the expected revenue net of the monitoring cost, over the agent's initial continuation value. While the revenue increases, the monitoring cost as a function of the agent's continuation value remains the same. To maximize her own value, the principal starts the incentive scheme with a higher initial continuation value to the agent. From there, the evolution is governed by the same monitoring cost. Note that the comparative static assumes the simple form because the agent's continuation value corresponds one-to-one to the length of employment, due to the minimal history dependence.



(a) Increased patience.

(b) Increased revenue/decreased monitoring cost.

Figure 6: Comparative statics of the optimal incentive scheme. Dashed lines represent the original scheme.

4 General effort recommendation

To incorporate effort recommendation into incentive provision, we study an extension where the principal can recommend shirking during employment. Minimal history dependence remains robust: conditional on the current effort, the signal history is determined by the action history because the agent is terminated *or tenured* upon Poisson arrival. Moreover, the decreasing turnover generalizes to a monotonic hazard rate of termination and tenure.

For some parameter values, the optimal scheme uses stationary two-sided Poisson monitoring, i.e., a compound Poisson monitoring of two possible arrivals, good and bad news, with stationary frequencies and precisions.

Theorem 2 *Depending on model parameters, the optimal incentive scheme takes one of four forms. All four forms feature Poisson monitoring, the possibility of tenure, and a trial period of deterministic length, during which Poisson monitoring becomes more precise and less frequent.*

1. *The first form monitors for Poisson bad news that leads to immediate termination during the trial period and, absent arrivals, tenures the agent at the end.*

2. *The second form monitors for Poisson good news that leads to immediate tenure*

during the trial period and, absent arrivals, terminates the agent at the end.

3. *The third form monitors for Poisson bad news that leads to immediate termination during the trial period, and switches to stationary two-sided Poisson monitoring with bad news that leads to termination and good news that leads to tenure.*

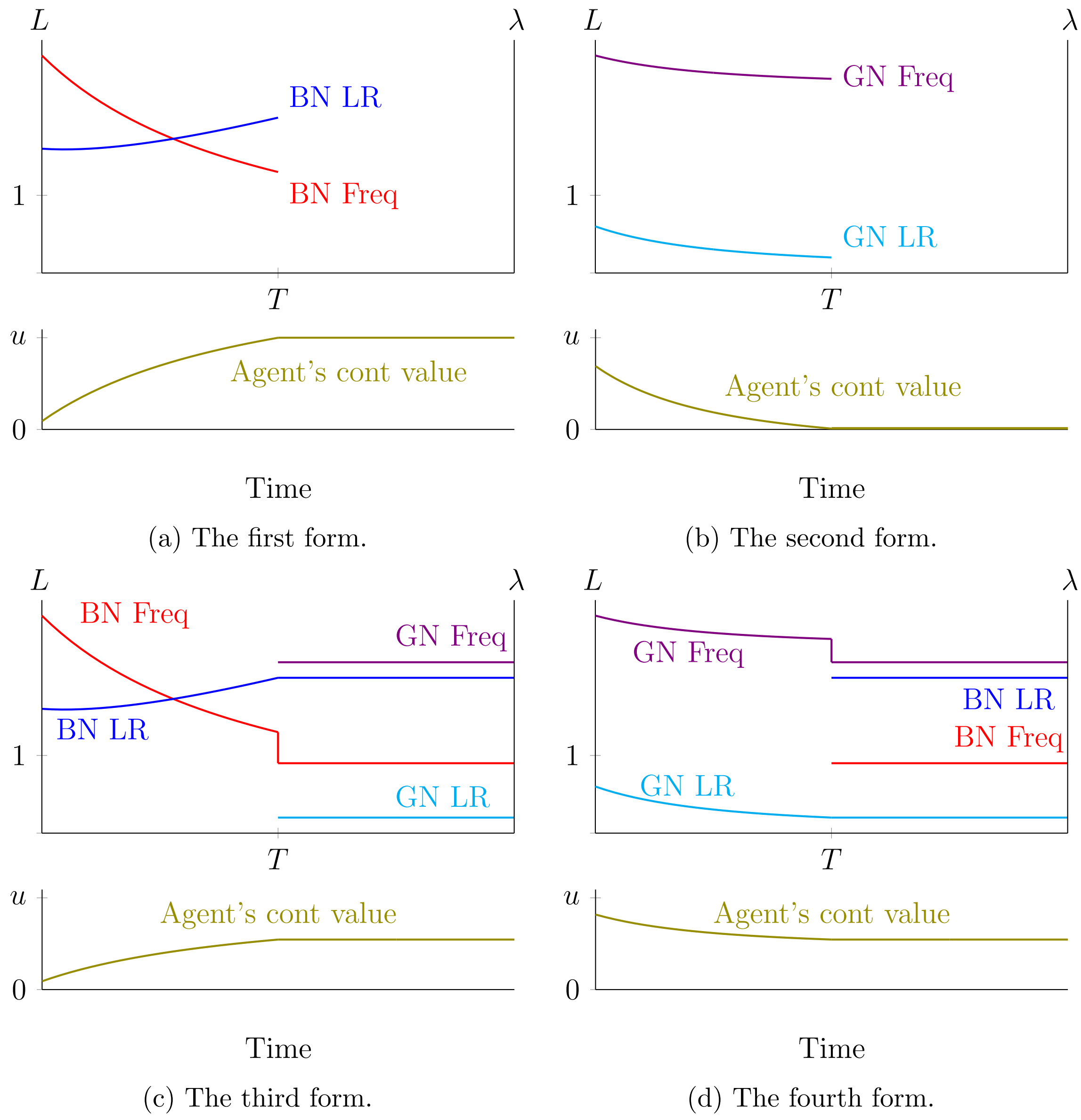
4. *The fourth form monitors for Poisson good news that leads to immediate tenure during the trial period, and switches to stationary two-sided Poisson monitoring with bad news that leads to termination and good news that leads to tenure.*

