Implementing Evidence Acquisition:

Time Dependence in Contracts for Advice*

Yingkai Li† Jonathan Libgober‡

First version: November 12, 2023

This version: September 26, 2024

Newest version

Abstract

An expert with no inherent interest in an unknown binary state can exert effort to acquire a piece of falsifiable evidence informative of it. A designer can incentivize learning using a mechanism that provides state-dependent rewards within fixed bounds. We show that eliciting a single report maximizes information acquisition if the evidence is revealing or its content predictable. This conclusion fails when the evidence is sufficiently imprecise, the failure to find it is informative, and its contents could support either state. Our findings shed light on incentive design for consultation and forecasting by showing how learning dynamics qualitatively shape effort-maximizing contracts.

Keywords— Scoring rules, dynamic contracts, dynamic moral hazard, Poisson learning.

---

*A preliminary version of this paper has been accepted in the Twenty-Fifth ACM Conference on Economics and Computation (EC'24) as a one-page abstract with the title *Optimal Scoring for Dynamic Information Acquisition*. Part of this work was completed while Jonathan Libgober was visiting Yale University and Yingkai Li was a postdoc at Yale University. The authors thank Dirk Bergemann, Tilman Börgers, Juan Carrillo, Tan Gan, Marina Halac, Jason Hartline, David Kempe, Nicolas Lambert, Elliot Lipnowski, Mallesh Pai, Harry Pei, Larry Samuelson, Philipp Strack, Guofu Tan, Kai Hao Yang and numerous seminar audiences for helpful comments and suggestions. Yingkai Li thanks Sloan Research Fellowship FG-2019-12378 and NUS Start-up Grant for financial support.

†Department of Economics, National University of Singapore. Email: `yk.li@nus.edu.sg`.

‡Department of Economics, University of Southern California. Email: `libgober@usc.edu`.

1 Introduction

Often when decision makers seek informed opinions, expertise takes time to acquire and skill to interpret. For instance, senior officers deciding whether to pre-emptively respond to a potential threat or regulators deciding whether to allow an aircraft type to fly after a mishap typically rely on other specialists to learn about and communicate the socially optimal decision.[1] Given the gap between authorities making decisions and experts acquiring information about them, the former must decide how to best incentivize the latter.

This paper shows how the set of implementable expertise acquisition strategies depends jointly on the learning technology and the ability to elicit information *dynamically*. Our research question contrasts with similar ones from past work on information elicitation in that we ask about both dynamic learning and dynamic contracts. Our contribution is to articulate interactions between these two dimensions. We do so restricting attention to learning technologies which we refer to as *finding evidence*. As the name suggests, this technology reflects experts who learn until some signal arrives (i.e., the evidence is found). This class facilitates tractability, describes economically meaningful applications, and allows us to capture qualitative properties of dynamic learning technologies.

On the one hand, toward minimality, we take the agent to have no intrinsic interest in the state being learned, so that no learning would take place absent provided incentives. Outsiders, on the other hand, lack the expertise necessary to recognize the exertion of effort or interpret its outcomes. Thus, our contracting problem features both moral hazard (as the expert's actions are unobserved) and soft information (as contracts must induce the expert to truthfully reveal what was learned). The only means to provide incentives is through rewards conditioning on the true state[2] once revealed.

Our main results are as follows. We show that a static contract—or, more precisely, eliciting a single report from the agent—suffices to induce maximal information acquisition when (1) no information is conveyed by the failure to find evidence, (2) evidence is always conclusive, or (3) the contents of the evidence are known. But we also identify a set of parameters jointly violating these conditions for which contracts involving dynamic reports *do* expand the set of strategies an agent can be induced to follow.

Finding evidence learning technologies generalize the discrete-time version of the foundational Poisson bandit model. We consider an agent who chooses (privately) over time whether to exert costly effort to learn about a binary state, the initial prior over which

[1]In these specific applications, determine the adversary's capabilities or investigate the incident.

[2]We note our results are maintained if rewards can depend on a signal correlated with the state, rather than the state itself. To minimize notation we suppress this possibility, but our results are equally relevant to applications where such signals are available.

is commonly known. Each time effort is exerted, the agent may (privately) observe an informative signal with fixed probability—reflecting the evidence having been found. Our assumption that both effort and learning are private reflects our interest in cases where evidence itself requires expertise to recognize and interpret. Many Poisson learning models from past work assume signal arrival can reveal one state (e.g., Keller et al., 2005). In our model, while a single arrival may occur—when the sought evidence is found—we do *not* necessarily assume either that (a) this arrival reveals the state or (b) it can support only one state. For instance, while aviation incidents often have a single cause, this cause need not definitively imply a defect with the aircraft type. Furthermore, the conclusion the evidence would support if found might not be known in advance, either: an incident's explanation may suggest either an underlying defect (e.g., a faulty software system) or a purely idiosyncratic issue (e.g., pilot error). Similar descriptions broadly characterize problems where experts seek some particular piece of information, aligning with our framework.

An appealing property of finding evidence settings is that the set of implementable strategies is determined by the maximum length of time that the agent can be induced to work absent evidence arrival. Indeed, the agent produces the most informative experiment precisely when working for this long. Thus, an equivalent formulation of our main question is whether varying rewards over time facilitates information production (as opposed to effort). These properties enhance tractability and intuition, making finding evidence settings a natural starting point for theoretically relating dynamic learning to incentive design.

Note that the implementation problem is trivialized if transfers or punishments can be set arbitrarily. To avoid this degeneracy and reflect practical limitations on the strength of incentives, we posit that rewards and punishments belong to some bounded range. The implementation problem may be of interest for various reasons. For one, determining the set of implementable effort strategies addresses whether a designer facing some social welfare function can induce *efficiency*, assuming transfers are welfare neutral (see Bergemann and Välimäki, 2002). The implementation problem is also of interest if the information acquisition problem is of first-order importance relative to the residual value of payments or if rewards themselves are non-monetary (e.g., a promotion or special title, recognition as in Halac and Prat (2016), a favorable recommendation, etc.).

The designer *could* ask the agent to report over time while acquiring information, allowing the aforementioned rewards to vary dynamically. We say a mechanism *has a static implementation* if it suffices to elicit just a single report of the aggregated information acquired by the agent, with the rewards provided being constant over time. Such an implementation, where rewards are provided as a function of the agent's reported (final) belief and the true state, is known in the literature on information elicitation as a *scoring rule*.[3]

In one-shot interactions, the scoring rule that provides the greatest gain from exerting effort will typically depend on the prior. To see why, consider a mechanism where the agent guesses a state and is rewarded if and only if the guess is correct. If the reward when correctly guessing the initially-more-likely state is very large, then the agent might feign the discovery of evidence that confirms the prior without exerting effort. But a more modest reward in this state could get the agent to actually learn whether it is worthwhile to go against the prior. So, if the agent becomes more confident in one state, then the scoring rule that maximizes the gain from exerting effort will lower the reward provided in that state.

Applying this intuition to the dynamic setting might seem to suggest that rewards should decrease in states that the agent views as more likely the longer he exerts effort, reflecting the aforementioned comparative static for one-shot information acquisition. However, we find that this is not the case in three economically meaningful instances:

1. Stationary environment (Theorem 1) where beliefs absent signals are constant;
2. Perfect-learning environment (Theorem 2) where signals fully reveal the state;
3. Single-signal environment (Theorem 3) where evidence moves beliefs in one direction.

In the first case, since the agent's belief does not change while exerting effort, the optimization problem is essentially the same at every point in time. Thus, there is no need to adjust the rewards. In the latter two cases, despite the intuition that there may be gains from "reoptimizing" rewards over time, doing so would also lower the agent's continuation payoff from exerting effort at earlier times, weakening incentives. Essentially, reoptimizing rewards at any history redistributes the agent's incentives to exert effort across multiple periods. No reoptimization is necessary in perfect-learning and single-signal environments.

Outside of these cases, the initial intuition that more effort can be incentivized by reoptimizing the scoring rule is valid. In general, contracts with decreasing reward structures implement maximum effort (Theorem 5)—intuitively, since increasing rewards simply encourages the agent to "shirk and lie." Perhaps more interesting in light of our other results, however, are sufficient conditions such that this type of dynamic contract strictly outperforms all scoring rules (Theorem 4). These conditions essentially boil down to a requirement that beliefs drift with a sufficiently strong violation of the second and third conditions.

We emphasize that our results do *not* imply dynamics are irrelevant when scoring rules implement maximum effort. First, it need not be the case that the effort-maximizing scoring

[3]To our knowledge, Brier (1950) first studied scoring rules as a way of evaluating (i.e., providing scores for) weather forecasts. In the subsequent literature, this terminology has been used to describe general reward schemes using payoffs that depend on belief reports and realized outcomes.

rule provides the strongest incentives at the prior, since incentives must be balanced as the agent's belief changes. As we discuss, this contrasts with the static case. Second, in *static* settings, any level of effort that can be induced given an arbitrary mechanism can be induced by a mechanism that only requires the agent to guess a state and provides a positive reward if the guess is correct (and no reward otherwise, Li et al. (2022)). Such scoring rule formats may fail to implement an implementable amount of effort in our setting, even when the effort-maximizing contract has an implementation as a scoring rule. Specifically, as a means of encouraging the agent to exert effort at earlier times, he may be provided the option to secure a strictly positive minimum wage even if his prediction of the state is wrong.

Given the recognized importance of dynamics in information acquisition, we view our central question—i.e., how dynamic reward provision influences implementable information acquisition strategies—as a natural starting point toward developing techniques applicable to the analysis of such settings, highlighting its theoretical significance. Yet our work has notable practical implications as well—for instance, shedding light on the use of performance pay in advising contracts (e.g., with consultants, forecasters, or analysts). Often such incentives take the form of bonuses paid if the advice provided turns out to be correct (Stahl, 2018). While idiosyncratic factors outside our simple model may determine the use and design of such contracts, understanding how to structure them to maximize the impact of incentives is still critical. Beyond our theoretical contribution, our findings provide guidelines regarding when dynamic contracts can better incentivize information acquisition. We hope these insights are practically relevant for settings where expert advice is crucial.

Our paper joins a long line of work in economic theory asking how to incentivize information acquisition or experimentation. A key theoretical novelty that arises in such settings is the introduction of *endogenous adverse selection* since different effort choices (typically themselves subject to moral hazard) will provide the agent with different beliefs over the relevant state. This basic interaction, where an agent exerts effort under moral hazard to acquire information, has been analyzed under varying assumptions regarding the underlying information acquisition problem and contracting abilities.[4]

In our model, mechanisms can condition rewards on the state as well as a report from the agent. Much of the literature on scoring rule design focuses exclusively on the *elicitation* of information (see, for instance McCarthy (1956); Savage (1971); Lambert (2022), as well as Chambers and Lambert (2021) for the dynamic setting). Our focus is instead on the question of how to incentivize its acquisition. To the best of our knowledge, Osband (1989) is the

[4]For instance, static information acquisition technologies where information is acquired before contracting (Crémer and Khalil, 1992) or after (Krähmer and Strausz, 2011), where the outcome of experimentation may be contractable (Yoder (2022), as well as Chade and Kovrijnykh (2016) in a repeated setting), or where decision rights but not transfers are available (Szalay, 2005).

earliest work to focus on this question. More recent work in economics and computer science include Häfner and Taylor (2022); Zermeno (2011); Carroll (2019); Li et al. (2022); Neyman et al. (2021); Hartline et al. (2023); Whitmeyer and Zhang (2023); Chen and Yu (2021). Our main point of departure from this line of work stems from our focus on dynamics, and in particular the ability to write dynamic contracts.[5] A related application of scoring rules is to *screening forecasters*, where mechanisms address initial adverse selection rather than moral hazard (see Deb et al. (2023); Dasgupta (2023)); in particular, Deb et al. (2018) study a dynamic problem with this application, making analogous contracting assumptions to us and also describing when the solution involves a single report from the agent.[6]

The Poisson information acquisition technology has been a workhorse for the analysis of how to structure *dynamic* contracts for experimentation.[7] Bergemann and Hege (1998, 2005) were early contributions studying a contracting problem under the assumption that a "success" reveals the state. The subsequent literature has considered variations on this basic environment (e.g., Hörner and Samuelson (2013) relaxes commitment; Halac et al. (2016) allow for ex-ante adverse selection and transfers; Guo (2016) considers delegation without transfers). The closest to our work is Gerardi and Maestri (2012), who assume a Poisson arrival technology and, as in the scoring rule literature, allow for state-dependent contracts. Our information acquisition technology generalizes this technology to allow for evidence that may support either state and be inconclusive. Our results show that *both* modifications are necessary for dynamic mechanisms to outperform static scoring rules.

On this note, Poisson bandits have been extensively utilized in strategic settings, even absent the ability to write contracts, since the influential work of Keller et al. (2005). An advantage of this setting is that it facilitates qualitative, economically-substantive properties of information acquisition and predicted behavior; a highly incomplete list of examples includes Strulovici (2010); Che and Mierendorff (2019); Damiano et al. (2020); Keller and Rady (2015); Bardhi et al. (2024); Lizzeri et al. (2024). Our exercise essentially amounts to *designing payoffs* in a single-agent environment. Note the agent's problem need not admit a simple stationary representation for arbitrary mechanisms in our framework.[8] This

[5]While Neyman et al. (2021); Hartline et al. (2023) and Chen and Yu (2021) allow dynamic information acquisition, all explicitly assume contracts must be static.

[6]A related line of work studies mechanism design with ex-post verifiability in multi-agent contexts (e.g., Deb and Mishra, 2014; DeMarzo et al., 2005). In this literature, Crémer (1987) noted the potential for trivialization without restrictions in the set of mechanisms (see discussions in Deb and Mishra (2014) and Skrzypacz (2013)). Mylovanov and Zapechelnyuk (2017) studies a setting with the ability to make contingent punishments, imposing restrictions on these punishments as we do.

[7]McClellan (2022); Henry and Ottaviani (2019) consider related models where information acquisition instead uses a *Brownian motion technology*, and an agent deciding when to stop experimenting.

[8]Ball and Knoepfle (2024) study monitoring using a Poisson framework; while they allow bidirectional signals, their design problem maintains recursivity, unlike ours.

contrasts with most of the settings where payoffs are exogenous, in which case such stationarity may be crucial for tractability. Partially for this reason, our approach does not require determining the agent's exact best response following an arbitrary dynamic contract.

2 Model

Our model considers an agent (who, depending on the application, may be an individual expert or a team working as a single entity) who can acquire information about an uncertain state $\theta \in \Theta = \{0, 1\}$ (e.g., whether the adversary is capable of an attack or whether an aircraft type has a design flaw) at discrete times $\{0, \Delta, 2\Delta, \ldots, T\}$. For conceptual simplicity, we take $T < \infty$, although our results apply equally to the high-frequency limit as $\Delta \to 0$. A mechanism designer shares a common prior with the agent over θ; we let D denote the initial probability that $\theta = 1$. We have in mind situations where the designer must make some decision at time $T + \Delta$ (e.g., whether to attack the adversary or ground the aircraft type), although as this plays no role we remain agnostic about the designer's precise preferences—aside from preferring (Blackwell) more information. We first describe the information acquisition technology and then describe the contracting environment.

2.1 Information Acquisition

The agent can acquire information at time t by paying a cost $c\Delta$, where $c > 0$ is an effort cost parameter. When this cost is paid, with some probability a piece of (falsifiable) evidence informative of θ arrives. We take the probability of evidence arrival to be $\lambda_\theta \Delta$; note that this probability may depend on the state. If no evidence arrives, the agent observes a *null signal*, which we denote by N. If the agent does not exert effort, then a null signal is observed with probability 1. Without loss of generality, we assume that $\lambda_1 \geq \lambda_0$, so that the agent's belief drifts toward state 0 in the absence of evidence arrival. Once the evidence arrives, no further information can be acquired (e.g., only one mechanical error responsible for an aircraft malfunction can be found; if an adversary's capabilities are determined, no further information is relevant for assessing attack probability, etc.).

When the evidence arrives, the agent observes some $s \in S$—so that, throughout the paper, S is the set of *non-null* signals. Our main results on static contracts implementing maximal effort relate to the following special cases of this model:

1. **Stationary environments,** where $\lambda_1 = \lambda_0$ (so null signals do not move beliefs).

2. **Perfect-learning environments,** where the state is observed following every $s \in S$.

3. **Single-signal environments,** where $|S| = 1$.

To simplify exposition without detracting from our main message, it suffices to consider the case where $|S| \leq 2$. We refer to non-null signals as either "good news" (e.g., a pilot miscalculation is determined) or "bad news" (e.g., a mechanical issue is found) where a "good news" signal G arrives with probability $\lambda_\theta^G \Delta$ and a "bad news" signal B arrives with probability $\lambda_\theta^B \Delta$ when the state is $\theta \in \{0, 1\}$ (so that $\lambda_\theta = \lambda_\theta^G + \lambda_\theta^B$). We take $\lambda_1^G > \lambda_0^G$ and $\lambda_1^B < \lambda_0^B$. We use the terminology of "bad news" and "good news" to distinguish signals from states, although we do not necessarily view one state as intrinsically better than another (although this may be the case in some settings). Rather, the meaningful difference is that if beliefs drift absent signal arrivals, this drift is toward state 0. As only one piece of evidence can be found, no effort is exerted following the arrival of G or B.

This information acquisition technology generalizes Poisson bandit learning (applied to contracting problems by Gerardi and Maestri (2012); Halac et al. (2016)), as signals (a) need not reveal the state and (b) can be either good or bad. For any $t \leq T$ we let μ_t^N denote the posterior belief that $\theta = 1$ if no Poisson signal arrived before time t (including t) and the agent has exerted effort for all periods until t. We refer to μ_t^N as the "no-information belief" of the agent that at time t. Using the convention that $\mu_0^N = D$, Bayes rule implies:

$$\mu_t^N = \frac{\mu_{t-\Delta}^N(1 - \lambda_1 \Delta)}{\mu_{t-\Delta}^N(1 - \lambda_1 \Delta) + (1 - \mu_{t-\Delta}^N)(1 - \lambda_0 \Delta)}.$$

Similarly, we let μ_t^G and μ_t^B denote the agent's posterior when receiving Poisson signals G and B (respectively) exactly at time t (working until then). We similarly obtain:

$$\mu_t^s = \frac{\mu_{t-\Delta}^N \lambda_1^s}{\mu_{t-\Delta}^N \lambda_1^s + (1 - \mu_{t-\Delta}^N)\lambda_0^s},$$

noting that signal arrival is off-path whenever $\lambda_1^s = \lambda_0^s = 0$ (so that beliefs following this event will play no role). Note that if $\lambda_1^G, \lambda_0^G \in (0, 1)$, good news does not reveal the state (similarly for bad news if $\lambda_1^B, \lambda_0^B \in (0, 1)$).

