Incentivizing Forecasters to Learn:

Summarized vs. Unrestricted Advice*

Yingkai Li† Jonathan Libgober‡

Abstract

How should forecasters be incentivized to acquire the most information when learning takes place over time? We address this question in the context of a novel dynamic mechanism design problem in which a designer incentivizes an expert to learn by conditioning rewards on an event's outcome and the expert's reports. Eliciting summarized advice at a terminal date maximizes information acquisition if an informative signal either fully reveals the outcome or has predictable content. Otherwise, richer reporting capabilities may be required. Our findings shed light on incentive design for consultation and forecasting by illustrating how learning dynamics shape the qualitative properties of effort-maximizing contracts.

Keywords— Scoring rules, dynamic contracts, dynamic moral hazard, Poisson learning.
JEL— D82, D83

*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, Alex Bloedel, Tilman Börgers, Juan Carrillo, Tan Gan, Yingni Guo, Marina Halac, Jason Hartline, David Kempe, Nicolas Lambert, Elliot Lipnowski, Andrew McClellan, 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 the Sloan Research Fellowship FG-2019-12378 and the 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

1.1 Overview

This paper asks whether adding complexity to a contract by making rewards sensitive to the timing of reports can incentivize more information acquisition than simpler, time-independent contracts. We ask this question in a minimal model where only a single prediction is sought and there is no exogenous time pressure. We show that the answer depends on whether incentivizing effort at one belief necessarily weakens incentives at another.

Specifically, we study a novel mechanism design problem that extends past work on *scoring rule design*[1] to a setting featuring dynamic acquisition of costly information. A scoring rule is a contract that provides a reward—over which the agent has fixed value—with some probability depending on (i) the agent's (reported) belief about the likelihood of an event and (ii) that event's outcome. While much past work studying scoring rules is concerned with incentivizing forecasters to *truthfully report* information, we share a focus with a smaller line of work on how to achieve this end *as well as* incentivize its acquisition.

A leading application of our work—and scoring rule design more generally—is to the practical question of how to best incentivize forecasters. Forecasting takes a plethora of forms, ranging from familiar (e.g., weather or economic outlooks) to predictions more broadly (e.g., national security threat assessments or medical diagnoses). Our interest is in situations where the event over which a forecast is sought is idiosyncratic. Thus, the only way to obtain advice is via asking a forecaster to actively gather information, while the forecaster cannot rely on passive knowledge to make an informed prediction.

The motivation for considering dynamic information acquisition together with forecaster incentive design stems from the observation that active learning takes time. Consider, for example, a decision maker hiring an analyst to conduct due diligence; e.g., before acquiring a company or funding a startup. In this context, the analyst's task is to forecast whether the investment target will meet its promised outcomes, making this determination by analyzing its complex financial records. In applications where this information is extensive and difficult to interpret, it would be up to the analyst to decide *when* to stop looking for a red flag. This timing decision makes the forecaster's problem inherently dynamic.

We focus on the question of whether eliciting a single prediction, made after all information is acquired, induces truthful reporting *and* maximizes incentives for information acquisition. Our results clarify when more complex contracting capabilities are more conducive to informed forecasts than simpler ones. Some relationships—such as when a consultant is

[1] This literature, starting with Brier (1950), considered how to evaluate (i.e., provide scores for) weather forecasts. We review this work in more depth in Section 1.4.

hired at arm's length to report at a prespecified date—may limit the scope for reports to be provided at any time. To the extent that closer relationships enable continuous reporting, our main results address whether this capability provides an avenue for such relationships to yield benefits.

1.2 Model Summary

In our model, an agent (forecaster) chooses privately—over time—whether to exert costly effort to learn about an uncertain future binary event. We refer to the outcome of this event as the *state*, and assume the initial prior over it is commonly known. We assume that the private information of the agent, when exerting effort, arrives through a discretized Poisson bandit process. This Poisson learning technology allows us to tractably capture key qualitative features of dynamic information acquisition, enabling sharp comparisons across contract designs despite the complexities stemming from the combination of dynamic learning with a rich contracting space.[2]

When exerting effort, the agent may (privately) observe two kinds of signals, each (in our main model) with fixed probability: (a) a "null signal" or (b) a "Poisson signal." We interpret a Poisson signal as some particular sought-after information—e.g., a red flag in the due diligence application or an indication of an adversary's capabilities in a national-security context. In contrast to some related work, a Poisson signal in our model need *not* reveal the outcome or suggest only one outcome. A red flag might narrow the set of failure scenarios without confirming whether any will materialize; likewise, an intelligence analyst might learn that an adversary has the capability to act on a given date without learning their intentions. This structure is relevant to many expert-forecasting problems.