The four forms of optimal incentive scheme are shown in Figure 7. Each of them features Poisson monitoring and minimal history dependence because, as in the main model, the principal can mix information and uninformative monitorings and aggregate signals across time periods. The new aspect is that tenure can also be an optimal reaction to Poisson arrivals; however, temporary shirking is suboptimal because it is payoff-equivalent to randomized tenure.

During the trial period, Poisson monitoring increases in precision and decreases in frequency, which results in the decreasing hazard rate of termination/tenure. In the case of Poisson good news, the agent's continuation value decreases absent arrivals because a more rewarding tenure eases future incentive provision. As the good news exposes more information, optimal monitoring increases in precision to match the higher cost of information exposure. The frequency needed to incentivize effort decreases with a more rewarding tenure and more precise signals. The symmetric argument applies to the case of Poisson bad news.

The first two forms tenure or terminate the agent by the end of the trial period, resembling up-or-out incentive schemes in accounting, consulting, and law firms and in academia. The first form applies when Poisson bad news that leads to termination remains optimal as the agent's continuation value increases to u (Figure 7a). Symmetrically, the second form applies when Poisson good news that leads to tenure remains optimal as the agent's continuation value decreases to 0 (Figure 7b).

The last two forms transitions to stationary two-sided Poisson monitoring after the trial period. In the third form (Figure 7c), the continuation value increases to a threshold absent bad news during the trial period. Two-sided monitoring begins when good news monitoring becomes equally optimal at the threshold. The bad news's precision is continuous over time due to the continuous cost of information exposure. However, its frequency and therefore the hazard rate of termination decrease



(a) The first form.

(b) The second form.

(c) The third form.

(d) The fourth form.

Figure 7: The likelihood ratio (LR) and frequency (Freq) of Poisson good news (GN) and bad news (BN) over time in the four forms of optimal incentive schemes, where T denotes the duration of the trial period.

discontinuously because the good news also provide some incentives. The fourth form is symmetric (Figure 7d).

The stationary monitoring can be interpreted as chattering between good and bad news. When the continuation value is just below the threshold, the principal monitors for Poisson bad news and so, absent arrivals, the continuation value increases above the threshold. Now that the continuation value is above the threshold, the principal monitors for Poisson good news and so, absent arrivals, the continuation value decreases below the threshold. The chattering continues until one of the Poisson news arrives.

We analyze how the optimal form depends on model parameters numerically by value function iteration (Figure 8). For the relative entropy cost function, the parameter space divides into two regions according to the agency cost and the monitoring cost. The first form is optimal when the costs are high and the third form is optimal when the costs are low. The fourth form become optimal in the low cost region if we augment the model by a sufficiently high initial outside option for the agent. The second form does not appear to be optimal for any parameter values.

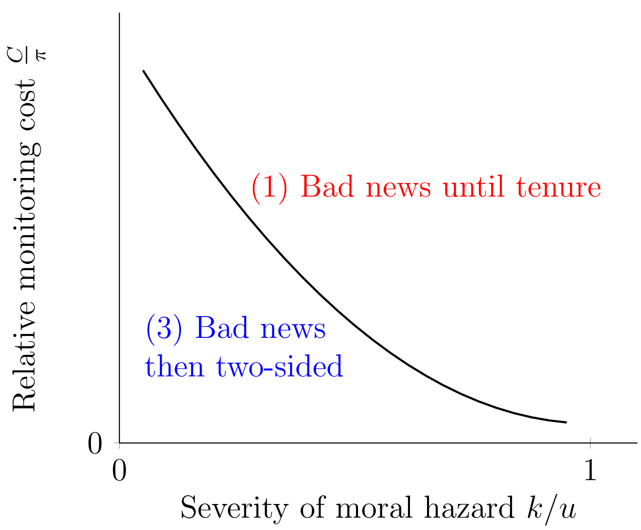


Figure 8: Numerical analysis of the optimal form over model parameters, for the relative entropy cost function.