2.2 Contracting and Main Question

Our goal is to characterize when static contracts can implement the maximum amount of second-best effort given access to arbitrary contracts with bounded rewards. The agent's effort choices and signal realizations are unobserved over the course of the interaction.

At any time t, let M_t be the message space of the agent. The history at time t is denoted as $h_t = \{m_{t'}\}_{t' \leq t}$. Let $\mathcal{H}_t$ be the set of all possible histories at time t. As our interest is in cases where rewards can condition on the true state ex-post, to avoid trivialization we take rewards to belong to a bounded set, which we take to be the unit interval for simplicity. Therefore, the agent receives a benefit according to the function:

$$R : \mathcal{H}_T \times \Theta \to [0, 1]$$

where $R(h_T, \theta)$ is the fraction of the total available reward provided to the agent when his history of reports is h_T, and the realized state is θ. An alternative interpretation is that the designer has an indivisible reward (e.g., a special recognition), with $R(h_T, \theta)$ denoting the probability that the agent receives the reward.[9] Thus, if the agent has exerted effort in $\tilde{t}$ periods, his final payoff is $R(h_T, \theta) - c\tilde{t}\Delta$.

Our main findings are qualitatively unchanged if, instead of θ itself, the designer can condition rewards on some binary signal $\tilde{s} \sim \mathcal{I}(\theta)$ which the agent has no (other) information about. For ease of exposition we do not add this possibility to our model, but in several of our applications such signals may be available even if the state itself is not—for instance, a concrete case is when $\tilde{s}$ is the outcome of some external investigation (e.g., whether the aircraft manufacturer identifies a defect or not).

The main result of our paper is to provide sufficient conditions under which it suffices to only require a single, final report from the agent. Using the terminology from the information elicitation literature, a *scoring rule* $P : \Delta(\Theta) \times \Theta \to \mathbb{R}$ for eliciting the agent's subjective belief is a mapping from the posterior space and the state space to a real number; we refer to the corresponding real number output by this function as the *score*. A scoring rule is essentially a static contract described in our model.

Definition 1 (Implementation as Scoring Rules).
A dynamic contract R can be implemented as a scoring rule P if the message space $M_T = \Delta(\Theta)$, and for any history of reports h_T with last message m_T, we have $R(h_T, \theta) = P(m_T, \theta)$ for all states $\theta \in \Theta$.

Scoring rules are substantially simpler compared to arbitrary dynamic contracts, since they only require a single report rather than richer sequences of reports and time-varying rewards. The main goal of our paper is to guide when dynamic reporting is truly necessary to incentivize maximal information acquisition. Simply put, this corresponds to the following question:

[9] While we allow randomization over the event that the agent receives the full reward, the contracts we defined above are essentially deterministic since the probability that the agent gets the full reward is a deterministic function of the history of reports. In this paper, we focus on deterministic contracts, deferring our discussion of how the possibility of randomization influences the results in Section 6.1.

Main Question: Can the maximum effort implementable using an arbitrary dynamic contract R also be implemented using some scoring rule P?

We answer this main question using characterizations of effort-maximizing contracts, which also reveal other properties beyond the use of dynamics. These other characteristics may themselves shed light on which incentive schemes best incentivize consultants or forecasters.

2.3 Preliminary Simplifications

Having completed the formal presentation of our model, we now turn to some immediate simplifications that facilitate our analysis and turn out to be without loss of generality.

2.3.1 Stopping Strategies

We first simplify the set of effort profiles we must consider toward answering our main question. We argue it is without loss to assume that the agent follows a simple *stopping strategy.* In general, the agent's information acquisition strategy can be arbitrary and quite complex. For example, the agent could wait for several periods before starting or randomize these decisions. Nevertheless, for the purposes of characterizing the *maximum amount of effort* implementable by some contract, it is enough to assume that all effort is front-loaded.

More precisely, call a *stopping strategy* an information acquisition strategy whereby the agent chooses to (i) exert effort at every time $t \leq \tau$, conditional on the evidence not having yet arrived, and (ii) stop exerting effort once either the signal arrives or τ has passed. Abusing notation slightly, let τ denote a stopping time under a stopping strategy.

Lemma 1 (Stopping Strategies are Without Loss)**.**
Given any contract R and any best response of the agent with maximum effort length z_R conditional on not receiving any Poisson signal,[10] *there exists a stopping strategy τ_R with $\tau_R \geq z_R$ that is also optimal for the agent.*

The intuition for Lemma 1 is simple. An agent working earlier can pretend to have only worked later, but the converse is not necessarily true as previous reports cannot be undone. The proof follows from considering a modification where the agent exerts effort for the same amount of time but front-loads effort. While the gains from doing this depend on the contract, such a modification cannot hurt the agent who incurs the same cost but always adopts a strategy that ensures no less of a reward, and possibly even a larger one.

Note that the maximum effort duration conditional on no Poisson signal arrival uniquely determines the information acquired by the agent. Moreover, whenever the agent works

[10] If the agent randomizes, we let z_R denote the maximum effort length among all realizations.

longer, the aggregated information acquired is Blackwell more informative. Thus, Lemma 1 implies that to characterize the maximum *amount of information* the agent can be induced to acquire (and convey), it is enough to consider stopping strategies.

2.3.2 Menu Representation

Our next simplification essentially amounts to a version of the taxation principle for our environment. As a stopping strategy involves the agent exerting effort every period until evidence arrives, it is correspondingly not necessary for the agent to report every period. It turns out to be without loss to focus on implementation via a *sequence of menu options*.

Specifically, we associate each complete sequence of signal realizations—which we think of as the agent's (endogenously determined) type—with a *reward function* that maps states into rewards provided to the agent. Note that any type of this form is determined by the length of time the agent exerted effort and the signal observed immediately prior to stopping. We represent a general reward function as $r = (r_0, r_1)$, where r_0 is the reward in state 0 and r_1 is the reward in state 1. Our menu representation asserts that it suffices to consider contracts that associate each terminal time-signal pair with a reward function. Thus, each *menu option* is one such reward function, while the *menu* offered at time t, $\mathcal{R}_t$, is the set of reward functions available at time t *and after*. The agent then makes a one-time, irrevocable choice from the set of available menus.

Notice that the menus the agent faces, $\mathcal{R}_t$, shrink over time, as agents who observe signal realizations early can deviate to later menu options but not vice versa. In this formulation, the menu option selected by the agent corresponds to a truthful report of $s = G$ or $s = B$ immediately once observed. Thus, letting $u(\mu, r) \triangleq \mathbf{E}_{\theta\sim\mu}[r(\theta)]$ denote the agent's expected payoff with belief μ under reward function r, incentive compatibility requires that, whenever the agent's selection is made:

$$u(\mu_t^s, r_t^s) \geq u(\mu_t^s, r), \quad \forall r \in \mathcal{R}_t, \tag{IC}$$

recalling that μ_t^N denotes the agent's belief if exclusively null signals have been observed between 0 and t (and working until then), while μ_t^G and μ_t^B are the agent's beliefs if the corresponding non-null signal is observed at time t. For any time t and any signal $s \in S$, we denote $u_t^s(R) \triangleq u(\mu_t^s, r_t^s)$ as the agent's utility if ceasing effort entirely after t with belief μ_t^s—in other words, the left-hand side of (IC). We omit R when clear from context.

The following result formally presents our representation:

Lemma 2 (Menu Representation)**.**
Any dynamic contract R implementing optimal stopping time τ_R is equivalent to a sequence

of menu options $\{r_t^s\}_{t \le \tau_R, s \in S} \cup \{r_{\tau_R}^N\}$ *where* $r_t^s : \Theta \to [0,1]$*. At any* $t \le \tau$*, the agent can select any element of* $\mathcal{R}_t \triangleq \{r_{t'}^s\}_{t \le t' \le \tau_R, s \in S} \cup \{r_{\tau_R}^N\}$*, and is rewarded according to this (single) selection after* T*. For any* $t \le \tau_R$ *and* $s \in S$*, or for* $t = \tau_R$ *and* $s = N$*, (IC) holds.*

Appendix A.1 contains the proof of Lemma 2. Note that as per Eq. (IC), given access to the set of contracts in $\mathcal{R}_t$, it suffices to consider the incentive constraints where the agent accepts their intended reward function immediately and does not delay. Lemma 2 dramatically simplifies the space of contracts; for instance, it converts the design of the effort-maximizing contracts into a sequence of linear programs for which the sequence of menu options in the effort-maximizing contract can be computed efficiently (see Appendix OA 2 for details).

The above descriptions of incentive compatibility in our menu representation ignore the possibility that the agent stops working prior to either a Poisson signal arrival or time τ_R. Of course, this need not be the case. For any $t < \tau_R$, we denote

$$r_t^N = \arg\max_{r \in \mathcal{R}_t} u(\mu_t^N, r)$$

as the menu option the agent would choose at time t with belief μ_t^N, and we let $u_t^N(R) \triangleq u(\mu_t^N, r_t^N)$. The moral hazard constraint requires that $u_t^N(R)$ is less than the expected utility from adopting a stopping strategy with stopping time τ_R.

3 Preliminary Intuition and Illustrations

3.1 Preview for Single-Signal Perfect Learning

We preview our findings by presenting the solution in the $\Delta \to 0$ limit, where a single signal reveals the state, i.e.,

$$\lambda_0^B = \lambda_1^B = \lambda_0^G = 0, \quad \lambda_1^G > 0.$$

We take the horizon T to be sufficiently large so that it will not be a binding constraint. Illustrating this solution will highlight the tensions in maximizing the incentives to exert effort at different points in time.

For this learning environment, we show (in Theorem 2 below) that a scoring rule with two reward functions implements maximum effort: $(1, 0)$, corresponding to a guess of state 0, and $(0, r_1)$, corresponding to a guess of state 1. The value of r_1, the reward when guessing state 1 (correctly), depends on the initial prior, D (which coincides with μ_0^N).

We describe how r_1 is determined. Providing a higher reward for guessing state 1 encourages the agent to continue exerting effort, even as this state appears increasingly

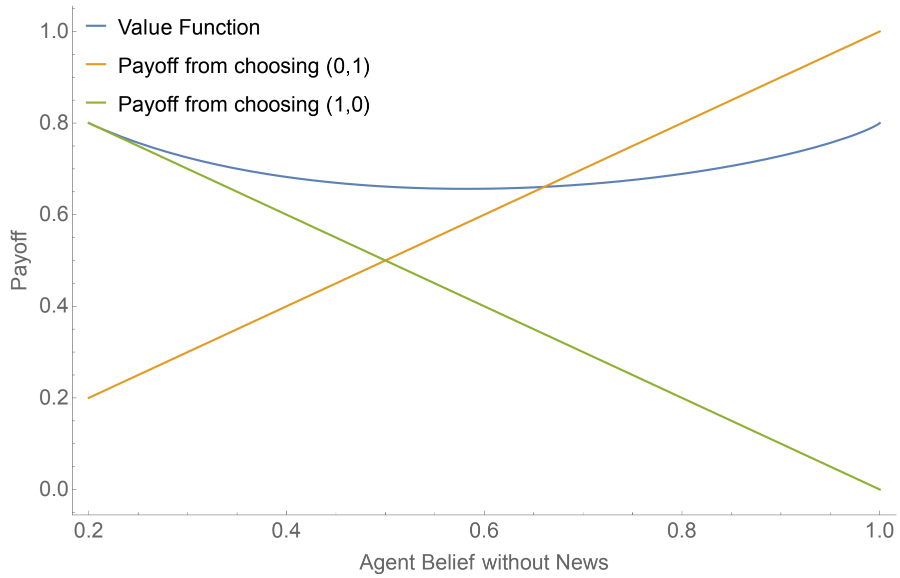

Figure 1: Value function with perfect learning; $r_1 = 1, c = .2, \lambda_1^G = 1$.

unlikely. In fact, one can show the agent is indifferent between continuing effort and selecting a reward when $\mu_\tau^N = \frac{c}{\lambda_1^G r_1}$ (see Appendix for details). If the event that $\theta = 1$ is not too likely according to the initial prior, then setting $r_1 = 1$ gets the agent to work for as long as possible. However, if the probability that $\theta = 1$ is initially high, setting $r_1 = 1$ may violate the agent's initial incentive constraints.

Figure 1 illustrates the agent's value function when $r_1 = 1$, assuming the agent works until $\mu_\tau^N = \frac{c}{\lambda_1^G}$, along with the expected payoff when selecting each reward function. While the adverse selection constraint holds for this contract, moral hazard is violated if the event that $\theta = 1$ is sufficiently likely initially. This can be seen by observing that the value function is below the expected payoff when choosing $(0, 1)$, so the agent would prefer to guess state 1 rather than exert any effort at all.

In this case, lowering r_1 is necessary to motivate the agent to begin working. The cost, of course, is that now the agent stops working once $\mu_\tau^N = \frac{c}{\lambda_1^G r_1}$, so that the agent stops sooner when r_1 is lower. Now, when $\mu_0^N < c/\lambda_1^G$, it is impossible to induce the agent to acquire any information at all. Outside of this range, a pair of thresholds, μ^*, μ^{**} satisfying $c/\lambda_1^G < \mu^* < \mu^{**} < 1$ determine the form of the effort-maximizing scoring rule:

- For $c/\lambda_1^G \leq \mu_0^N \leq \mu^*$, the effort-maximizing scoring rule sets $r_1 = 1$.
- For $\mu^* < \mu_0^N \leq \mu^{**}$, the effort-maximizing scoring rule sets $r_1 < 1$, with the exact value pinned down by the condition that at time 0, the agent is indifferent between (i) working absent signal arrival until their belief is $c/(\lambda_1^G r_1)$ and (ii) never working.

- For $\mu_0^N > \mu^{**}$, it is impossible to induce the agent to acquire any information.

Of course, when $\mu_0^N \in (\mu^*, \mu^{**})$, if the agent works beliefs may eventually leave this region. If the agent had *started* at such a belief, more effort could be induced by setting $r_1 = 1$. On the other hand, it is clear why this reoptimization cannot help. When $\mu_0^N \in (\mu^*, \mu^{**})$, $r_1 < 1$ will be set so that the initial moral hazard constraint binds—but if the agent expected r_1 to increase later, he would simply shirk and claim to have found evidence once the reward is increased to 1. Given that such adjustments are impossible, effort-maximizing mechanisms cannot utilize dynamics and are therefore static.

The lesson is that the tensions between optimizing incentives to exert effort both earlier and later may be unavoidable. A designer may be unable to re-optimize rewards because the re-optimized rewards would violate an incentive constraint at some other time. One deceptive aspect of this example is that the agent's moral hazard constraint only ever binds at the stopping belief and (possibly) at their initial belief. A technical challenge we face in Section 4.3, for instance, is that if signals are not fully revealing, it may be that the moral hazard constraint binds somewhere "in between." This feature will imply extra reward functions should be provided to the agent in the effort-maximizing scoring rule. Still, this example illustrates the intuition on how effort-maximizing rewards vary with the agent's beliefs, which will be useful for understanding the form of effort-maximizing contracts.

3.2 A Set of Static Problems from the Dynamic Problem

We now introduce key tools that will be useful in proving our results. Specifically, we decompose the dynamic problem into a sequence of static problems. Each static game in our decomposition is indexed by a pair of times, $t, t' \in [0, T]$, which we refer to as the *continuation game between* t *and* t', denoted $\mathcal{G}_{t,t'}$. Each continuation game considers the agent's problem at time t, but drops all incentive constraints except for (a) the one at time t, which says the agent is willing to start working, and (b) the one at t', which says they are willing to report the truth after having followed the stopping strategy which stops at t'.

More formally, the continuation game $\mathcal{G}_{t,t'}$ is a static game with prior belief $\mu_{t-\Delta}^N$, where the agent faces a single, binary effort choice: the agent can either (a) not exert effort or (b) exert effort up to and including time t' or until a Poisson signal arrives, incurring the associated costs of effort. For any contract R with stopping time $\tau_R \le T$, we say $\mathcal{G}_{t,\tau_R}$ is the continuation game at time t for contract R. In the rest of the paper, we omit the subscript of τ_R from the notation when the contract R is clear from context.

In the original dynamic environment, the agent has incentives to exert effort at time t if there exists $t' \ge t$ such that the agent has incentives to exert effort in the continuation

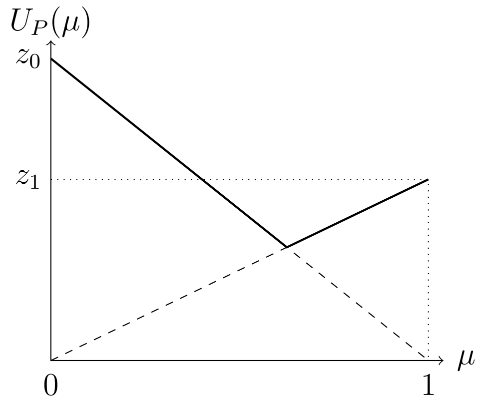

Figure 2: Expected score $U_P(\mu)$ of a V-shaped scoring rule P with parameters z_0, z_1.

game $\mathcal{G}_{t,t'}$. If the agent has incentives to work in every continuation game $\mathcal{G}_t$, then by backward induction the contract can implement τ_R. The natural question is whether effort is implementable in a given continuation game, a question whose answer follows immediately from existing results for the static case. To describe this we use the following concept:

Definition 2 (V-shaped Scoring Rules).
A scoring rule P is V-shaped with parameters z_0, z_1 if

$$P(\mu, \theta) = \begin{cases} z_1 & \mu \geq \frac{z_0}{z_1+z_0} \text{ and } \theta = 1 \\ z_0 & \mu < \frac{z_0}{z_1+z_0} \text{ and } \theta = 0 \\ 0 & \text{otherwise.} \end{cases}$$

We say P is a V-shaped scoring rule with kink at D if the parameters z_0, z_1 satisfies $z_0 = 1, z_1 = \frac{1-D}{D}$ if $D \geq \frac{1}{2}$ and $z_0 = \frac{D}{1-D}, z_1 = 1$ if $D < \frac{1}{2}$.

The terminology of the scoring rule as "V-shaped" comes from the property that the expected score $U_P(\mu) \triangleq \mathbf{E}_{\theta\sim\mu}[P(\mu, \theta)]$ is a V-shaped function, which is illustrated in Figure 2. Furthermore, given any V-shaped scoring rule P with kink at D, the agent with prior belief D is indifferent between guessing the state is either 0 or 1.