Past work on forecaster incentives has emphasized that a primary motivator is reputation (Marinovic et al., 2013). If the forecaster is employed by a decision maker, the latter's recommendation has a fixed value insofar as it shapes external perceptions. The aforementioned literature on scoring rule design used this observation to note that the truthful revelation of information could be incentivized by conditioning endorsements on predictions together with realized outcomes. But such schemes can not only incentivize forecasters to truthfully reveal their private information, but also to exert unobservable effort.[3] In our model, the designer's contract specifies how to provide such an endorsement, modeled as a

[2]In particular, the agent's continuation problem typically fails to be stationary for general contracts. Much existing work on costly information acquisition relies on stationarity to obtain tractability, requiring us to develop new techniques to analyze contract design in our setting.

[3]See Zellner et al. (2021) for a recent survey on forecasting, including a discussion of the relevance of scoring rules for forecaster incentives. Gneiting and Raftery (2007) provides theoretical background on scoring rules, while Frongillo and Waggoner (2024) provides a more recent survey.

reward of fixed value (i.e., a score). The objective of the designer is to incentivize the agent to acquire the most precise information by exerting maximum effort through the choice of how such endorsements are provided.

To provide sharp guidance on whether dynamic reporting provides the strongest incentives, we keep the contracting environment otherwise identical between single-elicitation and dynamic cases, except for the ability to receive reports from the agent at any time. Given this, the designer's problem in either contracting environment is to decide how to provide the endorsement or recommendation as a function of (a) the forecaster's static or dynamic prediction and (b) the outcome that arises.

Note that a crucial feature of our model is that both effort and learning are private. This reflects our interest in cases where the sought-after signal itself requires expertise to recognize and interpret. In particular, the assumption that learning is private separates us from the literature on contracting for experimentation, which typically allows contracts to increase rewards in later periods to incentivize continued effort. This is not the case in our problem: with private learning, the agent can conceal or misreport results, undermining any direct performance-based incentives.

The literature on contract design for forecasters has a much shorter tradition when applied to *dynamic* settings. We share this focus with Deb et al. (2018) who study the design of dynamic contracts to *screen forecaster ability* in the absence of moral hazard. While our models make similar assumptions about the available contracts, the key difference is our focus on dynamic moral hazard rather than initial adverse selection.

1.3 Results and Intuition

Our interest is in whether maximum effort can be implemented by a contract in which the principal promises to give the forecaster the desired endorsement with a probability depending on (i) a single belief report and (ii) the realized outcome. Such a contract is static if the reward options offered to the agent do not depend on time in any way. The challenge is that the incentives necessary to induce effort depend on the forecaster's current belief, and these belief-contingent incentive requirements generally conflict as beliefs evolve over time. Our first result, Theorem 1, characterizes effort-maximizing dynamic contracts themselves. Such contracts admit a decreasing no-information reward path: increasing rewards over time instead creates profitable "shirk-and-lie" deviations. This characterization sharply reduces the relevant contract space and provides the starting point for our main question, namely whether time variation in rewards is necessary for effort-maximization.

We briefly describe why belief evolution creates a fundamental tension for incentive provision—and why moving beyond static contracts could potentially help resolve this ten-

sion. Consider first a one-shot information acquisition problem, and notice that the scoring rule that provides the greatest gain from effort typically depends on the prior. Intuitively, if rewards place too much weight on the state that is currently more likely, an agent may prefer to stop and report a favorable signal rather than exert effort; the agent may similarly lack incentives to learn if rewards place insufficient weight on the initially less-likely state. However, the same logic implies that the reward structure required for continued effort *flips* once the agent becomes highly confident in the *other* state. Thus, rewards maximizing incentives at one belief can sharply reduce incentives at another, suggesting that different rewards should be available at different times (and under different beliefs). While this argument uses extreme beliefs to illustrate, the same issue arises throughout the belief space. We present a more precise illustration of how effort-maximizing rewards should be calibrated to the agent's belief—and that the incentive constraints associated with different beliefs place conflicting requirements on rewards—in Section 3.4.