5 Conclusion

Monitoring is central to incentive provision. This paper provides a dynamic incentive provision framework that allows flexible endogenous monitoring, and characterizes the optimal incentive scheme. We find that optimal incentive provision takes a simple form: Poisson bad news that leads to immediate termination. The agent's backloaded payoff offers a moral hazard theory for the decreasing hazard rate of termination in employment relationships. Our comparative statics results provide testable predictions for this theory.

We explore how endogenous monitoring changes our understanding of dynamic incentive provision. While the insight of adapting future actions to past signals still applies, the intricate signal dependence found in exogenous monitoring models does not, because it is absorbed by signal aggregation over time. Minimal history dependence arises from the compound reduction assumption which isolates agency friction from monitoring friction. In the presence of monitoring and other frictions in applications, we believe that endogenous monitoring will remain a critical incentive component. Therefore, our model stands as a benchmark for dynamic monitoring design.

6 Bibliography

Dilip Abreu, Paul Milgrom, and David Pearce. Information and timing in repeated partnerships. *Econometrica*, pages 1713–1733, 1991.

Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.

Alexander W. Bloedel and Weijie Zhong. The cost of optimally-acquired information. *Working Paper*, 2020.

Kenneth Burdett. A theory of employee job search and quit rates. *American Economic Review*, 68(1):212–220, 1978.

Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015.

Liang Dai, Yenan Wang, and Ming Yang. Dynamic contracting with flexible monitoring. *Working Paper*, 2021.

Peter M. DeMarzo and Yuliy Sannikov. Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance*, 61(6):2681–2724, 2006.

Tommaso Denti, Massimo Marinacci, and Aldo Rustichini. Experimental cost of information. *American Economic Review*, 112(9):3106–23, 2022.

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

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

Arash Fahim, Simon Gervais, and R. Vijay Krishna. Equity prices and the dynamics of corporate governance. *Working Paper*, 2021.

Drew Fudenberg and David K. Levine. Continuous time limits of repeated games with imperfect public monitoring. *Review of Economic Dynamics*, 10(2):173–192, 2007.

Drew Fudenberg and David K. Levine. Repeated games with frequent signals. *Quarterly Journal of Economics*, 124(1):233–265, 2009.

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

Alkis Georgiadis-Harris. Information acquisition and the timing of actions. *Working Paper*, 2021.

Robert Gibbons and Michael Waldman. Careers in organizations: Theory and evidence. *In O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics*, 3: 2373–2437, 1999.

Marina Halac and Andrea Prat. Managerial attention and worker performance. *American Economic Review*, 106(10):3104–3132, 2016.

Benjamin Hébert and Weijie Zhong. Engagement maximization. *Working Paper*, 2022.

Johannes Hörner, Nicolas Klein, and Sven Rady. Overcoming free-riding in bandit games. *Review of Economic Studies*, 89(4):1948–1992, 2022.

Boyan Jovanovic. Job matching and the theory of turnover. *Journal of Political Economy*, 87(5):972–990, 1979.

Boyan Jovanovic. Matching, turnover, and unemployment. *Journal of Political Economy*, 92(1):108–122, 1984.

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

Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

Andrew Koh and Sivakorn Sanguanmoo. Attention capture. *Working Paper*, 2022.

Anqi Li and Ming Yang. Optimal incentive contract with endogenous monitoring technology. *Theoretical Economics*, 15(3):1135–1173, 2020.

Qingmin Liu. Information acquisition and reputation dynamics. *The Review of Economic Studies*, 78(4):1400–1425, 2011.

Iván Marinovic, Andrzej Skrzypacz, and Felipe Varas. Dynamic certification and reputation for quality. *American Economic Journal: Microeconomics*, 10(2):58–82, 2018.

Jacob Mincer and Boyan Jovanovic. Labor mobility and wages. In *S. Rosen (ed.), Studies in Labor Markets*, pages 21–64. University of Chicago Press, 1981.

James A. Mirrlees. The theory of moral hazard and unobservable behaviour: Part i. *Review of Economic Studies*, 66(1):3–21, 1999.

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

Tomasz Piskorski and Mark M. Westerfield. Optimal dynamic contracts with moral hazard and costly monitoring. *Journal of Economic Theory*, 166:242–281, 2016.

Luciano Pomatto, Philipp Strack, and Omer Tamuz. The cost of information: The case of constant marginal costs. *American Economic Review*, 113(5):1360–1393, 2023.

Doron Ravid. Ultimatum bargaining with rational inattention. *American Economic Review*, 110(9):2948–63, 2020.

William P. Rogerson. Repeated moral hazard. *Econometrica*, pages 69–76, 1985.

Ariel Rubinstein. An optimal conviction policy for offenses that may have been committed by accident. In *S. Brams, A. Schotter, and G. Schwodiauer (eds.), Applied Game Theory*, pages 406–413. Springer, 1979.

Tomasz Sadzik and Ennio Stacchetti. Agency models with frequent actions. *Econometrica*, 83(1):193–237, 2015.