Proposition 1. *Consider any continuation game $\mathcal{G}_t$ where the agent prefers to exert effort rather than not. Then the agent prefers to exert effort in the continuation game $\mathcal{G}_t$ under the V-shaped scoring rule with kink at μ_t^N.*

We omit the proof of the Proposition, as it follows immediately from Li et al. (2022), which proved a more general version of this result for static environments. Briefly, V-shaped scoring rules maximize the expected score at all posteriors subject to (a) the constraint that the indirect utility is convex (as a consequence of incentive compatibility) and (b) the expected score at the prior is a constant (so that incentives for the agent to exert effort are maximized). For any fixed information structure, adding curvature to the indirect utility in Figure 2 would only decrease the expected agent utility under that information structure—thus diminishing the incentives to exert effort. And moving the kink increases the payoff from not exerting effort by more than the expected utility from exerting effort.

Proposition 1 illustrates the tensions involved in designing dynamic contracts. As the agent's posterior evolves over time, the priors for the continuation games at different time periods vary, leading to inconsistencies in the scoring rules that maximize incentives for effort across these periods. In particular, as illustrated in Section 3.1, to implement maximum effort in dynamic environments, the moral hazard constraints bind at both time 0 and the stopping time τ when the signals are perfectly revealing. A V-shaped scoring rule with a kink at μ_τ^N leads to insufficient incentives for the agent to exert effort at time 0, resulting in the agent not starting work at all. Conversely, a V-shaped scoring rule with a kink at μ_0^N results in insufficient incentives for the agent to exert effort at time τ, causing the agent to stop prematurely. Our illustration demonstrates that to balance the incentives for exerting effort across different time periods, the V-shaped scoring rule may need a kink located at some interior belief. Moreover, as we will show later in Section 4.3, the optimal scoring rule may not take a V-shaped form to provide balanced incentives in dynamic environments.

4 Implementing Maximum Effort via Scoring Rules

This section presents our main results on the implementation of the effort-maximizing contract as a static scoring rule in three canonical environments: stationary, perfect-learning and single-signal. Our proof strategy is to show that for any dynamic contract R, there exists a scoring rule P that provides stronger incentives for exerting effort in every continuation game. Appendix B contains the missing proofs in this section. Section 5 provides a partial converse to these results by showing that dynamic structures can be necessary to implement the effort-maximizing contract when all three conditions are sufficiently violated.

A key step in some of our replacement arguments is identifying a time at which strengthening incentives to exert effort *at that time* induces the agent to work for longer and remains incentive compatible at other times. The challenge is to ensure that the resulting replacement does not stop the agent from working at other times. This task is not immediate; in fact, Section 4.3 describes cases where this replacement may violate the incentives for exerting effort at earlier times. In those cases, we show how to restore those incentives by adding additional scores that provide strictly positive rewards in all states. The resulting scoring rules that implement maximum effort will qualitatively differ between the dynamic model and the corresponding "all-at-once" static model. These differences are driven by the dynamic nature of our environment.

4.1 Stationary Environment

The simplest case of interest is the stationary environment, where the agent's no information belief is $\mu_t^N = D$ for all $t \leq T$. In this simple case, all continuation games at any time t share the same prior belief. But Proposition 1 showed that the prior determines the effort-maximizing scoring rule for all continuation games. Thus, it is immediate that the V-shaped scoring rule with kink at the prior D can implement any implementable effort level.

Theorem 1 (Stationary Environment)**.**
In the stationary environment, a V-shaped scoring rule with kink at D is effort-maximizing.

Thus, the complexities of interest that emerge in our other two environments of interest (perfect-learning and single-source) ultimately stem from the dynamics in the agent's beliefs absent signal arrival. In these cases, it need not be the case that effort-maximizing scoring rules have a kink located at the prior D.

4.2 Perfect-learning Environment

Outside of the stationary environment, the no-information belief μ_t^N of the agent drifts over time, so the aforementioned tension in designing effort-maximizing scoring rules *does* arise: In particular, it could be that there exist scoring rules which can implement effort from 0 to $\tilde{t}$ and from $\tilde{t}$ to τ, but no dynamic contract that can implement effort from 0 to τ. This idea was mentioned in Section 3.1, where we argued that (a) a scoring rule with $r_1 < 1$ might be necessary to induce the agent to exert effort from some $\mu_0^N > \mu^*$ to $\mu_t^N = 1/2$ (for instance), while (b) a scoring rule with $r_1 = 1$ would be necessary to induce the agent to exert effort from $\mu_t^N = 1/2$ to $\mu_\tau^N = \frac{c}{\lambda_1^G}$, but (c) no contract (static or dynamic) could induce the agent to exert effort between μ_0^N and $\frac{c}{\lambda_1^G}$. Thus, adjusting the solutions to the relaxed problems is necessary to find a scoring rule that is also an effort-maximizing contract.

We now show that any implementable effort level can be implemented via a scoring rule under perfect-learning. Unlike for the stationary environment, while a V-shaped scoring rule implements maximum effort, its kink need not be at the prior D:

Theorem 2 (Perfect-learning Environment)**.**
In the perfect-learning environment, a V-shaped scoring rule with kink at $\frac{1}{1+r_1} \in (1/2, 1)$ implements maximum effort.

We first sketch the ideas behind the proof of Theorem 2. First, we show that under perfect-learning, incentives to exert effort are maximized if the agent receives the maximum reward in state 0 whenever (a) receiving a bad news signal or (b) no Poisson signal—where we recall that 0 is the state that the agent's posterior belief drifts to absent a Poisson signal:

Lemma 3. *In the perfect-learning environment, for any prior D and any signal arrival probabilities λ, there exists an effort-maximizing contract R with a sequence of menu options $\{r_t^s\}_{t\leq\tau_R,s\in S} \cup \{r_{\tau_R}^N\}$ such that $r_t^B = r_{\tau_R}^N = (1, 0)$ for any $t \leq \tau_R$.*

The argument in the proof of Lemma 3 replaces an arbitrary dynamic contract with one where the agent obtains the full reward in state 0 conditional on either receiving a bad news signal or no signal by time τ_R. Crucially, this outcome occurs with probability 1 if the state is 0 under perfect learning. Therefore, whenever the state is 0, the agent enjoys the full increase in rewards in any continuation game $\mathcal{G}_t$ for exerting effort. By contrast, the increase in the expected reward when not exerting effort in the continuation game $\mathcal{G}_t$ is at most the prior probability of state 0 times the reward increase. Thus, the agent has stronger incentives to exert effort in all continuation games in the replacement, which implies that the replacement induces more effort from the agent.

The next step is to show that offering a single menu option for rewarding good news signals is also sufficient. Intuitively, the dynamic incentives of the agent imply that in any contract R, the agent's reward for receiving a good news signal must decline over time. We show that by decreasing earlier rewards to make the agent's utility from early stopping sufficiently low and increasing later rewards to make the reward from continuing effort sufficiently high, the resulting contract has an implementation as a scoring rule, and the agent's incentive for exerting effort increases weakly in all continuation games given the new contract. Complete formal details are provided in the proof in Appendix B.2.

The proof of Theorem 2 does *not* identify an effort-maximizing choice of r_1, but rather argues that any dynamic contract (within the class identified in Lemma 3) can be replaced by a scoring rule while increasing the incentives to exert effort. We provide some additional results to show how to identify the effort-maximizing choice of r_1 and, in particular, show why this kink need not be at the prior (in contrast the static case). This step also provides formal details behind an interesting dynamic effect in our model alluded to in Section 3.1: for long time horizons, there exist $\mu_1 > \mu_2 > \mu_3$ such that the agent can be incentivized to work when the no information belief would drift from (a) μ_1 to μ_2 or (b) μ_2 to μ_3, but not when this belief would drift from μ_1 to μ_3.

To describe this argument, we momentarily ignore the constraint imposed by the time horizon T. Given a V-shaped scoring rule P with parameters $r_0 = 1$ and $r_1 \in [0, 1]$, recall

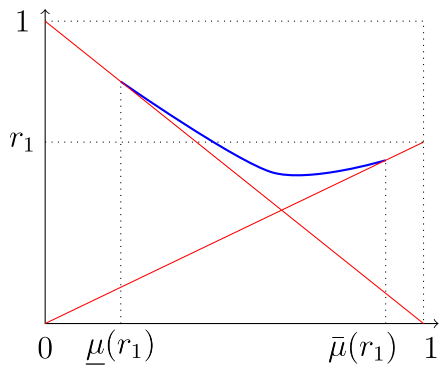

Figure 3: The red lines are the agent's utility $u^N_{t-\Delta}$ when selecting a reward function; the blue curve is the agent's value function when exerting (non-zero) effort optimally, U^+_t. Both are a function of $\mu^N_{t-\Delta}$, the no information belief.

that u^N_t is the agent's utility when not exerting effort after time t. Let U^+_t be the value function of the agent at the prior belief $\mu^N_{t-\Delta}$: That is, the agent's payoff when exerting effort optimally in at least one period starting from (and including) time t. It is straightforward that U^+_t is convex in $\mu^N_{t-\Delta}$, with its derivative between -1 and r_1. We let $\underline{\mu}(r_1) \leq \bar{\mu}(r_1)$ be the beliefs such that U^+_t intersects u^N_t.[11] The agent has incentives to exert effort at time t given scoring rule P if and only if $\mu^N_t \in [\underline{\mu}(r_1), \bar{\mu}(r_1)]$. Figure 3 illustrates how $\underline{\mu}(r_1)$ and $\overline{\mu}(r_1)$ are determined, namely as the intersection between the agent's value function U^+_t and the payoff attainable without exerting any further effort.

Lemma 4. *Both $\underline{\mu}(r_1)$ and $\bar{\mu}(r_1)$ are weakly decreasing in r_1.*

In particular, the agent has incentives to exert effort initially if and only if $D \in [\underline{\mu}(r_1), \bar{\mu}(r_1)]$. Lemma 4 implies that, by increasing the reward r_1 for prediction state 1 correctly in the scoring rule, the agent has incentives to exert effort for longer ($\underline{\mu}(r_1)$ is smaller) but the agent has a weaker incentive to exert effort at time 0 ($\bar{\mu}(r_1)$ is smaller).

Let $\mu^* = \max_r\{\bar{\mu}(r) : \bar{\mu}(r) \leq 1\}$. That is, μ^* is the maximum belief of the agent that can be incentivized to exert effort given any scoring rule. Thus, we can focus on the case where $D \leq \mu^*$. When there is a time horizon constraint T for exerting effort, let $r_T = 1$ if $\underline{\mu}(1) > \mu^N_T$ and let $r_T \in [0,1]$ be the minimum parameter such that $\underline{\mu}(r_T) = \mu^N_T$ otherwise.

Proposition 2 (Effort-Maximizing Scoring Rule)**.**
In the perfect-learning environment, no contract incentivizes the agent to exert effort if

[11] More formally: Assuming that $U^+_t \geq u^N_{t-\Delta}$ for some $t \geq 0$, we define $\underline{\mu}(r_1) = \min\{\mu^N_t, t \geq 0 : U^+_t \geq u^N_{t-\Delta}\}$ and $\bar{\mu}(r_1) = \max\{\mu^N_t, t \geq 0 : U^+_t \geq u^N_{t-\Delta}\}$. If $U^+_t < u^N_{t-\Delta}$ holds for all t, then the agent does not work and we take $\underline{\mu}(r_1) = \overline{\mu}(r_1)$ to be undefined.

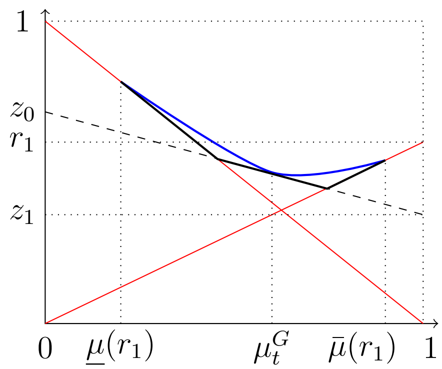

Figure 4: The red lines are the agent's utility $u_{t-\Delta}^N$ for not exerting effort, and the blue curve is the agent's utility U_t^+ from exerting effort in at least one period, both as a function of the no information belief $\mu_{t-\Delta}^N$. The black curve is the agent's utility for not exerting effort in the alternative scoring rule with additional menu option (z_0, z_1).

$D > \mu^*$ or $D < \underline{\mu}(1)$. Otherwise, the effort-maximizing value of the parameter r_1 in a V-shaped scoring rule is the maximum value below r_T such that $\bar{\mu}(r_1) \geq D$.

Intuitively, the effort-maximizing parameter r_1 maximizes the stopping time subject to the constraint that the agent has incentives to exert effort at time 0. If the latter does not bind, then $r_1 = 1$; if it does, then r_1 is set so that the agent is indifferent between not working at all and working until their belief reaches $\underline{\mu}(1)$ absent a signal. Stepping back, we see that the kink of the scoring rule depends on the parameter r_1, which is chosen to balance the agent's incentive to exert effort at both time 0 and the stopping time τ_R. This balance drives our earlier observation that the kink need not be at the prior.

4.3 Single-signal Environment

We now consider the single-signal environment where the agent's belief drifts towards 0 in the absence of a Poisson signal and jumps towards 1 when a Poisson signal arrives. We emphasize that we do *not* assume the signal is fully revealing.

Theorem 3 (Single-signal Environment)**.**
In the single-signal environment, there exists a scoring rule implementing maximum effort.

A notable feature is that in the single-signal environment, although the effort-maximizing contract can be implemented as a static scoring rule, this scoring rule may not be V-shaped. Put differently, effort-maximizing scoring rules need not simply involve the agent guessing the state and being rewarded for a correct guess. It may be necessary to reward the agent *even* when the guess is wrong. Consider an interpretation of the minimum reward across the two states as the "base reward" and the difference between the rewards as the "bonus reward." From this perspective, our results indicate that it may be necessary to consider scoring rules where the base reward is strictly positive. This observation may seem counterintuitive, as providing a strictly positive base reward strictly decreases the maximum bonus reward for the agent since rewards are constrained to the unit interval. In principle, providing a positive base reward should lower the agent's utility increase when correctly guessing the state and, hence, subsequently lower the agent's incentive for exerting effort.

The correct intuition is as follows: while providing a strictly positive base reward at time t decreases the agent's incentive to exert effort at time t, it increases the agent's incentive to exert effort at earlier times $t' < t$; indeed, the agent may expect high rewards for exerting effort, even if mistakes are made in guessing a state after receiving imperfect signals. This modification induces more effort if the agent's incentive constraint for exerting effort initially binds at time 0 but becomes slack at intermediate time $t \in (0, \tau_R)$. As illustrated in Figure 4 and in Section 4.2, by implementing the effort-maximizing V-shaped scoring rule, the agent's incentive for exerting effort is binding only at the extreme time 0 with belief $D = \bar{\mu}(r_1)$ and time τ_R with belief $\mu^N_{\tau_R} = \underline{\mu}(r_1)$. In this case, since the signals are not perfectly revealing, there may exist a time t such that $\mu^G_t < D$. By providing an additional menu option with a strictly positive base reward in the scoring rule to increase the agent's utility at beliefs μ^G_t (e.g., the additional menu option (z_0, z_1) illustrated in Figure 4), the agent's incentive constraint for exerting effort at time 0 is relaxed and the contract thus provides the agent incentives to exert effort following more extreme prior beliefs without influencing the stopping belief $\underline{\mu}(r_1)$.

We now illustrate the main ideas for proving Theorem 3. Given any (static) scoring rule, the utility u^N_t of the agent from not exerting effort is a convex function of his no information belief μ^N_t (see Lemma 2). We first show that it is without loss to focus on *dynamic* contracts where u^N_t is convex in μ^N_t.

Lemma 5 (Convexity in Utilities).
In the single-signal environment, an effort-maximizing contract exists with the no-information utility u^N_t convex in μ^N_t.

Intuitively, in the effort-maximizing contract, if the no-information utility is not convex, one of the following cases holds at time $\bar{t}$, the earliest time such that the utility is non-convex:

- The agent's incentive to exert effort is slack at time $\bar{t}+\Delta$. In this case, we can increase the no-information utility at time $\bar{t}$ by increasing the rewards in menu option r^N_t such that either (a) the incentive for exerting effort at time $\bar{t} + \Delta$ will bind, or (b) the no-information utility will become convex at $\bar{t}$.

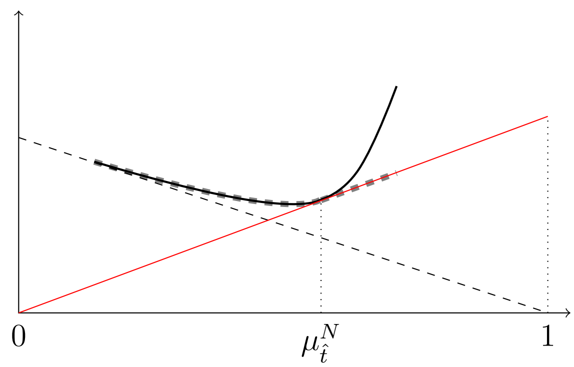

Figure 5: The black curve is the convex no-information utility of the agent, and $\hat{t}$ is the minimum time with a bounded tangent line (red line). The thick dashed line is the no-information utility of the agent given a feasible scoring rule that offers a menu option that corresponds to the red line instead of the black curve for belief $\mu \geq \mu_{\hat{t}}^N$.

- The agent's incentive for exerting effort is slack at time $\bar{t} + \Delta$. In this case, the combination of the no-information utility function's non-convexity and the contract's incentive constraint actually implies that the agent will have a strict incentive to stop exerting effort at time $\bar{t}$. See Figure 7 in Appendix B.3 for an illustration.

While convexity of the no-information utility function does imply that the effort-maximizing contract has an implementation as a scoring rule—since convex functions are equal to the upper envelope of the linear functions below them—the resulting scoring rule need not satisfy the reward constraint, i.e., that $r_\theta \in [0, 1]$. For instance, consider a convex no-information utility function $u_t^N = (\mu_t^N)^2$. A simple dynamic contract that implements this no-information utility function is to offer a constant reward $(\mu_t^N)^2$ at time t regardless of the realization of the state. However, to implement this utility function using a scoring rule, by Lemma 2, the menu option for belief $\mu \in [0, 1]$ must be $(-\mu^2, 2\mu - \mu^2)$, which violates the ex post individual rationality constraint.

Primarily, a violation of the constraint that rewards lie in $[0, 1]$ emerges because the no-information utility function is too convex. In this case, we can flatten the no-information utility by decreasing the reward to the agent at earlier times. We can show that by flattening the no-information utility, the decrease in no-information utility is weakly larger than the decrease in continuation payoff in all continuation games, and hence, the agent has stronger incentives to exert effort. Figure 5 illustrates this idea, with details provided in Appendix B.3.