Now, if beliefs in the dynamic setting were constant before a Poisson signal arrival—i.e., no information is conveyed by the failure to find a red (or green) flag—such adjustments would clearly be unnecessary: the optimization problem is identical at every point in time. We refer to this case as a *Stationary environment* and confirm that indeed static contracts suffice then (Theorem 2). The aforementioned tension emerges when beliefs evolve over time. As the agent becomes more confident in one state, should rewards in that state decrease to maintain incentives? The answer turns out to be no in the following cases:

1. If a red (or green) flag reveals the future outcome perfectly: That is, a *Perfect-learning* environment (Theorem 3) where signals fully reveal the state;

2. If only a red flag can arrive: That is, a *Single-signal* environment (Theorem 4) where a Poisson signal arrival always moves beliefs in one direction.

Outside of these cases, dynamic reports may expand the set of strategies that an agent can be induced to follow. We identify a set of parameters jointly violating these conditions for which richer contracts are necessary (Theorem 5).[4]

Our key insight for perfect-learning and single-signal environments is that, despite belief drift, adjusting rewards over time is counterproductive. Lowering rewards in increasingly-likely states might seem to sharpen incentives at later beliefs. But doing so reduces the agent's continuation value from working, weakening earlier incentives. In these environments, any such time-varying contract can be replaced by a static scoring rule that provides weakly stronger incentives at all times. However, when beliefs drift and signals can push

[4]In particular, as we describe in Section 6, these parameters essentially amount to a "sufficient violation" of the conditions of Theorems 3 and 4.

posteriors in either direction without fully revealing the state, conflicts arise that cannot be resolved by static contracts alone.

Our unifying message is that the benefits from dynamic contracts require informative, but imperfect, Poisson signals that can move beliefs in the same direction as the belief drift absent signal arrival. For *these* signals, adjusting rewards over time can strengthen incentives. In addition, the environment must feature *sufficiently important dynamics*. Section OA 6 shows that static and dynamic implementations coincide with only two periods—an agent cannot be incentivized to work for two periods by a dynamic contract if this is impossible under a static contract. Our characterization in Theorems 2 to 4 is notable because the conditions on the learning environment *do not* depend on the time horizon or effort cost. The necessity of dynamic contracts emerges once the horizon is sufficiently long.

We mention that effort-maximizing contracts can in general be computed numerically via linear programming, as we describe in Sections 3 and 4. This numerical procedure yields explicit solutions for effort-maximizing contracts, illustrating concretely how dynamics can influence contract design. The analytical results that we obtain serve a complementary goal: isolating conditions on primitives with a transparent economic interpretation.

Two further points clarify how dynamics influence the design of optimal contracts even in cases where a static implementation exists (Section 7.1). First, it need not be the case that the effort-maximizing scoring rule provides the strongest incentives at the prior, since incentives must be balanced as the agent's belief changes, as alluded to above. Second, the optimal scoring rule may require offering intermediate options that provide strictly positive rewards even for incorrect predictions—contrasting with results for static cases Li et al. (2022); Szalay (2005). This distinction reflects that beliefs evolve: an option that is "intermediate" early may be "extreme" later. While our analysis also speaks to other aspects of scoring rule design, we focus on the problem of static implementation as this is a sharp, qualitative contract characteristic whose practical relevance is transparent.

1.4 Related Literature

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.[5]

As noted above, 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). To the best of our knowledge, Osband (1989) is the earliest work sharing our focus on the question of incentivizing the *acquisition* of information. Other work relevant to the application of forecasting is Elliott and Timmermann (2016), which reviews the statistical properties of forecasting models. Aside from Deb et al. (2018), other papers focused on the problem of screening forecasters include Deb et al. (2026); Dasgupta (2023). More recent work in economics and computer science related to scoring rules 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); Bloedel and Segal (2024). Our main point of departure from this line of work stems from our focus on dynamics.[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.[8] 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 Poisson signals that may support either state and be inconclusive. Our main message is that dynamic contracts can outperform static scoring rules only under *both* modifications.

On this note, Poisson bandits have been extensively utilized in economic settings since the influential work of Keller et al. (2005). An advantage of this framework 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

[5]For instance, static information acquisition technologies where information is acquired after contracting (Krähmer and Strausz, 2011) or where the outcome of experimentation may be contractable (Yoder (2022), as well as Chade and Kovrijnykh (2016) in a repeated setting).

[6]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. Bloedel and Segal (2024) study dynamic contracts, but their framework models dynamics across different agents, whereas our model focuses on dynamics in which a single agent acquires information over time.

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

[8]For instance, Hörner and Samuelson (2013) relaxes commitment; Halac et al. (2016) allow for ex-ante adverse selection and transfers; Guo (2016) considers delegation without transfers; Fu (2024) studies optimal approval rules in a private sequential experimentation problem with selectively disclosed hard evidence, comparing environments with conclusive good news and conclusive bad news.