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

Yuliy Sannikov and Andrzej Skrzypacz. The role of information in repeated games with frequent actions. *Econometrica*, 78(3):847–882, 2010.

Halil Mete Soner. Optimal control of jump-Markov processes and viscosity solutions. In *W. Fleming and P.-L. Lions (eds.), Stochastic Differential Systems, Stochastic Control Theory and Applications*, pages 501–511. Springer, 1988.

Stephen E. Spear and Sanjay Srivastava. On repeated moral hazard with discounting. *Review of Economic Studies*, 54(4):599–617, 1987.

Felipe Varas, Iván Marinovic, and Andrzej Skrzypacz. Random inspections and periodic reviews: Optimal dynamic monitoring. *Review of Economic Studies*, 87(6): 2893–2937, 2020.

Yishu Zeng. *On Information Policy Design and Strategic Interactions*. PhD thesis, 2022.

Weijie Zhong. Optimal dynamic information acquisition. *Econometrica*, 90(4):1537–1582, 2022.

A Proof of Theorem 1

In this section, we show the optimality of Poisson monitoring and immediate termination. Omitted proofs can be found in the online appendix.

A.1 Discrete-time incentive provision

We formulate the discrete-time monitoring design problem and study the optimal incentive scheme. For $\Delta t > 0$, the stage game in discrete time follows the timeline in Section 2 in each period $n \in \{1, 2, ...\}$.

A discrete-time incentive scheme consists of complete filtered probability space $(\Omega, \{\mathcal{F}_n\}_n, \mathbb{P})$, predictable employment decision h and effort recommendation a, and monitoring technology $\mathbf{\Gamma}$ which is a martingale with $\mathbf{\Gamma}_0 = 0$. We write the monitoring in period n as $\mathbf{L}_n := \mathbf{\Gamma}_n - \mathbf{\Gamma}_{n-1} + 1$.

The principal's problem is to offer an incentive scheme to maximize her value

$$\mathbb{E}^a \left[\sum_{n=1}^{\infty} e^{-(n-1)r\Delta t} h_n \left(\left(1 - e^{-r\Delta t}\right) \pi a_n - C\left(\mathbf{L}_n\right) \right) \right]$$

subject to the agent's incentive compatibility constraint

$$a \in \max_{a'} \mathbb{E}^{a'} \left[\sum_{n=1}^{\infty} e^{-(n-1)r\Delta t} \left(1 - e^{-r\Delta t}\right) h_n \left(u - k a'_n\right) \right].$$

Without loss of optimality, we restrict attention to incentive schemes that do not monitor the effort when it is not recommended, e.g., $a_n = 0 \implies \mathbf{L}_n = \boldsymbol{\delta}_1$.

Following Spear and Srivastava (1987), we denote by $V_{\Delta t}(W)$ the principal's value as a function of the agent's continuation value

$$W_n := \mathbb{E}^a_{n-1} \left[\sum_{m=n}^{\infty} e^{-(m-n)r\Delta t} h_m \left(1 - e^{-r\Delta t}\right) \left(u - k a_m\right) \right].$$

In the main model, the principal must incentivize effort during employment, i.e.,

$h_n = 1 \implies a_n = 1$.

In Section S1.1, we show that the optimal incentive scheme can be constructed iteratively by taking the maximizer $(\mathbf{L}_n, \mathbf{J}_n)$ of the Bellman equation (Lemma S2). The maximizer is supported on at most four points by Lemma S3. The iteration continues until $N := \inf\{n : W_n = 0\}$ when the agent is terminated, i.e., $h_n = a_n = \mathbf{1}_{n<N}$. We refer to this optimal incentive scheme simply by $\{\mathbf{L}_n, \mathbf{W}_n\}$.

A.2 Recursive formulation via Poisson monitoring

We establish an HJB equation about the value function by replicating discrete-time incentive schemes with compound Poisson monitoring.

A.2.1 Compound Poisson incentive schemes

Definition 1 *An incentive scheme $(\Omega, \mathbb{F}, \mathbb{P}, \Gamma, h, a)$ is a compound Poisson incentive scheme if there exist optimal Δt-incentive scheme $\{\mathbf{L}_n, \mathbf{W}_n\}$ and an independent Poisson process $\{N_t : t \geq 0\}$ of frequency $\lambda := \frac{e^{-r\Delta t}}{1-e^{-r\Delta t}} r$ such that*

- *the cumulative excess likelihood ratio is a compound Poisson process*

$$\mathbf{\Gamma}_t = \sum_{n=1}^{N_t \wedge N} (\mathbf{L}_n - 1)\,;$$

- *the filtration $(\Omega, \mathbb{F}, \mathbb{P})$ is the augmented natural filtration of $\mathbf{\Gamma}$;*
- *the employment decision and effort recommendation are $h_t = a_t = \mathbf{1}_{N_{t-}<N}$.*

Note that $\mathbf{\Gamma}$ is a càdlàg martingale, and h and a are left-continuous and thus predictable. Although signal L_n arrives stochastically instead of deterministically, the choice of λ implies that the agent's continuation value W equals the corresponding one in Δt, i.e., $W_t = W_{N_t}$, and satisfies the *instantaneous* incentive compatibility constraint for $n \leq N-1$:

$$\lambda \mathbb{E}_{n-1}\left[(\mathbf{\Gamma}_n - \mathbf{\Gamma}_{n+1})(W_{n+1} - W_n)\right] = rk\,.$$

The instantaneous incentive compatibility constraint in turn implies incentive compatibility and the instanteous promise keeping condition, i.e.,

$$\lambda \mathbb{E}_{n-1}\left[W_{n+1} - W_n\right] = r(W_n - u + k)\,.$$

The principal's value increases because she acquires the costly signals right before adapting to it, instead of Δt beforehand.