5 Dynamic Rewards

5.1 Conditions for the Insufficiency of Scoring Rule

This section shows that the effort-maximizing scoring rule must be dynamic when (i) signals are noisy and (ii) the drift is sufficiently slow. Specifically, we consider cases where both good and bad news signals can arrive with strictly positive probability if the agent exerts effort, so that $\lambda_1^G > \lambda_0^G > 0$ and $0 < \lambda_1^B < \lambda_0^B$. Recall that we assumed without loss that $\lambda_1^G + \lambda_1^B > \lambda_0^G + \lambda_0^B$, so beliefs drift toward state 0 absent news. Taking the drift to be slow means that $\lambda_1^G + \lambda_1^B$ is only slightly larger than $\lambda_0^G + \lambda_0^B$. Signals being noisy means that they do not reveal the state.

Let $\mu_{\lambda,c} \triangleq \min\{\frac{1}{2}, \frac{c}{\lambda_1^G - \lambda_0^G}\}$ and let $T_{\lambda,D,c}$ be the maximum time such that $\mu^N_{T_{\lambda,D,c}-\Delta} \geq \mu_{\lambda,c}$. Intuitively, $T_{\lambda,D,c}$ is the maximum calendar time such that the agent can be incentivized to exert effort in any contract when the prior is D.

Lemma 6. *The stopping time τ_R satisfies $\tau_R \leq T_{\lambda,D,c}$, given any prior $D \in (0,1)$, arrival rates λ, cost of effort c, and contract R with rewards belonging to $[0,1]$.*

Appendix OA 1 contains this section's missing proofs. The following result provides our sufficient conditions necessitating the use of complex dynamic structures in the effort-maximizing contract:

Theorem 4 (Strictly Less Effort Under Scoring Rules)**.**
Fix any prior $D \in (0, \frac{1}{2})$, any cost of effort c, and any constant $\kappa_0 > 0$, $\frac{1}{4\Delta} \geq \bar{\kappa}_1 > \underline{\kappa}_1 > 0$. There exists $\epsilon > 0$ such that for any λ satisfying:

- $\lambda_1^G - \lambda_0^G \geq \frac{1}{D}(c + \kappa_0)$; *(sufficient-incentive)*
- $\lambda_1^B, \lambda_0^B, \lambda_0^G, \lambda_1^G \in [\underline{\kappa}_1, \bar{\kappa}_1]$; *(noisy-signal)*
- $\lambda_1^G + \lambda_1^B \in (\lambda_0^G + \lambda_0^B, \lambda_0^G + \lambda_0^B + \epsilon)$, *(slow-drift)*

and $T \geq T_{\lambda,D,c}$, any static scoring rule implements effort strictly less than the maximum.

We discuss the role of each of the four conditions on λ. Two essentially avoid degenerate cases. Specifically, the condition $T \geq T_{\lambda,D,c}$ implies the time horizon T will not be a binding constraint for the agent to exert effort. The sufficient incentive condition avoids trivial solutions by ensuring that the agent has incentives to exert effort for a strictly positive length of time in the effort-maximizing contract. More substantive in light of our previous results, however, are the other two conditions: The noisy-signal condition rules out the perfect-learning environment and the single-signal environment, and the slow-drift

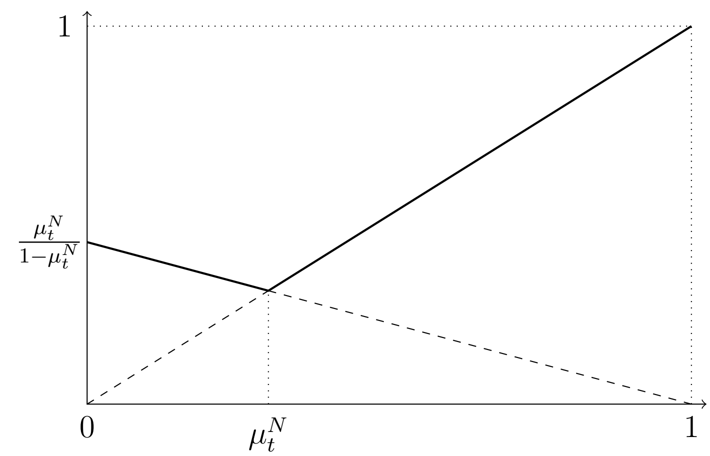

Figure 6: Illustration of myopic-incentive contract. The solid line is the expected reward of the agent, as a function of his belief, for not exerting effort and reporting his belief truthfully at time t.

condition rules out the stationary environment. Moreover, the single-signal environment can also be viewed as the extreme opposite of the slow-drift condition, since the difference in arrival rates of the signals is maximal and the belief drifts to 0 quickly in the absence of a Poisson arrival.

A substantive assumption, but one that appears relatively harmless, is that $D \in (0, \frac{1}{2})$. Note that combining this with our restriction to environments where the belief absent a Poisson signal drifts towards 0, our theorem only applies to environments where the agent's posterior belief absent a Poisson signal becomes more polarized towards the more likely states given the prior belief. This restriction ensures that the following myopic-incentive contract we define is incentive compatible for the agent with rewards in $[0, 1]$.

We prove Theorem 4 by identifying a particular contract that can outperform any static scoring rule. Recall that an arbitrary dynamic contract can be represented via a sequence of reward options that vary over time.

Definition 3 (Myopic-incentive Contract)**.**
When prior $D \in (0, \frac{1}{2})$, a contract R with menu options $\{r_t^s\}_{t\leq\tau_R, s\in S} \cup \{r_{\tau_R}^N\}$ is a myopic-incentive contract *if $r_{\tau_R}^N = r_t^G = (1, 0)$ and $r_t^B = (\frac{\mu_t^N}{1-\mu_t^N}, 0)$ for any $t \geq 0$.*

Figure 6 illustrates the reward offered at a given time t under the myopic-incentive contract. Note that the condition that the belief absent a Poisson signal becomes increasingly polarized implies that the rewards in menu options decrease over time. This property is necessary for the constructed contract to be incentive-compatible.

Notice that the menu options in the myopic-incentive contract at any time t resembles the V-shaped scoring rules in Section 3.2 that implement effort, if possible, in the continuation game when the agent's belief is μ_t^N. However, the myopic-incentive contract typically does not provide maximal incentives to exert effort in all continuation games. Indeed, myopic-incentive contracts involve strictly decreasing rewards, so that an agent who observes a Poisson signal after time t receives a reward less under the effort-maximizing scoring rule for continuation game $\mathcal{G}_t$. While this implies the myopic-incentive contract may not implement maximal effort, we show that they nevertheless provide incentives *close* to the effort-maximizing contract under slow-drift:

Lemma 7 (Approximate Effort Maximization of Myopic-incentive Contracts).
Given any prior $D \in (0, \frac{1}{2})$, any cost of effort c, any constant $\underline{\kappa}_1 > 0$, and any $\eta > 0$, there exists $\epsilon > 0$ such that for any $T \geq T_{\lambda,D,c}$ and any λ satisfying that $\lambda_\theta^s \leq \frac{1}{4\Delta}$ for all $s \in S$ and $\theta \in \{0, 1\}$, and

- $\lambda_1^G - \lambda_0^G \geq \frac{1}{D}(c + \kappa_0)$; *(sufficient-incentive)*
- $\lambda_1^G + \lambda_1^B \leq \lambda_0^G + \lambda_0^B + \epsilon$, *(slow-drift)*

letting R be the myopic-incentive contract (Definition 3), we have $\mu_{\tau_R}^N - \mu_{T_{\lambda,D,c}}^N \leq \eta$.

When the drift of the no information belief is sufficiently slow compared to the arrival rates of the Poisson signals, with sufficiently high probability, the agent will receive a Poisson signal before the no information belief drifts far away from his initial belief. In this case, the decrease in rewards over time in the myopic-incentive contract does not significantly weaken the agent's incentive for exerting effort. Hence, the myopic-incentive contract approximately implements maximal effort.

By contrast, if the agent instead faced a contract that can be implemented as a scoring rule, given stopping time τ_{R^*} in the effort-maximizing contract R^*, this scoring rule must be close to the effort-maximizing scoring rule in the continuation game $\mathcal{G}_t$ for any t that is sufficiently close to τ_{R^*} in order to incentivize the agent to exert effort in continuation game $\mathcal{G}_t$. However, when the signals are noisy, as illustrated in Appendix OA 4, such scoring rules fail to provide sufficient incentives for the agent to exert effort at time 0, leading to a contradiction. Intuitively, the myopic-incentive contract avoids this conflict in incentives by providing higher rewards to the agent upon receiving a bad new signal B without affecting the agent's incentive to report the acquired information truthfully. In particular, higher rewards upon Poisson signal arrivals strengthen the agent's incentives to exert effort at any time t. Interestingly, when $\varepsilon = 0$, the environment is stationary, in which case a scoring rule is again optimal; however, as long as drift is non-zero, scoring rules will fail to approximate the optimum in contrast to the myopic-incentive contracts.

5.2 Effort-Maximizing Dynamic Contracts

Lemma 7 shows that the myopic-incentive contract approximately implements maximum effort when the drift is slow and the belief absent a Poisson signal is polarizing. Although the stated contract is generally not fully effort-maximizing, a similar decreasing reward structure does, in fact, characterize effort-maximizing mechanisms.

We now present a formal characterization of reward dynamics in effort-maximizing dynamic contracts, applicable when (static) scoring rules cannot implement maximum effort. Recall that any menu option r has a representation as a tuple (r_0, r_1), where r_0 is provided in state 0 and r_1 is provided in state 1. For any pair of menu options r, r', we define $r' \preceq r$ if $r'_0 \le r_0$ and $r'_1 \le r_1$. That is, $r' \preceq r$ if r' is at most r in all components.

The qualitative features in myopic-incentive contracts preserved in the effort-maximizing dynamic contracts are:

1. Decreasing rewards for "bad news" signals. That is, the effort-maximizing sequence of rewards for signal B, denoted as r_t^B, satisfies the condition that $r_t^B \preceq r_{t'}^B$ for any $t \le t' \le \tau_R$. Moreover, this decrease in rewards has a particular structure: the contract initially maintains the maximum score for state 0 and reduces the reward for state 1. This is followed by a decrease in the reward for state 0 while maintaining the minimum score for state 1.

2. Maximal rewards for "good news" signals subject to incentive constraints. That is, the rewards for receiving "good news" signals are uniquely determined by the "bad news" signals. Specifically, the reward r_t^G at time t is determined by finding the reward vector that maximizes the expected reward for posterior belief μ_t^G, subject to the constraints that the no-information belief $\mu_{t'}^N$ weakly prefers the option $r_{t'}^B$ over r_t^G for any time $t' \le t$.

Theorem 5 shows that contracts with these features maximize effort.

The only feature of myopic-incentive contracts that does not extend to effort-maximizing contracts is the rate of decrease for rewards following "bad news" signals. In myopic-incentive contracts, rewards for "bad news" signals are chosen such that the menu options (r_t^G, r_t^B) offered at time t consist of a V-shaped scoring rule with a kink at belief μ_t^N. Such a rate of decrease is shown to approximately implement maximum effort (Lemma 7), but it implements strictly less than the maximum in general. In general environments, the effort-maximizing rate of decreasing rewards for bad news signals depends on primitives and may not admit closed-form characterizations. However, such rewards can be computed efficiently by solving a family of linear programs (Appendix OA 2).

Theorem 5 (Effort-Maximizing Dynamic Contracts).
For any prior D and any signal arrival rates λ, there exists an effort-maximizing contract R with optimal stopping time τ_R and a sequence of menu options $\{r_t^s\}_{t\leq\tau_R, s\in S} \cup \{r_{\tau_R}^N\}$ with $r_{\tau_R}^N = r_{\tau_R}^B$ such that

1. *decreasing rewards for signal B: $r_{t'}^B \preceq r_t^B$ for all $t' \geq t$; and $r_{t,0}^B = 1$ if $r_{t,1}^B > 0$;*

2. *maximal rewards for signal G: for any $t \leq \tau_R$,*

$$r_t^G = \underset{r:\Theta\to[0,1]}{\arg\max}\ u(\mu_t^s, r)$$
$$\text{s.t.} \quad u(\mu_{t'}^N, r_{t'}^B) \geq u(\mu_{t'}^N, r), \quad \forall t' \in [0, t].$$

Compared to the menu representation in Lemma 2, the simplification in Theorem 5 is that effort-maximizing contracts involve a decreasing sequence of rewards when receiving a bad news signal B in both states. In fact, rewards following B decrease first in state 1, and subsequently in state 0 once the reward in state 1 hits 0. Furthermore, the rewards for good news signals are uniquely determined based on the rewards for bad news signals. Note that the rewards for bad new signals are only weakly decreasing. Indeed, the theorem covers cases where our main theorems imply that the effort-maximizing contract is attained by keeping the rewards r_t^B unchanged over time.

6 Extensions and Discussions

We now turn to some extensions of the main model, particularly those related to randomized contracts, non-stationary information acquisition, and ex-ante reward constraints. We show how to generalize our results and techniques to such environments. Finally, we conclude our paper and propose several open questions in Section 7.

6.1 Randomized Contracts

Our goal has been to determine the maximum effort implementable within the class of contracts defined in Section 2.2. As histories only include messages sent by the agent, we implicitly rule out the use of randomization (as in, for instance, Deb et al. (2018)). Deterministic mechanisms have significant practical appeal, as it is not always obvious what a randomization might correspond to and how a mechanism designer might implement this. These issues have been discussed extensively in the contracting literature; we refer the reader to discussions in Laffont and Martimort (2002) as well as Bester and Strausz (2001) on this point to avoid detours. Now, randomization provides no additional benefit in implementing maximum effort with probability 1. But it is natural to ask whether some designer objectives may yield benefits to randomization. Our analysis speaks to this question as well.

We discuss randomization formally. Let $\varsigma = \{\varsigma_t\}_{t \leq T}$ be a sequence of random variables with ς_t drawn from a uniform distribution in $[0,1]$. A *randomized contract* is a mapping

$$R(\cdot|\varsigma) : \mathcal{H}_T \times \Theta \to [0,1].$$

Crucially, in randomized contracts, at any time t, the history of the randomization device $\{\varsigma_{t'}\}_{t' \leq t}$ is publicly revealed to the agent before determining his choice of effort or the message sent to the designer. The randomization revealed before t affects the agent's incentives after time t, and without it, such contracts reduce back to deterministic contracts.

Stopping strategies may be *with* loss under a randomized contract. In particular, the agent may decide whether to work or not depending on the public randomization's past realizations. The agent may also strategically delay exerting effort to wait for the realization of the public randomization. As a result, the simplification of the objective to maximizing the stopping time of the agent is not appropriate for randomized contracts.

Nevertheless, our analysis suggests randomization can expand the set of implementable strategies. We previously observed that it may be possible to get the agent to work from μ_1 to μ_2 and to get the agent to work from μ_2 to μ_3, but not from μ_1 to μ_3. This would occur if the reward necessary to get the agent to work to μ_3 were so high that the agent would "shirk-and-lie" at μ_1. However, the agent may be willing to start working at μ_1, not knowing whether the reward will be "high" or "low"—but once the agent starts working, the designer can randomly inflate or decrease the rewards of the agent. Once time has passed, the outcome of the randomization can be revealed, and if the rewards inflate, the agent has incentives to exert effort for longer absent a Poisson signal arrival—so that the realized stopping time increases for *some* realization of the randomization.

We illustrate this intuition more formally and rigorously for the environment from Section 3.1, where learning is perfect and only good news signal G arrives with positive probability. As discussed, under deterministic contracts, a V-shaped scoring rule with parameters $r_0 = 1$ and $r_1 \in [0,1]$ implements the effort-maximizing contract R; in particular, Section 3.1 characterized when in fact effort-maximizing contracts require $r_1 < 1$. Recall that we denote τ_R as the stopping time in the effort-maximizing deterministic contract and let $\mu^N_{\tau_R}$ be the stopping belief when no Poisson signal is observed. Let $\delta \in (0, 1 - r_1]$ be the maximum number such that (1) the agent has strict incentives to exert effort until time $\frac{2\tau_R}{3}$ absent signal arrival given menu options $(1,0)$ and $(0, r_1 - \delta)$; and (2) the agent can be incentivized to exert effort given menu options $(1,0)$ and $(0,r_1+\delta)$ given belief $\mu^N_{\frac{\tau_R}{2}}$. Now consider the randomized contract $\hat{R}$ that provides menu options $(1,0)$ and $(0,r_1-\delta)$ from time 0 to $\frac{\tau_R}{2}$, and after time $\frac{\tau_R}{2}$, offers menu options $(1,0)$ and $(0,r_1+\delta)$ with probability ϵ^2, offers menu options $(1,0)$ and $(0,0)$ with probability ϵ, and offers the same menu options $(1,0)$ and $(0,r_1-\delta)$ otherwise. With sufficiently small $\epsilon>0$, the agent still has incentives to exert effort at any time $t\le\frac{\tau_R}{2}$. Moreover, after time $\frac{\tau_R}{2}$, with probability ϵ^2, the realized menu options are $(1,0)$ and $(0,r_1+\delta)$, and the agent can be incentivized to exert effort to a time strictly larger than τ_R in the absence of a Poisson signal.[12]

6.2 Non-invariant Environments

Our model assumes that both the cost of acquiring information and the signal arrival probabilities when exerting effort are fixed over time. On the other hand, many of our techniques and results do not rely on these assumptions, particularly as we focus on maximizing the incentive for the agent to exert effort. Our results extend unchanged if the cost of exerting effort is $c(\tilde{t})\Delta$, whenever the agent has exerted effort for $\tilde{t}$ units of time and $c(\cdot)$ is a non-decreasing function. In this case, the agent's strategy again without loss is a stopping time, and identical arguments imply that scoring rules maximize the incentive to exert effort under any of the three environments discussed.

Moreover, we can also allow for more general time-dependent cost functions. In particular, there exist settings where the cost of acquiring information is lower closer to the decision deadline, regardless of the previous efforts exerted by the agent. In these applications, given a dynamic contract, the best response of the agent may not be a stopping strategy. Scoring rules still implement maximal effort in this extension under one of three conditions in Section 4—in the sense that given any dynamic contract R and any best response of the agent, there exists another contract $\hat{R}$ that can be implemented as a scoring rule, and the agent's best response in contract $\hat{R}$ first order stochastic dominates his best response in contract R.[13] Therefore, the information acquired under contract $\hat{R}$ is always weakly Blackwell more informative compared to the information acquired under contract R.

[12] This kind of modification may be of interest to a designer who only values extreme posterior beliefs. For instance, suppose the designer faced a decision problem where the possible decisions belonged to a set $A=\{0,1\}$. Consider a designer payoff function of $v(0,\theta)=0$ for all $\theta\in\Theta$, and $v(1,0)=1, v(1,1)=-\frac{1-\mu^N_{\tau_R}}{\mu^N_{\tau_R}}$; plainly, the designer only seeks to change her action from 1 to 0 if the posterior belief is below $\mu^N_{\tau_R}$. In this case, the payoff under any deterministic contract is 0. However, under the randomized contract outlined, the designer would obtain a positive payoff when implementing the identified randomized contract.

[13] Let $z_t,\hat{z}_t\in\{0,1\}$ be the effort decision of the agent given contracts $R,\hat{R}$ respective conditional on not receiving any Poisson signal before t. We can show that given any time t, $\sum_{i\le t} z_t \le \sum_{i\le t}\hat{z}_t$.

We can similarly allow for time dependence in the informational environment. Specifically, we can allow for the arrival rates of signals at any time to depend on the amount of effort the agent has exerted until that point. That is, suppose that if the agent has exerted effort for $\tilde{t}$ units of time, then exerting effort produces a good news signal arrives in state θ with probability $\lambda^G_{\theta,\tilde{t}}\Delta$, and a bad news signal with probability $\lambda^B_{\theta,\tilde{t}}\Delta$ (and no signal with complementary probability). While seemingly minor, this modification induces more richness in the set of possible terminal beliefs as a function of the effort history—for instance, if the terminal beliefs are always in the set $\{\underline{p}, \overline{p}\}$, despite drifting over time.

Our proof techniques did not make use of the particular belief paths induced by constant arrival rates and hence extend to this case, with the minor exception that Theorem 3 requires μ^G_t to be weakly monotone as a function of time (a property that holds when the arrival rate is constant). Otherwise, as long as parameters stay within each environment articulated in Section 4, the proofs of these results extend unchanged.

6.3 Valuing Transfers and Ex-Ante Reward Constraints

As our analysis focused entirely on the problem of implementing maximum effort, we have (a) avoided degeneracy by imposing bounds in the rewards to the agent and (b) not considered the ex-ante expected value of transfers (as stated in the introduction, either viewing them as non-monetary or zero-sum). An alternative to part (a) would be to impose an *ex-ante* reward constraint on the designer rather than ex-post as in our main model; solving this problem for any bound allows the designer to identify tradeoffs between expected payments and information acquisition, thus addressing part (b).

Even if there is an ex-ante upper bound on the payments the agent can be provided, limited liability is still necessary to avoid degeneracy—i.e., the reward should satisfy $r \geq 0$ ex-post. To implement an arbitrary stopping strategy τ, simply find a decision problem where exerting effort until time τ is optimal. If the only constraint is that payments should have expectation zero, then τ could be implemented using a "selling-the-firm-to-the-agent" mechanism—that is, charging the agent the ex-ante expected utility from that decision problem up front and then providing a reward (or punishment) equal to the ex-post utility under the decision problem. Thus, we discuss issues under this modification by focusing on the model with ex-ante reward constraints and ex-post limited liability constraints.

We show that in the special case where the conditions in both perfect-learning environments and single-signal environments are satisfied, i.e., when $\lambda^B_0 = \lambda^B_1 = \lambda^G_0 = 0$ and only $\lambda^G_1 > 0$, the effort-maximizing contract can also be implemented as a scoring rule.

Appendix OA 5 contains details for this claim.[14] This observation illustrates that the optimality of static contracts extends under ex-ante reward constraints under natural assumptions. Continuing with the above, these results also imply that when only $\lambda_1^G > 0$, the value of ex-ante expected payments implementing some effort level τ cannot be minimized by resorting to dynamic contracts. We conjecture that our result also extends when only conditions in one of our canonical environments are satisfied (rather than both), although the formal arguments for this conjecture may require additional novel ideas.

7 Conclusion

We have articulated how dynamic rewards can expand the set of implementable strategies in a simple yet fundamentally dynamic information acquisition problem. The economic importance of contracting for information acquisition is self-evident, and most natural stories for why information acquisition is costly involve some dynamic element. Our goal has been to take such dynamics seriously, for learning technologies with a natural interpretation in terms of investigators seeking evidence, and under a class of contracts reflecting the provision of bonuses for correct advice.

As our focus is on contracting under a general class of mechanisms, a fundamental difficulty underlying our exercise is the lack of any natural structure (e.g., stationarity) under an arbitrary dynamic contract. Such assumptions are often critical in similar settings. Despite this fundamental challenge, we provided simple, economically meaningful conditions such that maximum effort is implementable by a scoring rule and explained the extent to which these conditions are necessary for this conclusion to hold.

There are many natural avenues for future work. Empirically, our results show how learning technologies influence the benefits of time variation to bonuses. A natural question is whether considerations our model has not captured influence the extent to which contracts do or do not involve such features. Theoretically, accommodating both dynamic moral hazard and adverse selection required us to focus on a class of technologies with recognized significance but less generality than, for instance, Chambers and Lambert (2021). Still, we do not doubt similar conclusions could emerge under different information arrival processes. More broadly, we view questions regarding whether or not simple contracts are limited in power relative to dynamic mechanisms as a worthwhile agenda overall. Any further insights on these questions would prove valuable toward understanding how dynamics influence mechanism design for information acquisition, for both theory and practice.

[14] This result follows immediately from Proposition 1 of Gerardi and Maestri (2012), but we present an independent proof for completeness.

References

Ball, I. and Knoepfle, J. (2024). Should the timing of inspections be predictable? Working Paper.

Bardhi, A., Guo, Y., and Strulovici, B. (2024). Early-career discrimination: Spiraling or self-correcting? Working paper.

Bergemann, D. and Hege, U. (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance*, 22:703–735.

Bergemann, D. and Hege, U. (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics*, 36(4):719–752.

Bergemann, D. and Välimäki, J. (2002). Information acquisition and efficient mechanism design. *Econometrica*, 70(3):1007–1033.

Bester, H. and Strausz, R. (2001). Contracting with imperfect commitment and the revelation principle: The single agent case. *Econometrica*, 69(4):1077–1098.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3.

Carroll, G. (2019). Robust incentives for information acquisition. *Journal of Economic Theory*, 181:382–420.

Chade, H. and Kovrijnykh, N. (2016). Delegated information acquisition with moral hazard. *Journal of Economic Theory*, 162:55–92.

Chambers, C. P. and Lambert, N. S. (2021). Dynamic belief elicitation. *Econometrica*, 89(1):375–414.

Che, Y.-k. and Mierendorff, K. (2019). Optimal dynamic allocation of attention. *American Economic Review*, 109(8):2993–3029.

Chen, Y. and Yu, F.-Y. (2021). Optimal scoring rule design. *Working Paper.*

Crémer, J. (1987). Auctions with contingent payments: Comment. *American Economic Review*, 77(4):746.

Crémer, J. and Khalil, F. (1992). Gathering information before signing a contract. *The American Economic Review*, 82(3):566–578.

Damiano, E., Li, H., and Suen, W. (2020). Learning while experimenting. *The Economic Journal*, 130(625):65–92.

Dasgupta, S. (2023). Optimal test design for knowledge-based screening. *Working Paper*.

Deb, R. and Mishra, D. (2014). Implementation with contingent contracts. *Econometrica*, 82(6):2371–2393.

Deb, R., Pai, M., and Said, M. (2018). Evaluating strategic forecasters. *American Economic Review*, 108(10):3057–3103.

Deb, R., Pai, M., and Said, M. (2023). Indirect persuasion. *Working Paper*.

DeMarzo, P., Kremer, I., and Skrzypacz, A. (2005). Bidding with securities— auctions and security design. *American Economic Review*, 95(4):936–959.

Gerardi, D. and Maestri, L. (2012). A principal–agent model of sequential testing. *Theoretical Economics*, 7:425–463.

Guo, Y. (2016). Dynamic delegation of experimentation. *American Economic Review*, 106(8):1969–2008.

Häfner, S. and Taylor, C. (2022). On young turks and yes men: Optimal contracting for advice. *Rand Journal of Economics*, 53(1):63–94.

Halac, M., Kartik, N., and Liu, Q. (2016). Optimal contracts for experimentation. *Review of Economic Studies*, 83:601–653.

Halac, M. and Prat, A. (2016). Managerial attention and worker performance. *American Economic Review*, 106(10):3104–3232.

Hartline, J. D., Shan, L., Li, Y., and Wu, Y. (2023). Optimal scoring rules for multi-dimensional effort. In *Thirty Sixth Annual Conference on Learning Theory*, pages 2624–2650.

Henry, E. and Ottaviani, M. (2019). Research and the approval process: The organization of persuasion. *The American Economic Review*, 109(3):911–955.

Hörner, J. and Samuelson, L. (2013). Incentives for experimenting agents. *The RAND Journal of Economics*, 44(4):632–663.

Keller, G. and Rady, S. (2015). Breakdowns. *Theoretical Economics*, 10(1):175–202.

Keller, G., Rady, S., and Cripps, M. (2005). Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68.

Krähmer, D. and Strausz, R. (2011). Optimal procurement contracts with pre-project planning. *The Review of Economic Studies*, 78(3):1015–1041.

Laffont, J.-J. and Martimort, D. (2002). *Theory of Incentives*. Princeton University Press.

Lambert, N. S. (2022). Elicitation and evaluation of statistical forecasts. *revised and resubmitted at Econometrica*.

Li, Y., Hartline, J. D., Shan, L., and Wu, Y. (2022). Optimization of scoring rules. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 988–989.

Lizzeri, A., Shmaya, E., and Yariv, L. (2024). Disentangling exploration from exploitation. Working paper.

McCarthy, J. (1956). Measures of the value of information. *Proceedings of the National Academy of Sciences of the United States of America*, 42(9):654.

McClellan, A. (2022). Experimentation and approval mechanisms. *Econometrica*, 90(5):2215–2247.

Mylovanov, T. and Zapechelnyuk, A. (2017). Optimal allocation with ex post verification and limited penalties. *American Economic Review*, 107(9):2666–2694.

Neyman, E., Noarov, G., and Weinberg, S. M. (2021). Binary scoring rules that incentivize precision. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 718–733.

Osband, K. (1989). Optimal forecasting incentives. *Journal of Political Economy*, 97(5):1091–1112.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801.

Skrzypacz, A. (2013). Auctions with contingent payments - an overview. *International Journal of Industrial Organization*, 31(5):666–675.

Stahl, J. (2018). *Impact of Consulting on the Performance of Small and Medium-Sized Enterprises*. PhD thesis, University of Halle-Wittenberg.

Strulovici, B. (2010). Learning while voting: Determinants of collective experimentation. *Econometrica*, 78(3):933–971.

Szalay, D. (2005). The economics of clear advice and extreme options. *Review of Economic Studies*, 72(4):1173–1198.

Whitmeyer, M. and Zhang, K. (2023). Buying opinions. Working paper.

Yoder, N. (2022). Designing incentives for heterogeneous researchers. *Journal of Political Economy*, 130(8):2018–2054.

Zermeno, L. (2011). A principal-expert model and the value of menus. *working paper*.

A Additional Preliminaries

A.1 Menu Representation

Proof of Lemma 2. Recall that by Lemma 1, it is without loss to assume that the agent uses a stopping strategy. In particular, if the agent were to randomize at any time, then since he must be indifferent between actions, he would be wiling to continue exerting effort at any point of indifference. Thus, the agent must follow a deterministic strategy in any contract. By the revelation principle, it is without loss to focus on contracts where the agent truthfully reports whether he exerts effort or not and the received signal at each time t. That is, the message space at any time t is $\{0, 1\} \times \{G, B, N\}$ where 1 represents exerting effort and 0 represents not exerting effort. Since the agent's strategy is deterministic, this mechanism is as well.

Next we construct the sequence of menu options $\{r_t^s\}_{t \le \tau_R, s \in S} \cup \{r_{\tau_R}^N\}$ that corresponds to a contract R that induces truth-telling. Let τ_R be the stopping time of contract R. For any time $t \le \tau_R$, for any signal $s \in S$, let h_t^s be the history of reports the agent sends if he receives signal s at time t. That is, in history h_t^s, the agent sends $(1, N)$ before time t, $(1, s)$ at time t, and $(0, N)$ after time t. Let $h_{\tau_R}^N$ be the history of reports the agent sends if he didn't receive any Poisson signal at all.

For any time $t \le \tau_R$ and any signal $s \in S$, let $r_t^s = R(h_t^s, \cdot)$ and let $r_{\tau_R}^N = R(h_{\tau_R}^N, \cdot)$. Given this constructed sequence of menu options, the incentive constraints in Eq. (IC) is satisfied since contract R induces truth-telling. Moreover, it is easy to verify that the agent's utility for stopping effort at any time $t \le \tau_R$ is the same given both the menu representation and the original contract R. Therefore, the stopping time of the agent given this menu representation is also τ_R. □

A.2 Details Behind Section 3.1

Here we provide some additional details behind the calculations in Section 3.1. We consider any scoring rule which involves the choice between $(r_0, 0)$ and $(0, r_1)$, with the former being chosen when stopping in the absence of any signal arrival and the latter being chosen if one does occur. We note that stopping and accepting contract $(r_0, 0)$ delivers payoff $r_0(1 - \mu_t^N)$; if, at time τ, the agent continues for a length Δ and then stops, the payoff is:

$$-c\Delta + (1 - \lambda_1^G \mu_\tau^N \Delta) r_0 (1 - \mu_{\tau+\Delta}^N) + \lambda_1^G \mu_\tau^N \Delta r_1.$$

Imposing indifference between stopping and continuing yields:

$$r_0\frac{\mu_\tau^N-\mu_{\tau+\Delta}^N}{\Delta}+\lambda_1^G\mu_\tau^N r_1-\lambda_1^G\mu_\tau^N(1-\mu_{\tau+\Delta}^N)r_0=c$$

As $\Delta\to 0$, $\frac{\mu_\tau^N-\mu_{\tau+\Delta}^N}{\Delta}\to\dot{\mu}_\tau^N$; substituting in for this expression and using continuity of beliefs yields the expression for the stopping belief and the stopping payoff:

$$r_0\lambda_1^G\mu_\tau^N(1-\mu_\tau^N)+\lambda_1^G\mu_\tau^N r_1-\lambda_1^G\mu_\tau^N(1-\mu_\tau^N)r_0=c.$$

Algebraic manipulations show this coincide with the expression for the stopping belief from the main text. In particular, note that this stopping belief is independent of r_0 (as in the main text). From this, it immediately follows that $r_0=1$ maximizes effort, as it does not influence the length of time the agent works but may make the agent more willing to initially start exerting effort.

We now solve for the agent's value function, $V(\mu_t^N)$, for all agent beliefs $\mu_t^N>\mu_\tau^N$ (assuming the agent works until time τ—recalling that beliefs "drift down"). Writing out the HJB yields:

$$V(\mu_t^N)=-c\Delta+\lambda_1^G\mu_t^N\Delta r_1+(1-\lambda_1^G\mu_t^N\Delta)V(\mu_{t+\Delta}).$$

From this, we obtain the following differential equation:

$$V'(\mu_t^N)\lambda_1^G\mu_t^N(1-\mu_t^N)=-c+\lambda_1^G\mu_t^N(r_1-V(\mu_t^N)).$$

Solving this first-order differential equation gives us the following expression for the value function, up to a constant k (which is pinned down by the condition $V(\mu_\tau^N)=(1-\mu_\tau^N)$):

$$V(\mu_t^N)=k(1-\mu_t^N)+\frac{r_1\lambda_1^G-c+c(1-\mu_t^N)\log\left(\frac{1-\mu_t^N}{\mu_t^N}\right)}{\lambda_1^G}.$$

Note that $V''(\mu_t^N)>0$ for this solution, as well as that $V'(c/(r_1\lambda_1^G))=-1$, so that the value function is everywhere above $(1-\mu_t^N)$. Thus, the agent would *never* shirk and choose option $(1,0)$ prior to time τ; so, as long as the value function is also above $r_1\mu_t^N$, the moral hazard constraint does not bind before τ. This shows that r_1 should be set so that the initial moral hazard constraint holds, as discussed in the main text.

B Effort-Maximizing Contracts as Scoring Rules

B.1 Stationary Environment

Proof of Theorem 1. For any contract R with stopping time $\tau_R \in [0, T]$, to show that there exists a static scoring rule P such that the agent has incentive to exert effort at least until time τ_R given static scoring rule P, it is sufficient to show that there exists a static scoring rule P such that the agent has incentive to exert effort at any continuation game $\mathcal{G}_t$ for any $t \in [0, \tau_R]$.

First note that to maximize the expected score difference for the continuation game at any time t, it is sufficient to consider a static scoring rule. This is because at any time $t' \in [t, \tau_R]$, we can allow the agent to pick any menu option from time t to τ_R. This leads to a static scoring rule where the agent's expected utility at time t for stopping effort immediately is not affected but the continuation utility weakly increases. Finally, by Proposition 1, the effort-maximizing static scoring rule that maximizes the expected score difference is the V-shaped scoring rule P with kink at prior D. □

B.2 Perfect-learning Environment

Proof of Lemma 3. For any contract R, by applying the menu representation in Lemma 2, let $\{r^s_t\}_{t \le \tau_R, s \in S}$ be the set of menu options for receiving Poisson signals and let $r^N_{\tau_R} = (z_0, z_1)$ be the menu option for not receiving any Poisson signal before the stopping time τ_R. Note that in the perfect-learning environment, it is without loss to assume that $r^B_{t,1} = 0$ for any $t \le \tau_R$ since the posterior probability of state 1 is 0 after receiving a Poisson signal B. Now consider another contract $\hat{R}$ with menus menu options $\{\hat{r}^s_t\}_{t \le \tau_R, s \in S \setminus \{N\}}$ and $(\hat{z}_0, \hat{z}_1)$, where

$$(\hat{z}_0, \hat{z}_1) = \arg\max_{z, z' \in [0,1]} z \quad \text{s.t.} \quad \mu^N_{\tau_R} z' + (1 - \mu^N_{\tau_R}) z = u^N_{\tau_R}, \tag{1}$$

and for any time $t \le \tau_R$ and any signal $s \in S$,

$$\hat{r}^s_t = \begin{cases} r^s_t & u(\mu^s_t, r^s_t) \ge u(\mu^s_t, (\hat{z}_0, \hat{z}_1)) \\ (\hat{z}_0, \hat{z}_1) & \text{otherwise.} \end{cases}$$

Essentially, contract R adjusts the reward function for no information belief $\mu^N_{\tau_R}$ such that the reward for state being 0 weakly increases, the reward for state being 1 weakly decreases, and the expected reward remains unchanged. Moreover, at any time $t \le \tau_R$, contract $\hat{R}$ allows the agent to optionally choose the addition option of $(\hat{z}_0, \hat{z}_1)$ to maximize his expected payoff for receiving an informative signal at time t.