A.2.2 Proof of Lemma 1

We show in fact a stronger version of Lemma 1. Recall that V is the value function of the continuous-time problem. Let V_{DL} denote the point-wise limit of discrete-time value functions (which exists by Lemma S5) and V_{CP} the value function of compound Poisson incentive schemes.

Lemma 3 *$V_{DL} = V = V_{CP}$.*

Proof. We prove the lemma by showing $V_{DL} \geq V$, $V \geq V_{CP}$, and $V_{CP} \geq V_{DL}$. First, $V_{DL} \geq V$ results from the restriction that continuous-time incentive schemes can be approximated by discrete-time incentive schemes in value. The supremum over discrete-time schemes is therefore at least as large as the supremum over continuous-time schemes. Second, $V \geq V_{CP}$ follows from inclusion because compound Poisson incentive schemes can be approximated from discrete time (Lemma S7). Third, $V_{CP} \geq V_{DL}$ holds because the corresponding compound Poisson incentive scheme attains a higher value than the optimal Δt-incentive scheme, by delaying costly monitoring. ■

A.2.3 Compound Poisson HJB

Definition 2 (Viscosity solution to compound Poisson HJB) *A concave continuous function V is a viscosity solution to the compound Poisson HJB*

$$v(W) = \pi + \sup_{\substack{\lambda, (\mathbf{L}, \mathbf{J}) \\ \mathrm{supp}(\mathbf{L}, \mathbf{J}) \leq 4}} \lambda \mathbb{E}\left[v(J) - v(W) - c(L)\right]$$

$$s.t. \begin{cases} \mathbb{E}\left[1 - L\right] & = 0 \\ \lambda \mathbb{E}\left[J - W\right] & = W - u + k \\ \lambda \mathbb{E}\left[(1 - L)(J - W)\right] & = k \end{cases}$$

if and only if

1. *it is a viscosity subsolution, i.e. for $\phi \in \mathcal{C}^2$ with $\phi \geq V$ and $\phi(W) = V(W)$,*

$$V(W) \leq \pi + \sup \lambda \mathbb{E}\left[v(J) - v(W) - c(L)\right]$$

2. *it is a viscosity supersolution, i.e. for $\phi \in \mathcal{C}^2$ with $\phi \leq V$ and $\phi(W) = V(W)$,*

$$V(W) \geq \pi + \sup \lambda \mathbb{E}\left[v(J) - v(W) - c(L)\right]$$

Our definition specializes the definition for general jump processes in Soner (1988) to compound Poisson processes.

Remark 4 To simply notation, frequency λ in Definition 2 and thereafter is normalized by discount rate r.

Proposition 3 *Value function V is a viscosity solution to the compound Poisson HJB.*

Proof. We first show that V is a subsolution. It suffices to consider concave ϕ because the Hamiltonian is the same for the concave envelope of[14] ϕ. Suppose that,

[14] If $\phi(W)$ is not on the concave envelope, then the RHS is infinity and so the inequality holds trivially.

for some $W = W_0 \in (0, u - k)$,

$$\phi(W) > \pi + \sup \lambda \mathbb{E}\left[\phi(J) - \phi(W) - c(L)\right] .$$

Then the strict inequality also holds in a δ-neighborhood of W_0 by the theorem of maximum.

Take $\delta, \theta > 0$ and exit time $\rho := \inf\{t : W_t \notin (W_0 - \delta, W_0 + \delta)\}$ from Lemma S8. For any compound Poisson incentive scheme, Itô's lemma implies

$$\begin{aligned}
\mathbb{E}\left[e^{-r\rho}\phi(W_\rho)\right] =& \phi(W_0) + \mathbb{E}\left[\int_0^\rho re^{-rt}\left(-\phi(W_t)\right) dt + \sum_{t \le \rho} \phi(W_t) - \phi(W_{t-})\right] \\
=& \phi(W_0) + \mathbb{E}\left[\int_0^\rho re^{-rt}\left(-\phi(W_t) + \lambda_t \mathbb{E}\left[\phi(J_t) - \phi(W_t)\right]\right) dt\right] .
\end{aligned}$$