It is easy to verify that for any signal $s \in S\backslash\{N\}$ and any time $t \leq \tau_R$, the expected utility of the agent for receiving an informative signal s is weakly higher, and hence, at any time t, the continuation payoff of the agent for exerting effort until time τ_R weakly increases in contract $\hat{R}$. Moreover, at any time t, the expected reward of the agent with belief μ_t^N in contract $\hat{R}$ satisfies $\hat{u}_t^N \leq u_t^N$. This is because by our construction, at any time $t \leq \tau_R$, fewer options are available to the agent in contract $\hat{R}$ except the additional option of $(\hat{z}_0, \hat{z}_1)$, while $u(\mu_t^N, (z_0, z_1)) \geq u(\mu_t^N, (\hat{z}_0, \hat{z}_1))$ since $\mu_t^N \geq \mu_{\tau_R}^N$ and both options (z_0, z_1) and $(\hat{z}_0, \hat{z}_1)$ gives the same expected reward for posterior belief of $\mu_{\tau_R}^N$. Combining both observations, we have $\tau_{\hat{R}} \geq \tau_R$ and $\hat{R}$ is also an effort-maximizing contract.

Note that in optimization program (1), it is easy to verify that $\hat{z}_1 = 0$ if $\hat{z}_0 < 1$. If $\hat{z}_0 = 1$, in this case, at any time t, by the incentive constraint of the agent for any belief μ_t^B, we must have $\hat{r}_{t,0}^B = 1$ as well. Therefore, the agent receives the maximum reward of 1 whenever he receives a bad news signal. In this case, we can also decrease $\hat{z}_1$ and $r_{t,1}^G$ for all $t \geq 0$ by $\hat{z}_1$, which does not affect the agent's incentive for effort and hence the effort-maximizing contract satisfies that $r_t^B = r_{\tau_R}^N = (1, 0)$ for any $t \leq \tau_R$.

Next we will focus on the case when $\hat{z}_0 < 1$ and hence $\hat{z}_1 = 0$. Now consider another contract $\bar{R}$ with menu options $\{\bar{r}_t^s\}_{t \leq \tau_{\hat{R}}, s \in S\backslash\{N\}}$ and $(\bar{z}_0, \bar{z}_1)$, where $(\bar{z}_0, \bar{z}_1) = (1, 0)$ and for any time $t \leq \tau_R$, $\bar{r}_t^G = \hat{r}_t^G$ and $\bar{r}_t^B = (1, 0)$. We show that this weakly improves the agent's incentive to exert effort until time $\tau_{\hat{R}}$ for any $t \leq \tau_{\hat{R}}$. Specifically, for any $t \leq \tau_{\hat{R}}$, the increases in no information payoff is

$$\bar{u}_t^N - \hat{u}_t^N \leq (1 - \mu_t^N)(1 - \hat{r}_{t,0}^B).$$

This is because in contract $\bar{R}$, either the agent prefers the menu option $r_{t'}^G$ for some $t' \geq t$, in which case the reward difference is 0, or the agent prefers the menu option $(1, 0)$, in which case the reward difference is at most $(1 - \mu_t^N)(1 - \hat{r}_{t,0}^B)$ since one feasible option for the agent in contract $\hat{R}$ is $\hat{r}_t^B$ with expected reward at least $(1 - \mu_t^N)\hat{r}_{t,0}^B$. Moreover, for any time $t \leq \tau_{\hat{R}}$, the increases in continuation payoff for exerting effort from $t + \Delta$ until $\tau_{\hat{R}}$ is at least $(1 - \mu_t^N)(1 - \hat{r}_{t,0}^B)$. This is because incentive constraints (IC) imply that $\hat{r}^B t, 0$ must decrease as t increases, which is due to the fact that in the perfect-learning environment, the posterior belief $\mu^B t$ assigns a probability of 1 to the state being 0 at any time t. Therefore, when the state is 0, the reward of the agent is deterministically 1 in contract $\bar{R}$ and the reward of the agent is at most $\hat{r}_{t,0}^B$ in contract $\hat{R}$, implying that the difference in expected reward is at least $(1 - \mu_t^N)(1 - \hat{r}_{t,0}^B)$. Combining the above observations, we have $\tau_{\bar{R}} \geq \tau_{\hat{R}}$, and hence $\bar{R}$ is also effort-maximizing. □

Proof of Theorem 2. By Lemma 3, it suffices to focus on contract R with a sequence of menu options $\{r_t^s\}_{t\le\tau_R,s\in S}\cup\{r_{\tau_R}^N\}$ such that $r_t^B = r_{\tau_R}^N = (1,0)$ for any $t\le\tau_R$. In addition, since signals are perfectly revealing, it is without loss to assume that $r_{t,0}^G = 0$ and the incentive constraints imply that $r_{t,1}^G$ is weakly decreasing in t.

Let $\hat{t}\in[0,\tau_R]$ be the maximum time such that an agent with no information belief $\mu_{\hat{t}}^N$ weakly prefers menu option $r_{\hat{t}}^G$ compared to $r_{\tau_R}^N$. Since both μ_t^N and $r_{t,1}^G$ are weakly decreasing in t, an agent with posterior belief μ_t^N weakly prefers $r_{t'}^G$ compared to $r_{\tau_R}^N$ for any $t,t'\le\hat{t}$, and weakly prefers $r_{\tau_R}^N$ compared to $r_{t'}^G$ for any $t,t'>\hat{t}$. Now consider another contract $\hat{R}$ that offers only two menu options, $r_{\hat{t}}^G$ and $r_{\tau_R}^N$, at every time $t\le\tau_R$. Contract $\hat{R}$ can be implemented as a V-shaped scoring rule with parameters $r_0 = 1$ and $r_1 = r_{\hat{t},1}^G\in[0,1]$. Moreover, at any time $t\le\tau_R$,

- if $t>\hat{t}+\Delta$, in the continuation game $\mathcal{G}_{t,\tau_R}$, the agent's utility for not exerting effort is the same in both contract R and $\hat{R}$ because the agent with no information belief $\mu_{t-\Delta}^N$ will choose the same menu option $r_{\tau_R}^N$. However, the agent's utility for exerting effort is weakly higher in contract $\hat{R}$ since the reward $r_{t,1}^G$ from receiving a good news signal at time t weakly decreases in t.

- if $t\le\hat{t}+\Delta$, in the continuation game $\mathcal{G}_{t,\tau_R}$, by changing the contract from R to $\hat{R}$, the decrease in agent's utility for not exerting effort is exactly $\mu_{t-\Delta}^N(r_{t-\Delta,1}^G - r_{\hat{t},1}^G)$ by changing the menu option for no information belief $\mu_{t-\Delta}^N$ from $r_{t-\Delta}^G$ to $r_{\hat{t}}^G$. However, the decrease in the agent's utility for exerting effort in $\mathcal{G}_{t,\tau_R}$ is at most $\mu_{t-\Delta}^N(r_{t-\Delta,1}^G - r_{\hat{t},1}^G)$ since the decrease in reward for receiving a good news signal G is at most $r_{t-\Delta,1}^G - r_{\hat{t},1}^G$ and it only occurs when the state is 1.

Therefore, given contract $\hat{R}$, the agent has stronger incentives to exert effort in all continuation games $\mathcal{G}_{t,\tau_R}$ with $t\le\tau_R$, which implies that $\tau_{\hat{R}}\ge\tau_R$ and hence contract $\hat{R}$ is also effort-maximizing. □

Proof of Lemma 4. Note that it is easy to verify that if there exists a belief such that the agent is incentivized to exert effort, the intersection belief $\underline{\mu}(r_1)$ is such that the agent with belief $\underline{\mu}(r_1)$ would prefer menu option $(1,0)$ to $(0,r_1)$ and $\bar{\mu}(r_1)$ is such that the agent with belief $\bar{\mu}(r_1)$ would prefer menu option $(0,r_1)$ to $(1,0)$.

Consider the case of decreasing the reward parameter from $r_1 = z$ to $r_1 = z'$ for $0\le z' < z\le 1$. The agent's utility for not exerting effort given menu option $(1,0)$ remains unchanged, but the agent's utility for exerting effort in at least one period decreases. Therefore, $\underline{\mu}(r_1)$ weakly increases. Moreover, given posterior belief $\mu_{t-\Delta}^N$, the agent's utility for not exerting effort given menu option $(0,r_1)$ decreases by $\mu_{t-\Delta}^N(z-z')$, while the the agent's utility for exerting effort in at least one period decreases by at most $\mu^N_{t-\Delta}(z - z')$ since the reward decrease can only occur when the state is 1. Therefore, $\bar{\mu}(r_1)$ also weakly increases. □

Proof of Proposition 2. By Theorem 2, it is without loss to focus on contracts that can be implemented as V-shaped scoring with parameters $r_0 = 1$ and $r_1 \in [0, 1]$. If $D > \mu^*$ or $\mu < \underline{\mu}(1)$, given any $r_1 \in [0, 1]$, we have $D \notin [\underline{\mu}(r_1), \bar{\mu}(r_1)]$ and hence the agent cannot be incentivized to exert effort.

If $D \in [\underline{\mu}(1), \mu^*]$, if $\bar{\mu}_{r_T} \geq D$, the agent can be incentivized to exert effort from time 0 to time T given parameter $r_1 = r_T$, and hence choosing $r_1 = r_T$ must be part of the effort-maximizing mechanism. If $\bar{\mu}_{r_T} < D$, let r_1 be the maximum number such that $\bar{\mu}(r_1) \geq D$. By the monotonicity in Lemma 4, we have $r_1 \leq r_T$. In this case, the time horizon is not a binding constraint and the agent's optimal strategy is to stop before time T. Therefore, the agent's optimal utility from exerting effort in at least one period is the same with and without the time horizon constraint T. In this case, the agent has incentive to exert effort at any time $t \geq 0$ such that $\mu^N_t \geq \underline{\mu}(r_1)$. Moreover, this is part of the effort-maximizing mechanism, since r_1 is chosen to maximize $\underline{\mu}(r_1)$ subject to the effort constraint at time 0. □

B.3 Single-signal Environment

Proof of Lemma 5. For any contract R, let $\underline{u}_t(\mu)$ be the convex hull of the no information payoff $u^N_{t'}$ for $t' \leq t$ by viewing $u^N_{t'}$ as a function of $\mu^N_{t'}$. Consider an effort-maximizing contract R with the following selection:

1. maximizes the time $\bar{t}$ such that $\underline{u}_{\bar{t}}(\mu^N_t) = u^N_t$ for any time $t \leq \bar{t} - \Delta$;
2. conditional on maximizing $\bar{t}$, selecting the one that maximizes the weighted average no information payoff after time $\bar{t}$, i.e., $\sum_{i \geq 0} e^{-\frac{i}{\Delta}} \cdot r^N_{\bar{t}+i\Delta}$.

The existence of an effort-maximizing contract given such selection rule can be shown using standard arguments since, recalling that we have a discrete-time model, the set of effort-maximizing contracts that satisfy the first criterion is compact and the objective in the second selection criterion is continuous. Let τ_R be the stopping time of the agent for contract R. We will show that $\bar{t} = \tau_R$.

Suppose by contradiction we have $\bar{t} < \tau_R$. At any time $t \leq \tau_R$, recall that $\mathcal{G}_t$ is the continuation game at time t with prior belief $\mu^N_{t-\Delta}$ such that the agent's utility for not exerting effort is $u^N_{t-\Delta}$ and the agent's utility for exerting effort in $\mathcal{G}_t$ is U_t. Note that

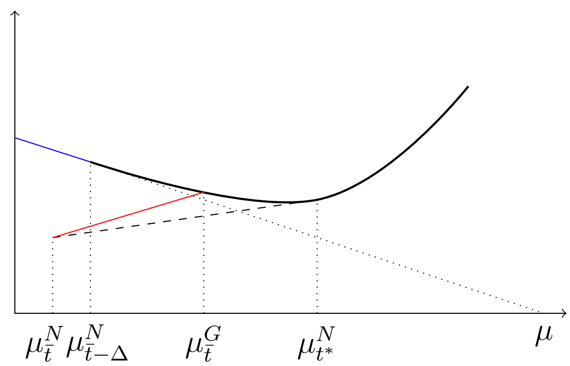

Figure 7: This figure illustrates the case when $\mu_{\bar{t}}^G \leq \mu_{t^*}^N$. The black curve is the function $\underline{u}_{\bar{t}}(\mu)$, blue line is the function $\bar{u}(\mu)$ and the red line is the function $y(\mu)$.

$U_t \geq u_{t-\Delta}^N$ for any $t \leq \tau_R$ for any $t \leq \tau_R$. We first show that the equality holds must hold at time $\bar{t}+\Delta$, i.e., $U_{\bar{t}+\Delta} = u_{\bar{t}}^N$.

Let $\bar{u}_t(\mu)$ be the upper bound on the expected reward at any belief μ at time t given that no Poisson signal has arrived before t. Specifically,

$$\bar{u}_t(\mu) = \max_{z_0,z_1\in[0,1]} \mu z_1 + (1-\mu) z_0 \quad \text{s.t.} \quad \mu_{t'}^N z_1 + (1-\mu_{t'}^N) z_0 \leq u_t^N, \forall t' \leq t.$$

It is easy to verify that function $\bar{u}_t(\mu)$ is convex in μ for all t and $\bar{u}_t(\mu_{t'}^N)$ is an upper bound on the no information utility $u_{t'}^N$ for all $t' > t$. Moreover, for any $\mu \leq \mu_t^N$, $\bar{u}_t(\mu)$ is a linear function in μ. The reward function $\bar{u}_t(\mu)$ for $t = \bar{t}$ and $\mu \leq \mu_{\bar{t}}^N$ is illustrated in Figure 7 as the blue straight line.

Since the no information utility is not convex at time $t = \bar{t}$, we have $\bar{u}_{\bar{t}-\Delta}(\mu_{\bar{t}}^N) > u_{\bar{t}}^N$. In this case, if $U_{\bar{t}+\Delta} > u_{\bar{t}}^N$, by increasing $u_{\bar{t}}^N$ to $\min\{U_{\bar{t}+\Delta}, \bar{u}_{\bar{t}-\Delta}(\mu_{\bar{t}}^N)\}$, the incentive of the agent for exerting effort is not violated. Moreover, selection rule (2) of maximizing the no information utility after time $\bar{t}$ is violated, a contradiction. Therefore, we can focus on the situation where the agent's incentive for exerting effort at time $\bar{t}+\Delta$ is binding.

By the construction of R, there exists $t \leq \bar{t}$ such that $\underline{u}_{\bar{t}}(\mu_t^N) < u_t^N$. Let t^* be the maximum time such that $\underline{u}_{\bar{t}}(\mu_{t^*}^N) = u_{t^*}^N$. That is, $\mu_{t^*}^N$ is the tangent point such that u_t^N coincides with it convex hull. See Figure 7 for an illustration. We consider two cases separately.

- $\mu_{\bar{t}}^* \geq \mu_{\bar{t}}^G$. In this case, let $y(\mu)$ be a linear function of posterior μ such that $y(\mu_{\bar{t}}^N) = u_{\bar{t}}^N$ and $y(\mu_{\bar{t}}^G) = \underline{u}(\mu_{\bar{t}}^G)$. Function y is illustrated in Figure 7 as the red line. Note that in this case, we have $y(\mu_{\bar{t}-\Delta}^N) < \underline{u}_{\bar{t}}(\mu_{\bar{t}-\Delta}^N) = u_{\bar{t}-\Delta}^N$. Moreover, $y(\mu_{\bar{t}-\Delta}^N)$ is the maximum continuation payoff of the agent for exerting effort at time $\bar{t}$ given belief

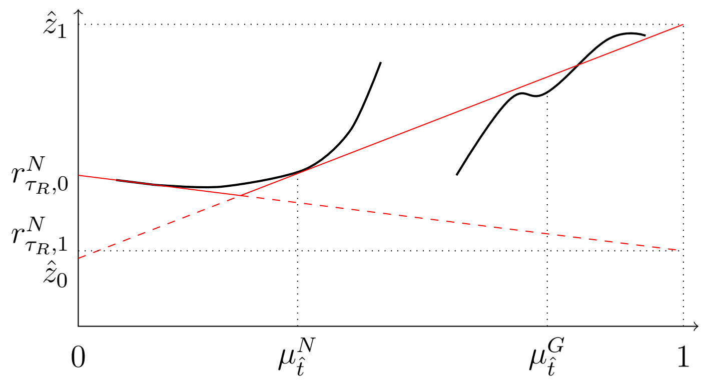

Figure 8: The black solid curves are the agent's expected utilities for not exerting effort as a function of his belief at any time t. The left curve is the expected utility for beliefs without receiving Poisson signals, and the right curve is the one receiving the Poisson signal G.

$\mu^N_{\bar{t}-\Delta}$. This is because by exerting effort, either the agent receives a Poisson signal G at time $\bar{t}$, which leads to posterior belief $\mu^G_{\bar{t}}$ with expected payoff $\underline{u}(\mu^G_{\bar{t}}) = y(\mu^G_{\bar{t}})$, or the agent does not receive a Poisson signal, which leads to belief drift to $\mu^N_{\bar{t}}$, with optimal continuation payoff being $U_{\bar{t}} = u^N_{\bar{t}} = y(\mu^N_{\bar{t}})$. However, $y(\mu^N_{\bar{t}-\Delta}) < u^N_{\bar{t}-\Delta}$ implies that the agent has a strict incentive to not exert effort at time $\bar{t}$, a contradiction.

- $\mu^*_{\bar{t}} < \mu^G_{\bar{t}}$. In this case, consider another contract $\bar{R}$ such that the no information utility in contract $\bar{R}$ is $u^N_{t;\bar{R}} = \underline{u}(\mu^N_t)$ for any $t \leq \bar{t}$. Note that in contract $\bar{R}$, the expected reward of the agent at any time t for receiving a Poisson signal is the same as in contract R, while the expected reward for not receiving Poisson signals weakly decreases. Therefore, contract $\bar{R}$ is also an effort-maximizing contract. However, the time such that the no information payoff is a convex function is strictly larger in $\bar{R}$, contradicting to our selection rule for R.