The value of any compound Poisson incentive scheme with $\mathbb{E}\left[1 - e^{-r\rho}\right] \geq \theta$ is thus bounded by

$$\begin{aligned}
& \mathbb{E}\left[\int_0^\rho re^{-rt}\left(\pi - \lambda_t \mathbb{E}\left[c(L_t)\right]\right) dt + e^{-r\rho}\phi(W_\tau)\right] \\
=& \phi(W_0) + \mathbb{E}\left[\int_0^\rho re^{-rt}\left(-\phi(W_t) + \pi + \lambda_t \mathbb{E}\left[\phi(J_t) - \phi(W_t) - c(L_t)\right]\right) dt\right] \\
<& V(W_0)
\end{aligned}$$

where the strict inequality follows from the strictly negative integrand over a set of strictly positive measure. The inequality contradicts V as the value of compound Poisson incentive schemes with $\mathbb{E}\left[1 - e^{-r\tau}\right] \geq \theta$ (Lemma S8).

The supersolution inequality holds because compound Poisson controls are admissible. ■

A.2.4 Poisson HJB

Definition 3 (Viscosity solution to Poisson HJB) *A concave continuous function V is a viscosity solution to the (simple) Poisson HJB*

$$
\begin{aligned}
v(W) =& \pi + (W-u+k)v'(W) \\
&+ \sup_{\substack{L,J \\ \text{s.t. } (1-L)(J-W)>0}} \frac{1}{1-L}\frac{k}{J-W}\left(v(J)-v(W)-(J-W)v'(W)-c(L)\right)
\end{aligned}
$$

if and only if

1. *it is a viscosity subsolution, i.e. for $\phi \in \mathcal{C}^2$ with $\phi \geq V$ and $\phi(W) = V(W)$,*

$$
\begin{aligned}
V(W) \leq& \pi + (W-u+k)\phi'(W) \\
&+ \sup \frac{1}{1-L}\frac{k}{J-W}\left(\phi(J)-V(W)-(J-W)\phi'(W)-c(L)\right)
\end{aligned}
$$

2. *it is a viscosity supersolution, i.e. for $\phi \in \mathcal{C}^2$ with $\phi \leq V$ and $\phi(W) = V(W)$,*

$$
\begin{aligned}
V(W) \geq& \pi + (W-u+k)\phi'(W) \\
&+ \sup \frac{1}{1-L}\frac{k}{J-W}\left(\phi(J)-V(W)-(J-W)\phi'(W)-c(L)\right)
\end{aligned}
$$

When $(\mathbf{L}, \mathbf{J})$ has finite support, we enumerate the probability p_i of each support (L_i, J_i) for integer i.

Proof of Proposition 1. We first show that V is a viscosity subsolution to the Poisson HJB. Take $\phi \in \mathcal{C}^2$ with $\phi \geq V$ and $\phi(W) = V(W)$. For any control $\{\lambda, (\mathbf{L}, \mathbf{J})\}$ of the compound Poisson HJB, denote the maximizer of the incentive–cost ratio $\frac{(1-L)(J-W)}{-(\phi(J)-V(W)-(J-W)\phi'(W))+c(L)}$ on its support by (L^*, J^*). The Hamiltonian then satisfies

$$
\begin{aligned}
&\pi + \lambda\mathbb{E}\left[\phi(J)-V(W)-c(L)\right] \\
=&\pi + (W-u+k)\phi'(W) + \lambda\mathbb{E}\left[\phi(J_i)-V(W)-(J_i-W)\phi'(W)-c(L_i)\right] \\
\leq&\pi + (W-u+k)\phi'(W) + \frac{1}{1-L^*}\frac{k}{J^*-W}\left(\phi(J^*)-V(W)-(J^*-W)\phi'(W)-c(L^*)\right).
\end{aligned}
$$

The equality follows from the promise keeping constraint, and the inequality from the maximizer (L^*, J^*) and binding incentive compatibility constraint. Because (L^*, J^*) is feasible for the Poisson HJB, we conclude that V inherits the subsolution inequality for the Poisson HJB from the compound Poisson HJB by taking the supremum.

We continue to show that V is a viscosity supersolution to the Poisson HJB by contraposition. Suppose that there exist $\phi \leq V$ with $\phi(W) = V(W)$, control (L, J), and $\epsilon > 0$ such that

$$\begin{aligned} V(W) <& \pi + (W - u + k)\phi'(W) \\ &+ \frac{1}{1-L}\frac{k}{J-W}\left(\phi(J) - V(W) - (J-W)\phi'(W) - c(L)\right) - \epsilon\,. \end{aligned}$$

We construct a binary control $\{\lambda, \{p_i, L_i, J_i\}_{i=1,2}\}$ parametrized by $\lambda > 0$. It is defined by $L_1 =: L$ and $J_1 := J$ with p_1, p_2, L_2, and J_2 determined by the law of total probability and the three constraints of Definition 2.

As $\lambda \to \infty$, it can be shown that $c(L_2) = o(\lambda^{-1})$ and

$$\phi(J_2) - V(W) - (J_2 - W)\phi'(W) = (J_2 - W)\phi'(W) + o(\lambda^{-1}) - (J_2 - W)\phi'(W) = o(\lambda^{-1})\,.$$

Therefore, the Hamiltonian of the compound control satisfies

$$\begin{aligned} &\pi + \lambda \sum_i p_i \left(\phi(J_i) - V(W) - c(L_i)\right) \\ =&\pi + (W - u + k)\phi'(W) + \lambda \sum_i p_i \left(\phi(J_i) - V(W) - (J_i - W)\phi'(W) - c(L_i)\right) \\ =&\pi + (W - u + k)\phi'(W) + \frac{1}{1-L}\frac{k}{J-W}\left(\phi(J) - V(W) - (J-W)\phi'(W) - c(L)\right) + o(\lambda^{-1}) \\ >&V(W) + \epsilon + o(\lambda^{-1}) \end{aligned}$$

where the first equality results from the promise keeping constraint. Therefore, V fails the supersolution inequality for the compound Poisson HJB for sufficiently large λ, which is a contradiction. ■

A.3 Optimality of termination upon Poisson arrival

We shall show the optimality of immediate termination in discrete time and then upon Poisson arrival by continuity.

A.3.1 Proof of Lemma 2

Let $(\mathbf{L}_1, \mathbf{J}_1)$ denote the maximizer to the discrete-time Bellman equation (Lemma S2). It is finitely supported by Lemma S3. Suppose that none of those signals lead to termination, i.e. $0 \notin \operatorname{supp} \mathbf{J_1}$. At each $J_1 \in \operatorname{supp} \mathbf{J}_1$, let $(\mathbf{L}_2, \mathbf{J}_2)$ denote the maximizer. The value at W_0 can be written as

$$
\begin{aligned}
& V_{\Delta t}(W_0) \\
=& (1-e^{-r\Delta t})(u-k) - \mathbb{E}\left[c(L_1)\right] + e^{-r\Delta t}\mathbb{E}\left[V_{\Delta t}(J_1)\right] \\
=& (1-e^{-r\Delta t})(u-k) - \mathbb{E}\left[c(L_1)\right] \\
& + e^{-r\Delta t}\mathbb{E}\left[(1-e^{-r\Delta t})(u-k) - \mathbb{E}\left[c(L_2(L_1))\right] + e^{-r\Delta t}\mathbb{E}\left[V_{\Delta t}(J_2)\right]\right] \\
=& (1+e^{-r\Delta t})(1-e^{-r\Delta t})(u-k) + \mathbb{E}\left[-c(L_1) - e^{-r\Delta t}c(L_2(L_1)) + e^{-2r\Delta t}V_{\Delta t}(J_2)\right] .
\end{aligned}
$$

For $p_0 := \frac{1}{1+e^{-r\Delta t}}$, consider the mixed control $(\tilde{\mathbf{L}}, \tilde{\mathbf{J}}) := p_0 \boldsymbol{\delta}_{(1,W_0)} + (1-p_0)\left(\mathbf{L}_1 \times \mathbf{L}_2(\mathbf{J_1}), \mathbf{J}_2(\mathbf{J}_1)\right)$. It can be verified that, due to the choice of p_0, the control is admissible (Equation (S1)). Therefore, the value at W_0 is greater than the value of this control, i.e.,

$$
\begin{aligned}
& V_{\Delta t}(W_0) \\
\geq & (1-e^{-r\Delta t})(u-k) + (1-p_0)\mathbb{E}\left[-c(L_1L_2) + e^{-r\Delta t}V_{\Delta t}(J_2)\right] + p_0 e^{-r\Delta t}V_{\Delta t}(W_0) \\
\geq & (1-e^{-r\Delta t})(u-k) + (1-p_0)\mathbb{E}\left[-c(L_1) - c(L_2) + e^{-r\Delta t}V_{\Delta t}(J_2)\right] + p_0 e^{-r\Delta t}V_{\Delta t}(W_0) \\
> & (1-e^{-r\Delta t})(u-k) + (1-p_0)\mathbb{E}\left[-c(L_1) - c(L_2) + e^{-r\Delta t}V_{\Delta t}(J_2)\right] + p_0 e^{-r\Delta t}V_{\Delta t}(W_0) .
\end{aligned}
$$

The second inequality results from compound reduction (Assumption 2).