Therefore, we have $\bar{t} = \tau_R$ and the no information utility of the agent is a convex function. □

Proof of Theorem 3. By Lemma 5, there exists a contract R with a sequence of menu options $\{r^s_t\}_{s\in S,t\leq\tau_R} \cup \{r^N_{\tau_R}\}$ in which the no information payoff is convex in the no information belief. If $\mu^N_t < u^N_t$ for all $t \leq \tau_R$, let $\hat{z}_1 = 1$ and let $\hat{z}_0 \leq 1$ be the maximum reward such that $\mu^N_t + \hat{z}_0(1-\mu^N_t) \leq u^N_t$ for all $t \leq \tau_R$. Otherwise, let $\hat{z}_0 = 0$ and let $\hat{z}_1 \leq 1$ be the maximum reward such that $\hat{z}_1 \cdot \mu^N_t \leq u^N_t$ for all $t \leq \tau_R$. Essentially, the straight line $(\hat{z}_0, \hat{z}_1)$ is tangent with the agent's utility curve for not receiving informative signals. Let $\hat{t}$ be the time corresponds to the rightmost tangent point. See Figure 8 for an illustration.

Let $\underline{u}(\mu)$ be the function that coincides with u_t^N for $\mu \leq u_{\hat{t}}^N$ and $\underline{u}(\mu) = (\hat{z}_1 - \hat{z}_0)\mu + \hat{z}_0$. Note that $\underline{u}$ is convex. Consider another contract $\hat{R}$ that is implemented by scoring rule $P(\mu, \theta) = \underline{u}(\mu) + \xi(\mu)(\theta - \mu)$ for all $\mu \in [0, 1]$ and $\theta \in \{0, 1\}$ where $\xi(\mu)$ is a subgradient of $\underline{u}$. It is easy to verify that the implemented scoring rule satisfies the bounded constraint on rewards. Next we show that $\tau_{\hat{R}} \geq \tau_R$ and hence contract $\hat{R}$ must also be effort-maximizing, which concludes the proof of Theorem 3.

In any continuation game $\mathcal{G}_t$, recall that $u_{t-\Delta}^N$ is the utility of the agent for not exerting effort and U_t is the utility of the agent for exerting effort given contract R. For any time $t \leq \tau_R$, the agent has incentive to exert effort at time t given contract R implies that $u_{t-\Delta}^N \leq U_t$. Given contract $\hat{R}$, we similarly define $\hat{u}_{t-\Delta}^N$ and $\hat{U}_t$ and show that for any time $t \leq \tau_R$, $U_t - \hat{U}_t \leq u_{t-\Delta}^N - \hat{u}_{t-\Delta}^N$. This immediately implies that the agent also has incentive to exert effort at any time $t \leq \tau_R$ given contract $\hat{R}$ and hence $\tau_{\hat{R}} \geq \tau_R$.

Our analysis for showing that $U_t - \hat{U}_t \leq u_{t-\Delta}^N - \hat{u}_{t-\Delta}^N$ is divided into two cases.

Case 1: $t \geq \hat{t}$. In this case, since $u_{t-\Delta}^N \geq \hat{u}_{t-\Delta}^N$ for any $t \leq \tau_R$ by the construction of contract $\hat{R}$, it is sufficient to show that $\hat{U}_t \geq U_t$ for any $t \in [\hat{t}, \tau_R]$. We first show that for any $t \in [\hat{t}, \tau_R]$, if $\mu_t^G \leq \mu_t^N$, we must have $\underline{u}(\mu_t^G) \geq u_t^G$ in order to satisfy the dynamic incentive constraint in contract R. Next we focus on the case where $\mu_t^G > \mu_t^N$ and show that $\hat{u}_t^G = \mu_t^G \hat{z}_1 + (1 - \mu_t^G)\hat{z}_0 \geq u_t^G$. We prove this by contradiction. Suppose that $u_t^G > \mu_t^G \hat{z}_1 + (1 - \mu_t^G)\hat{z}_0$. Recall that $(r_{t,0}^G, r_{t,1}^G)$ are the options offered to the agent at time t that attains expected utility u_t^G under belief μ_t^G. Moreover, in our construction, either $\hat{z}_0 = 0$, or $\hat{z}_1 = 1$, or both equality holds. Therefore, the bounded constraints $r_{t,0}^G, r_{t,1}^G \in [0, 1]$ and the fact that agent with belief μ_t^G prefers $(r_{t,0}^G, r_{t,1}^G)$ over $(\hat{z}_0, \hat{z}_1)$ imply that

$$r_{t,0}^G \geq \hat{z}_0 \text{ and } r_{t,1}^G \geq \hat{z}_1.$$

See Figure 9 for an illustration. Since $\mu_{\hat{t}}^N < \mu_t^G$, this implies that the agent's utility at belief $\mu_{\hat{t}}^N$ given option $(r_{t,0}^G, r_{t,1}^G)$ is strictly larger than his utility under $(\hat{z}_0, \hat{z}_1)$, i.e.,

$$\mu_{\hat{t}}^N r_{t,1}^G + (1 - \mu_{\hat{t}}^N) r_{t,0}^G > \mu_{\hat{t}}^N \hat{z}_1 + (1 - \mu_{\hat{t}}^N)\hat{z}_0 = u_{\hat{t}}^N.$$

However, option $(r_{t,0}^G, r_{t,1}^G)$ is a feasible choice for the agent at time $\hat{t}$ in dynamic scoring rule S since $t \geq \hat{t}$, which implies that $\mu_{\hat{t}}^N r_{t,1}^G + (1 - \mu_{\hat{t}}^N) r_{t,0}^G \leq u_{\hat{t}}^N$. This leads to a contradiction.

Finally, for $t \in [\hat{t}, \tau_R]$, conditional on the event that the informative signal did not

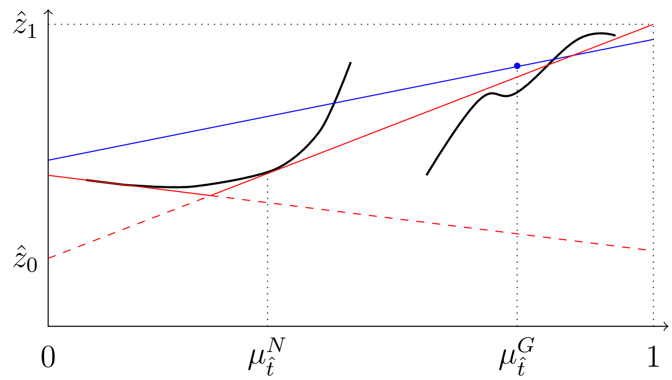

Figure 9: The blue line is the expected utility of the agent for choosing the menu option $(r^G_{t,0}, r^G_{t,1})$ if the utility at belief $\mu^G_{\hat{t}}$ is higher than the red line.

arrive at any time before t, since the agent expected utility given contract $\hat{R}$ is weakly higher compared to contract R given any arrival time of the Poisson signal, taking the expectation we have $\hat{U}_t \geq U_t$.

Case 2: $t < \hat{t}$. In this case, the continuation value for both stopping effort immediately and exerting effort until time τ_R weakly decreases. However, we will show that the expected decrease for stopping effort is weakly higher. For any $t < \hat{t}$, let $\hat{r}_t$ be the reward such that $u^N_t = \mu^N_t \hat{r}_t + (1-\mu^N_t)\hat{z}_0$. Note that $\hat{r}_t \geq \hat{z}_1$ and it is possible that $\hat{r}_t \geq 1$. The construction of quantity $\hat{r}_t$ is only used in the intermediate analysis, not in the constructed scoring rules. Let $\tilde{u}_t(\mu) \triangleq \mu\hat{r}_t + (1-\mu)\hat{z}_1$ be the expected utility of the agent for choosing option $(\hat{z}_0, \hat{r}_t)$ given belief μ. This is illustrated in Figure 10.

By construction, the expected utility decrease for not exerting effort in $\mathcal{G}_t$ is

$$u^N_t - \hat{u}^N_t = \tilde{u}_t(\mu^N_t) - \hat{u}^N_t = \mu^N_t(\hat{r}_t - \hat{z}_1).$$

Next observe that for any time $t' \in [t, \tau_R]$, $u^G_{t'} \leq \tilde{u}_t(\mu^G_{t'})$. This argument is identical to the proof in Case 1, and hence omitted here. Therefore, the expected utility decrease for exerting effort until τ_R is

$$\begin{aligned} U_t - \hat{U}_t &= \int_{t+\Delta}^{\tau_R} \left(u^G_{t'} - \hat{u}^G_{t'}\right) \mathrm{d}F_t(t') \leq \int_{t+\Delta}^{\tau_R} \left(\tilde{u}_t(\mu^G_{t'}) - \hat{u}^G_{t'}\right) \mathrm{d}F_t(t') \\ &= (\hat{r}_t - \hat{z}_1) \cdot \int_{t+\Delta}^{\tau_R} \mu^G_{t'} \,\mathrm{d}F_t(t') \leq (\hat{r}_t - \hat{z}_1) \cdot \mu^N_t \leq (\hat{r}_{t-\Delta} - \hat{z}_1) \cdot \mu^N_{t-\Delta} \end{aligned}$$

where the second inequality holds by Bayesian plausibility and the last inequal-

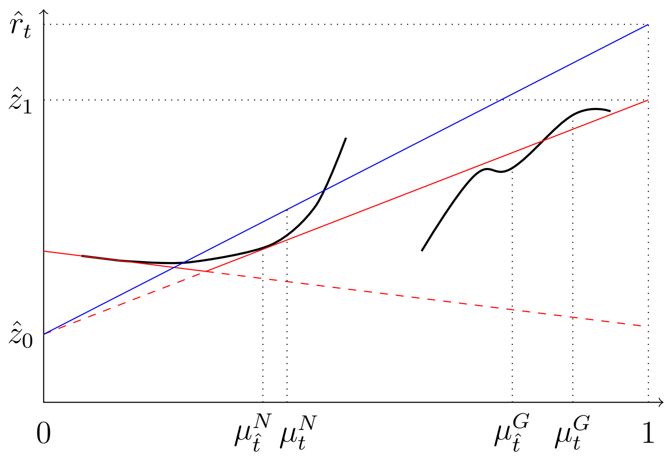

Figure 10: The blue line is the utility function $\tilde{u}_t$, which serves as an upper bound on the utility $u^G_{t'}$ for any $t' \geq t$.

ity holds since the no information belief drifts towards state 0. Combining the inequalities, we have $U_t - \hat{U}_t \leq u^N_{t-\Delta} - \hat{u}^N_{t-\Delta}$.

Combining the above two cases, we have $U_t - \hat{U}_t \leq u^N_{t-\Delta} - \hat{u}^N_{t-\Delta}$ for any $t \leq \tau_R$. Since the agent's optimal effort strategy is to stop at time τ_R given contract R, this implies that at any time $t \leq \tau_R$, if the agent has not received any informative signal by time t, the agent also has incentive to exert effort until time τ_R given contract $\hat{R}$ that can be implemented as a scoring rule. □

Online Appendix for "Implementing Evidence Acquisition: Time Dependence in Contracts for Advice"

Yingkai Li and Jonathan Libgober

OA 1 General Environments

OA 1.1 Results for Dynamic Contracts

Proof of Lemma 6. If $T \leq T_{\lambda,D,c}$, the lemma holds trivially. Next we focus on the case $T > T_{\lambda,D,c}$.

Suppose there exists a contract R such that $\tau_R > T_{\lambda,D,c}$. The prior belief in the continuation game $\mathcal{G}_{\tau_R}$ is $\mu^N_{\tau_R-\Delta} < \mu_{\lambda,c} \leq \frac{1}{2}$. By the definition of τ_R, the agent's optimal strategy is to exert effort for one period given contract R. This implies that the agent has incentive to exert effort in continuation game $\mathcal{G}_{\tau_R}$ given the effort-maximizing scoring rule for $\mathcal{G}_{\tau_R}$. By Proposition 1, the effort-maximizing scoring rule for $\mathcal{G}_{\tau_R}$ is the V-shaped scoring rule P with kink at $\mu^N_{\tau_R-\Delta}$. By simple algebraic calculation, the expected utility increase given scoring rule P for exerting effort in $\mathcal{G}_{\tau_R}$ is $\mu^N_{\tau_R-\Delta}(\lambda^G_1 - \lambda^G_0)\Delta$, which must be at least the cost of effort $c\Delta$. However, this violates the assumption that $\mu^N_{\tau_R-\Delta} < \mu_{\lambda,c}$, a contradiction. □

Proof of Lemma 7. For any time $t \geq 0$, given any information arrival probabilities λ such that $\lambda^G_1 + \lambda^B_1 \leq \lambda^G_0 + \lambda^B_0 + \epsilon$, we have

$$\begin{aligned}
\mu^N_{t-\Delta} - \mu^N_t &= \mu^N_{t-\Delta} - \frac{\mu^N_{t-\Delta}(1-\lambda^G_1\Delta - \lambda^B_1\Delta)}{\mu^N_{t-\Delta}(1-\lambda^G_1\Delta - \lambda^B_1\Delta) + (1-\mu^N_{t-\Delta})(1-\lambda^G_0\Delta - \lambda^B_0\Delta)} \\
&\leq \mu^N_{t-\Delta}\left(1 - \frac{(1-\lambda^G_1\Delta - \lambda^B_1\Delta)}{(1-\lambda^G_1\Delta - \lambda^B_1\Delta) + (1-\mu^N_{t-\Delta})\epsilon\Delta}\right) \\
&\leq 2\mu^N_{t-\Delta}(1-\mu^N_{t-\Delta})\epsilon\Delta \leq \frac{1}{2}\epsilon\Delta. \qquad (2)
\end{aligned}$$

the second inequality holds since $\lambda^G_1\Delta + \lambda^B_1\Delta \leq \frac{1}{2}$ and the last inequality holds since $\mu^N_{t-\Delta}(1-\mu^N_{t-\Delta}) \leq \frac{1}{4}$. For any $\eta > 0$, there exists ϵ_0 such that

For any $\eta > 0$, let $\epsilon = \frac{2\eta\kappa_0}{D} > 0$. Given the myopic-incentive contract R, the agent's utility increase for exerting effort in one period at time t is

$$\mu^N_t\lambda^G_1\Delta + (1-\mu^N_t)(1-\lambda^B_0\Delta)\cdot\frac{\mu^N_t}{1-\mu^N_t} - \mu^N_{t-\Delta} = \mu^N_t(\lambda^G_1 - \lambda^B_0)\Delta + (\mu^N_t - \mu^N_{t-\Delta}).$$

If $\mu_t^N \geq \mu_{T_{\lambda,D,c}}^N + \eta$, we have $\mu_t^N \geq \mu_{\lambda,c} + \eta$ and hence the expected utility increase is at least

$$\mu_{\lambda,c}(\lambda_1^G - \lambda_0^B)\Delta + \eta(\lambda_1^G - \lambda_0^B)\Delta + (\mu_t^N - \mu_{t-\Delta}^N) \geq \mu_{\lambda,c}(\lambda_1^G - \lambda_0^B)\Delta$$

where the inequality holds by the definition of ϵ and the sufficient incentive condition. Note that this is at least the cost of effort $c\Delta$ by the definition of $\mu_{\lambda,c}$, and hence the agent has incentive to exert effort at time t. Therefore, the stopping time given the myopic-incentive contract satisfies $\mu_{\tau_R}^N \leq \mu_{T_{\lambda,D,c}}^N + \eta$. □

To prove Theorem 4, we also utilize the following lemma to bound the difference in expected scores when the posterior beliefs differ by a small constant of ϵ given any bounded scoring rule.

Lemma 1. *For any bounded static scoring rule P with expected reward function $U_P(\mu)$ given posterior belief μ, we have*

$$|U_P(\mu + \epsilon) - U_P(\mu)| \leq \epsilon, \qquad \forall \epsilon > 0, \mu \in [0, 1 - \epsilon].$$

Proof. For any static scoring rule P, the subgradient of U_P evaluated at belief μ equals its difference in rewards between realized states 0 and 1, which is bounded between $[-1, 1]$ since the scoring rule is bounded within $[0, 1]$. This further implies that $|U_P(\mu + \epsilon) - U_P(\mu)| \leq \epsilon$ for any $\epsilon > 0$ and $\mu \in [0, 1 - \epsilon]$. □

Proof of Theorem 4. By Lemma 7, it is sufficient to show that there exists $\eta > 0$ and $\epsilon > 0$ such that when the slow-drift condition is satisfied for constant ϵ, for any contract R that can be implemented as a scoring rule, we have $\mu_{\tau_R}^N - \mu_{T_{\lambda,D,c}}^N > \eta$.

Suppose by contradiction there exists a contract R that can be implemented as a scoring rule and $\mu_{\tau_R}^N - \mu_{T_{\lambda,D,c}}^N \leq \eta$. Let P be the scoring rule that implements contract R and let $U_P(\mu) = \mathbf{E}_{\theta \sim \mu}[P(\mu), \theta]$ be the expected score of the agent. Let $\underline{U}(\mu)$ be a linear function such that $\underline{U}(\mu_{\tau_R}^B) = U_P(\mu_{\tau_R}^B)$ and $\underline{U}(\mu_{\tau_R}^N) = U_P(\mu_{\tau_R}^N)$. Let $\overline{U}(\mu)$ be a linear function such that $\overline{U}(\mu_{\tau_R}^G) = U_P(\mu_{\tau_R}^G)$ and $\overline{U}(\mu_{\tau_R}^N) = U_P(\mu_{\tau_R}^N)$. See Figure 11 for an illustration. Let $f_t^s \triangleq \mu_t^N \lambda_1^G \Delta + (1 - \mu_t^N)\lambda_0^G$. At time τ_R, the agent has incentive to exert effort, which implies that the cost of effort $c\Delta$ is at most the utility increase for exerting effort

$$\begin{aligned} &f_{\tau_R-\Delta}^G \Delta \cdot U_P(\mu_{\tau_R}^G) + f_{\tau_R-\Delta}^B \Delta \cdot U_P(\mu_{\tau_R}^B) + (1 - f_{\tau_R-\Delta}^G \Delta - f_{\tau_R-\Delta}^B \Delta) \cdot U_P(\mu_{\tau_R}^N) - U_P(\mu_{\tau_R-\Delta}^N) \\ &= f_{\tau_R-\Delta}^G \Delta \cdot (U_P(\mu_{\tau_R}^G) - \underline{U}(\mu_{\tau_R}^G)) + \underline{U}(\mu_{\tau_R-\Delta}^G) - U_P(\mu_{\tau_R-\Delta}^N) \leq f_{\tau_R-\Delta}^G \Delta \cdot (U_P(\mu_{\tau_R}^G) - \underline{U}(\mu_{\tau_R}^G)) \end{aligned}$$

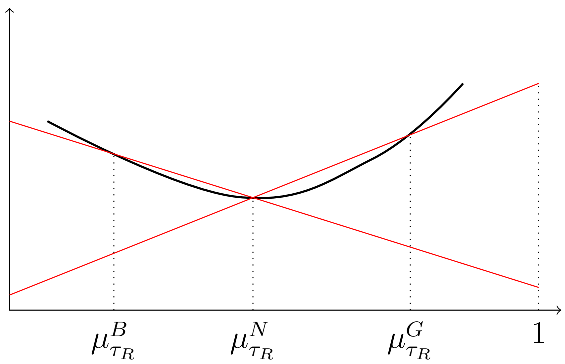

Figure 11: The black curve is the expected score function U_P. The red lines are linear functions $\underline{U}$ and $\overline{U}$ respectively.

where the equality holds by linearity of expectation and the inequality holds by the convexity of utility function U_P. Therefore, we have

$$U_P(\mu^G_{\tau_R}) - \underline{U}(\mu^G_{\tau_R}) \geq \frac{c}{f^G_{\tau_R-\Delta}} = \frac{c}{\mu^N_{\tau_R-\Delta}\lambda^G_1 + (1-\mu^N_{\tau_R-\Delta})\lambda^G_0}$$
$$= \frac{(\lambda^G_1 - \lambda^G_0)(1 - \frac{\eta}{\mu^N_{\tau_R-\Delta}})}{\lambda^G_1 + \frac{1}{\mu^N_{\tau_R-\Delta}}(1-\mu^N_{\tau_R-\Delta})\lambda^G_0} \geq \frac{(\lambda^G_1 - \lambda^G_0)}{\lambda^G_1 + \frac{1}{\mu^N_{\tau_R-\Delta}}(1-\mu^N_{\tau_R-\Delta})\lambda^G_0} - \frac{2\eta}{\underline{\kappa}_1}$$

where the second inequality holds since $\mu^N_{\tau_R} \leq \mu^N_{T_{\lambda,D,c}} + \eta \leq \mu_{\lambda,c} + \eta$, and the last inequality holds since $\lambda^B_0 \geq \underline{\kappa}_1$.