By applying the inequality repeatedly to $V_{\Delta t}(W_0)$ on the RHS, we obtain

$$V_{\Delta t}(W_0) > \left((1-e^{-r\Delta t})(u-k) + (1-p_0)\mathbb{E}\left[-c(L_1) - c(L_2) + e^{-r\Delta t}V_{\Delta t}(J_2)\right]\right)\sum_{n=0}^{N}\left(e^{-r\Delta t}p_0\right)^n$$
$$+ \left(p_0 e^{-r\Delta t}\right)^{N+1} V_{\Delta t}(W_0)$$

where the difference between the LHS and the RHS is increasing in N. Therefore, we take $N \to \infty$ to obtain

$$\begin{aligned}
&V_{\Delta t}(W_0)\\
>&\left((1-e^{-r\Delta t})(u-k) + (1-p_0)\mathbb{E}\left[-c(L_1) - c(L_2) + e^{-r\Delta t}V_{\Delta t}(J_2)\right]\right)\frac{1}{1-p_0e^{-r\Delta t}}\\
=&(1+e^{-r\Delta t})(1-e^{-r\Delta t})(u-k) - \mathbb{E}\left[-e^{-r\Delta t}c(L_1) - e^{-r\Delta t}c(L_2) + e^{-2r\Delta t}V_{\Delta t}(J_2)\right]\\
>&V_{\Delta t}(W_0)
\end{aligned}$$

where the equality results from $\frac{1-p_0}{1-p_0e^{-r\Delta t}} = e^{-r\Delta t}$, and the second inequality from $e^{-r\Delta t} < 1$. We have $V_{\Delta t}(W_0) > V_{\Delta t}(W_0)$, which is a contradiction.

A.3.2 The value function satisfies Termination HJB

Proposition 4 (Termination HJB) *Value function V is a classical solution to the Termination HJB*

$$\begin{aligned}
v(W) =&\pi + (W-u+k)v'(W)\\
&+ \max_L \frac{1}{1-L}\frac{k}{0-W}\left(v(0) - v(W) - (0-W)v'(W) - c(L)\right) \quad \forall W \in (0, u-k)\,.
\end{aligned}$$

Proof. Because V is concave, it is twice differentiable almost everywhere by the Alexandrov theorem. We first show the classical subsolution inequality for $W \in (0, u-k)$ such that $V''(W)$ exists. For $\epsilon > 0$, take $\phi \geq V$ with $\phi(W) = V(W)$ such that $\phi''(W) \in (V''(W), V''(W)+\epsilon)$. It satisfies $\phi'(W) = V'(W)$ because V is differentiable (Lemma S9). By Lemma 2.2 in Soner (1988), the viscosity subsolution inequality of Poisson HJB is equivalent to, for any $\delta > 0$,

$$\begin{aligned} V(W) \leq & \pi + (W - u + k)V'(W) \\ & + \max\Bigg\{ \sup_{\substack{L,J \\ |J-W|<\delta}} \frac{1}{1-L}\frac{k}{J-W}\left(\phi(J) - V(W) - (J-W)V'(W) - c(L)\right), \\ & \qquad \sup_{\substack{L,J \\ |J-W|\geq\delta}} \frac{1}{1-L}\frac{k}{J-W}\left(V(J) - V(W) - (J-W)V'(W) - c(L)\right) \Bigg\}. \end{aligned} \tag{4}$$

For the conditional maximizer L given by FOC (3), value function V satisfies

$$\lim_{J \to W} \frac{1}{1-L}\frac{k}{J-W}\left(V(J) - V(W) - (J-W)V'(W) - c(L)\right) = k\left(c''(1)\left(-V''(W)\right)\right)^{\frac{1}{2}}$$

because it is twice differentiable at W. With V replaced by ϕ, the analogous term converges to

$$\lim_{J \to W} \frac{1}{1-L}\frac{k}{J-W}\left(\phi(J) - V(W) - (J-W)V'(W) - c(L)\right) = k\left(c''(1)\left(-\phi''(W)\right)\right)^{\frac{1}{2}}.$$

Because $|\phi''(W) - V''(W)| < \epsilon$, Equation (4) now reads

$$\begin{aligned} V(W) \leq & \pi + (W - u + k)V'(W) \\ & + \sup_{L,J} \frac{1}{1-L}\frac{k}{J-W}\left(V(J) - V(W) - (J-W)V'(W) - c(L)\right) + O(\epsilon) + o_\delta(1) \\ = & \pi + (W - u + k)V'(W) \\ & + \max_{L} \frac{1}{1-L}\frac{k}{0-W}\left(V(0) - V(W) - (0-W)V'(W) - c(L)\right) + O(\epsilon) + o_\delta(1) \end{aligned}$$

where the equality follows from Lemma S11. Because ϵ and δ are arbitrary, we obtain the classical subsolution at W.

Symmetrically, we derive the classical supersolution inequality. Therefore, the value function V solves the termination HJB whenever it is twice differentiable. Because $V \in \mathcal{C}^1$ (Lemma S9), it also solves the termination HJB for all $W \in (0, u-k)$ by continuity. ■

B Proof of Proposition 2

Part 1 holds because value function $\tilde{V} = V$ and control variables $\left(\tilde{\lambda} = \frac{\tilde{r}}{r}\lambda, \tilde{L} = L, \tilde{J} = J = 0\right)$ solve the HJB and boundary conditions in Proposition 1. Similarly, Part 2 holds because of value function $\tilde{V}(W) = V(W) + \frac{\tilde{\pi}-\pi}{u-k}W$ and control variables $\left(\tilde{\lambda} = \lambda, \tilde{L} = L, \tilde{J} = J = 0\right)$.