For any constant $\gamma > 0$, let time t be the time such that $\mu^N_{t-\Delta} - \mu^N_{\tau_R-\Delta} > \gamma$. First note that the convexity of U_P and the constraint on rewards belonging to the unit interval at state 1 implies that

$$\overline{U}(\mu^N_{t-\Delta}) - U_P(\mu^N_{t-\Delta}) \leq \frac{2\eta}{\underline{\kappa}_1} \cdot \frac{1-\mu^N_{\tau_R}}{1-\mu^G_{\tau_R}} \cdot \frac{\mu^G_{\tau_R} - \mu^N_{t-\Delta}}{\mu^G_{\tau_R} - \mu^N_{\tau_R}} \leq \frac{2\eta(\underline{\kappa}_1 + \bar{\kappa}_1)}{\underline{\kappa}_1^2}$$

where the last inequality holds since $(1-\mu^N_{\tau_R}) \cdot \frac{\mu^G_{\tau_R} - \mu^N_{t-\Delta}}{\mu^G_{\tau_R} - \mu^N_{\tau_R}} \leq 1$ and $\frac{1}{1-\mu^G_{\tau_R}} \leq \frac{\underline{\kappa}_1 + \bar{\kappa}_1}{\underline{\kappa}_1}$. Moreover,

$$U_P(\mu^G_t) - \overline{U}(\mu^G_t) \leq \frac{2\eta}{\underline{\kappa}_1} \cdot \frac{1-\mu^N_{\tau_R}}{\mu^G_{\tau_R} - \mu^N_{\tau_R}} \leq \frac{2\eta}{\underline{\kappa}_1} \cdot \frac{\lambda^G_1}{\mu^N_{\tau_R}(\lambda^G_1 - \lambda^G_0)} \leq \frac{2\eta\bar{\kappa}_1}{\underline{\kappa}_1 c}.$$

Therefore, the utility increase for exerting effort in one period at time t is

$$
\begin{aligned}
& f^G_{t-\Delta}\Delta \cdot U_P(\mu^G_t) + f^B_{t-\Delta}\Delta \cdot U_P(\mu^B_t) + (1 - f^G_{t-\Delta}\Delta - f^B_{t-\Delta}\Delta) \cdot U_P(\mu^N_t) - U_P(\mu^N_{t-\Delta}) \\
& \leq \left(f^G_{t-\Delta} \cdot \overline{U}(\mu^G_t) + f^B_{t-\Delta} \cdot \underline{U}(\mu^B_t) - (f^G_{t-\Delta} + f^B_{t-\Delta}) \cdot \overline{U}(\mu^N_t) + \epsilon + \frac{2\eta\bar{\kappa}_1}{\underline{\kappa}_1 c} + \frac{2\eta(\underline{\kappa}_1 + \bar{\kappa}_1)}{\underline{\kappa}_1^2} \right) \Delta \\
& \leq \left(\mu^N_{\tau_R}(\lambda^B_0 - \lambda^B_1) - \gamma\underline{\kappa}_1 + \epsilon + \frac{2\eta\bar{\kappa}_1}{\underline{\kappa}_1 c} + \frac{2\eta(\underline{\kappa}_1 + \bar{\kappa}_1)}{\underline{\kappa}_1^2} \right) \Delta.
\end{aligned}
$$

Since $c = \mu_{\lambda,c}(\lambda^G_1 - \lambda^G_0) \geq (\mu^N_{\tau_R} - \eta)(\lambda^B_0 - \lambda^B_1)$, the agent suffer from a loss at least

$$
\left(\gamma\underline{\kappa}_1 - \eta\bar{\kappa}_1 - \epsilon - \frac{2\eta\bar{\kappa}_1}{\underline{\kappa}_1 c} - \frac{2\eta(\underline{\kappa}_1 + \bar{\kappa}_1)}{\underline{\kappa}_1^2} \right) \Delta
$$

for exerting effort in one period.

Now consider the utility increase for exerting effort from belief μ^N_t to $\mu^N_{\tau_R}$. Note that for any $\delta > 0$, with probability at least

$$
1 - (1 - f^G_{t-\Delta}\Delta - f^B_{t-\Delta}\Delta)^{\frac{\delta}{\epsilon\Delta}} \leq 1 - \exp\left(-\frac{\delta}{\epsilon}(f^G_{t-\Delta} + f^B_{t-\Delta}) \right),
$$

the agent receives a Poisson signal and stops before the no information belief drifts for a δ distance. Moreover, the loss is at least

$$
\left(\gamma\underline{\kappa}_1 - 2\delta - \eta\bar{\kappa}_1 - \epsilon - \frac{2\eta\bar{\kappa}_1}{\underline{\kappa}_1 c} - \frac{2\eta(\underline{\kappa}_1 + \bar{\kappa}_1)}{\underline{\kappa}_1^2} \right) \Delta
$$

in each period before the no information belief drifts a δ distance. In contrast, the benefit from exerting effort after the no information belief drifts a δ distance is at most 1, but it only occurs with probability at most $\exp\left(-\frac{\delta}{\epsilon}(f^G_{t-\Delta} + f^B_{t-\Delta})\right)$. Therefore, the agent's utility for exerting effort is smaller than not exerting effort in continuation game $\mathcal{G}_t$ when parameters δ, η, ϵ are chosen to be sufficiently small compared to γ. This leads to a contradiction since the agent at time t will not choose to exert effort given scoring rule P. □

OA 1.2 Effort-Maximizing Dynamic Contracts

Proof of Theorem 5. By Lemma 2, the effort-maximizing contract can be represented as offering a sequence of menu options $\{r^s_t\}_{t \leq \tau_R, s \in S} \cup \{r^N_{\tau_R}\}$ that satisfies the incentive constraints.

Recall that r^N_t is the menu option chosen by agent with belief μ^N_t given set of available menu options $\mathcal{R}_t$ at time t. For any time $t \leq \tau_R$, let

$$\hat{r}_t^N = \arg\max_{(r_0,r_1)\in[0,1]^2} \quad r_0$$
$$\text{s.t.} \qquad \mu_t^N r_1 + (1-\mu_t^N) r_0 = \mu_t^N r_1^N + (1-\mu_t^N) r_0^N.$$

That is, the agent with belief μ_t^N has the same expected reward given both menu options r_t^N and $\hat{r}_t^N$, and $\hat{r}_t^N$ maximizes the reward for state 0. Note that $\hat{r}_t^N$ is also the menu option that maximizes the agent's utility with belief μ_t^B without violating the incentive constraints. Therefore, it suffices to consider contracts where $\hat{r}_t^B = \hat{r}_t^N$ for all $t \leq \tau_R$.

Moreover, by the incentive constraints over time, the menu options $\hat{r}_t^B = \hat{r}_t^N$ are decreasing over time, and the decrease happens first for state 1 since $\hat{r}_t^N$ maximizes the rewards for state 0, which implies that conditions (1) in Theorem 5 holds. Finally, the menu option for belief μ_t^G can be computed by maximizing the agent's utility without violating the incentive constraints from previous time, i.e.,

$$r_t^G = \arg\max_{r:\Theta\to[0,1]} u(\mu_t^G, r)$$
$$\text{s.t.} \quad u(\mu_{t'}^N, r_{t'}^N) \geq u(\mu_{t'}^N, r), \quad \forall t' \in [0,t].$$

This is because such menu option maximizes the agent's continuation utility for exerting effort at any time $t \leq \tau_R$ without affecting the agent's utility for stopping immediately. Therefore, condition (2) in Theorem 5 is satisfied as well. $\square$

OA 2 Efficient Algorithms for Computing Effort-Maximizing Contracts

Let $\tau \leq T$ be the stopping time in the effort-maximizing contract. We know that there exists a contract R such that the agent has incentive to exert effort in all continuation $\mathcal{G}_{t,\tau}$ with $t \leq \tau$. Therefore, to compute the effort-maximizing contract, it is sufficient to enumerate all stopping time τ and verify if there exists such a contract for incentivizing the agent to work until τ. The effort-maximizing contract is the contract for the largest τ.

To show this, fixing a stopping time τ, we consider the menu representation in Lemma 2, i.e, a sequence of menu options $\{r_t^s\}_{t\leq\tau,s\in S} \cup \{r_\tau^N\}$. Moreover, we construct additional variables $\{r_t^N\}_{t<\tau}$ to simplify the constraints we need in the optimization problem. For any $t \leq t' \leq \tau$ and $s \in S$, let $f_t^s(t')$ be the probability of receiving Poisson signal s at time t' conditional on not receiving Poisson signals before time t. Let $F_t^s(t')$ be the corresponding cumulative probability. The constraints for providing sufficient incentives for the agent to exert effort is

$$\left(1 - \sum_{s\in S} F_t^s(\tau)\right) \cdot u(\mu_\tau^N, r_\tau^N) + \sum_{t'=t}^{\tau}\sum_{s\in S} u(\mu_{t'}^s, r_{t'}^s) \cdot f_t^s(t') \geq u(\mu_{t-\Delta}^N, r_{t-\Delta}^N)$$

for all $t \leq \tau$ and $s \in S$. Moreover, the contract R is incentive compatible if

$$u(\mu_t^s, r_t^s) \geq u(\mu_t^s, r_{t'}^{s'})$$

for all $t \leq t' \leq \tau$ and $s, s' \in S \cup \{N\}$. Note that since the utility function $u(\mu, \cdot)$ is a linear function, in the discrete time model, the above set of constraints is a finite set of linear constraints on the menu options. This implies that whether a solution exists and finding a solution if it exists can be computed in polynomial time, and hence the effort-maximizing contract can also be found in polynomial time.

OA 3 Additional Review of Scoring Rules

A scoring rule is *proper* if it incentive the agent to truthfully report his belief to the mechanism, i.e.,

$$\mathbf{E}_{\theta\sim\mu}[P(\mu, \theta)] \geq \mathbf{E}_{\theta\sim\mu}\left[P(\mu', \theta)\right], \quad \forall \mu, \mu' \in \Delta(\Theta).$$

By revelation principle, it is without loss to focus on proper scoring rules when the designer adopts contracts that can be implemented as a scoring rule.

Lemma 2 (McCarthy, 1956)**.** *For any finite state space Θ, a scoring rule P is proper if there exists a convex function $U_P : \Delta(\Theta) \to \mathbb{R}$ such that*

$$P(\mu, \theta) = U_P(\mu) + \xi(\mu) \cdot (\theta - \mu)$$

for any $\mu \in \Delta(\Theta)$ and $\theta \in \Theta$ where $\xi(\mu)$ is a subgradient of U_P.[15]

[15]Here for finite state space Θ, we represent $\theta \in \Theta$ and $\mu \in \Delta(\Theta)$ as $|\Theta|$-dimensional vectors where the ith coordinate of θ is 1 if the state is the ith element in Θ and is 0 otherwise, and the ith coordinate of posterior μ is the probability of the ith element in Θ given posterior μ.

OA 4 Comparative Statics of Scoring Rules

As discussed in Section 4, the ideal situation cannot be implemented in dynamic contracts since we claim that the effort-maximizing static scoring rule at time τ is not effort-maximizing at time $t < \tau$. However, this argument alone is insufficient since in earlier time, the agent is more uncertain about the states, and hence is easier to be incentivized. In this appendix, we formalize this intuition using a comparative statics on static scoring rules.

To simplify the exposition, we consider a specific static environment where if the agent exerts effort, the agent may receive an informative signal in $\{G, B\}$ that is partially informative about the state. Otherwise, the agent does not receive any signals and the prior belief is not updated. Let $f_{\theta,s} \in (0, 1)$ be the probability of receiving signal s conditional on state θ. That is, signals are not perfectly revealing. We focus on the case when the prior $D < \frac{1}{2}$.

Proposition 1 shows that the utility function of the effort-maximizing scoring rule is V-shaped with a kink at the prior, that is, the effort-maximizing scoring rule offers the agent the following two options: $(0, 1)$ and $(\frac{D}{1-D}, 0)$. The agent with prior belief D is indifferent between these two options. Moreover, any belief $\mu > D$ would strictly prefer $(0, 1)$ and any belief $\mu < D$ would strictly prefer $(\frac{D}{1-D}, 0)$.

Next we conduct comparative statics. The expected score increase for exerting effort under the effort-maximizing scoring rule is

$$\mathrm{Inc}(D) \triangleq (1 - D) \cdot f_{0,B} \cdot \frac{D}{1 - D} + D \cdot f_{1,G} - D = D(f_{0,B} + f_{1,G} - 1).$$

Since $f_{0,B} > f_{1,B}$ and $f_{1,G} > f_{1,B}$, we have $f_{0,B} + f_{1,G} > 1$. Therefore, the expected score increase is monotone increasing in prior D. That is, the closer the prior is to $\frac{1}{2}$, the easier to incentivize the agent to exert effort.

Next we conduct comparative statics on prior D' by fixing the scoring rule to P be effort-maximizing for D, i.e., P is the V-shaped scoring rule with kink at D. The expected score increase for exerting effort given scoring rule P is

$$\begin{aligned} \mathrm{Inc}(D'; D) &\triangleq (1 - D') \cdot f_{0,B} \cdot \frac{D}{1 - D} + D' \cdot f_{1,G} - D' \\ &= D'(f_{1,G} - 1 - f_{0,B} \cdot \frac{D}{1 - D}) + f_{0,B} \cdot \frac{D}{1 - D}. \end{aligned}$$

Since $f_{1,G} < 1$, the strength of the incentives the designer can provide is strictly decreasing in prior D'. Therefore, even though when prior is closer to $\frac{1}{2}$, it is easier to incentivize the agent to exert effort, the effort-maximizing scoring rule for lower priors may not be sufficient to incentivize the agent (by assuming that the cost of effort is the same in both settings).

OA 5 Ex Ante Reward Constraints

Proposition 1. *In environments where both perfect-learning and single-signal conditions are satisfied, and the designer's reward constraint is imposed in ex ante, there exists an effort-maximizing contract that can be implemented as a V-shaped scoring rule.*

Proof. Note that even with ex ante reward constraints, the menu representation of the effort-maximizing contracts still applies. Since only good news signal G arrives with positive probability, it is sufficient to consider effort-maximizing contract R with menu representation $\{r^G_t\}_{t\le\tau_R} \cup \{r^N_{\tau_R}\}$.

First note that it is without loss to assume that $r^N_{\tau_R,1} = 0$ and $r^G_{t,0} = 0$ for all $t \le \tau_R$. The latter holds because signals are perfectly revealing and hence lowering the reward for state 0 upon receiving good news signals does not affect the agent's incentive or the ex ante reward constraint of the designer. The reason why $r^N_{\tau_R,1} = 0$ is because we can always increase the reward $r^N_{\tau_R,0}$ and decrease the reward $r^N_{\tau_R,1}$ such that the agent's expected reward given belief $\mu^N_{\tau_R}$ remains unchanged, and the dynamic incentives of the agent is not affected under single-signal environments. In this case, the dynamic incentives of the contract implies that $r^G_{t,1}$ is weakly decreasing over time.

Now consider another contract $\hat{R}$ that only offers menu options $r^N_{\tau_R}$ and $\hat{r} = (0, \hat{r}_1)$ at all time t with parameter $\hat{r}_1 \le r^G_{\tau_R,1}$ chosen such that the agent's incentive for exerting effort is binding at time τ_R. Note that contract $\hat{R}$ can be implemented as a V-shaped scoring rule. We will show that $\tau_{\hat{R}} = \tau_R$ and the expected reward of the agent given contract $\hat{R}$ is lower, which implies that contract $\hat{R}$ is also effort-maximizing under the ex ante reward constraint.

Let $\hat{t}$ be the maximum time such that $\mu^N_{\hat{t}-\Delta}$ weakly prefers menu option $\hat{r}$ over $r^N_{\tau_R}$. Note that given contract $\hat{R}$, by the Envelope Theorem, the agent's utility for exerting effort optimally is convex in his current belief with derivative larger than $-r^N_{\tau_R,0}$. Therefore, for any time $t \in (\hat{t}, \tau_R]$, the agent always has incentive to exert effort given contract $\hat{R}$. For any time $t \le \hat{t}$, using the identical argument in the proof of Theorem 2 for the perfect-learning environments, by decreasing the rewards in menu options r^G_t, the agent's decrease in utility for exerting effort in continuation game $\mathcal{G}_t$ is less than his decrease in utility for not exerting effort. Therefore, the agent also has incentive to exert effort at any time $t \le \hat{t}$ in the dynamic model given contract $\hat{R}$. This implies that $\tau_{\hat{R}} \ge \tau_R$. Since the menu option $\hat{r}$ is chosen such that the agent's incentive for exerting effort is binding at time τ_R, the agent does not have incentive to exert effort at time $\tau_R + \Delta$ and hence $\tau_{\hat{R}} = \tau_R$. Finally, since $\hat{r}_1 \le r^G_{t,1}$ for any $t \le \tau_R$, the agent's expected reward for receiving good news signals at any time t is weakly lower, and hence the ex ante reward constraint is satisfied given contract $\hat{R}$. □