Mierendorff (2019); Damiano et al. (2020); Keller and Rady (2015); Bardhi et al. (2024); Lizzeri et al. (2024). Our exercise essentially amounts to *designing a (possibly dynamic) single-agent decision problem*. Note that the agent's problem need not admit a simple stationary representation for arbitrary mechanisms in our framework.[9] This 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) as a forecaster. A mechanism designer shares a common prior with the agent over an uncertain event $\theta \in \Theta = \{0, 1\}$ (e.g., whether an investment will achieve a certain target outcome, whether an adversary will attack on a certain date, etc.); we let D denote the initial probability that $\theta = 1$.

The agent is able to acquire information about θ at discrete times $\{1, 2, \ldots, T\}$.[10] We refer to θ as the unknown *state*. This state is contractible and is only realized after time T (e.g., a successful prototype or an attempted attack from the adversary). For conceptual simplicity, we take $T < \infty$, and interpret T as the deadline for the designer to make a payoff relevant decision regarding the unknown state (e.g., whether to invest in the company or attack the adversary). The detailed decision environment plays no role in our analysis; we assume only that the designer weakly prefers more information (a la Blackwell).

2.1 Information Acquisition

At any time $t \leq T$, the agent can acquire information by paying a cost $c > 0$. When this cost is paid, a piece of (falsifiable) evidence informative of θ may be discovered by the agent with some probability. We refer to this evidence as a *Poisson signal*. We take the arrival probability of the Poisson signal to be λ_θ. Without loss of generality, we assume that $\lambda_1 \geq \lambda_0$, so the agent's belief drifts toward state 0 as no Poisson signal arrives.

If no Poisson signal 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 a probability of 1. When the Poisson signal arrives, various pieces of information may potentially be observed

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

[10] All of our results extend if we consider a variable time interval Δ and take the high-frequency limit as $\Delta \to 0$.

by the agent, leading to different beliefs. Specifically, we denote the set of *non-null* signals as S, and the agent observes a signal $s \in S$ with probability λ_θ^s by exerting effort, where $\sum_{s\in S} \lambda_\theta^s = \lambda_\theta$ for all states $\theta \in \{0, 1\}$. By Bayes' rule, at any time $t \in \{1, \ldots, T\}$, assuming that the agent has exerted effort for all periods until t, we let μ_t^N denote the posterior belief that $\theta = 1$ if no Poisson signal arrived before time t (including t), and let μ_t^s denote the agent's posterior when receiving Poisson signals s exactly at time t:

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

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

where $\mu_0^N = D$ is the prior belief.

Once the Poisson signal arrives, no further information can be acquired (e.g., if a red flag reveals the main point of concern so that additional scrutiny will not meaningfully alter the assessment of the investment's viability; if an adversary's capabilities are determined, no further information is relevant for assessing attack probability, etc.).[11] Note that our information acquisition technology generalizes Poisson bandit learning (as in Halac et al. (2016), for instance), because signals (a) need not reveal the state and (b) can increase the posterior probability of either state.

In addition, our main results on static contracts implementing maximal effort focus on the following canonical cases of this model:

1. **Stationary environments,** where $\lambda_1 = \lambda_0$.
2. **Perfect-learning environments,** where either $\lambda_0^s = 0$ or $\lambda_1^s = 0$ for every $s \in S$.
3. **Single-signal environments,** where $|S| = 1$.

Intuitively, in stationary environments, the mere passage of time conveys no information: silence (null signals) leaves beliefs unchanged, so only the content of an arriving non-null signal can move the posterior. In perfect-learning environments, every non-null signal is state-exclusive. Once any such signal arrives, the state is fully revealed. In single-signal environments, there is only one kind of evidence; thus, when the signal arrives, it shifts

[11] The assumption of a single-signal arrival represents an extreme case of information attrition in Strulovici (2022), which discusses various compelling practical instances where the number of available signals is naturally limited. In our case, this feature enhances tractability by allowing us to associate the amount of information produced with the length of time the agent works absent the arrival of a Poisson signal. It also avoids known complications in determining the agent's payoffs under Poisson learning when signals are not perfectly revealing.

beliefs in a single direction, toward the state that more readily generates that signal. If beliefs drift toward 0 when the Poisson signal does not arrive, they *must* jump toward 1 whenever it does in the single-signal environment.

To reiterate, effort choices and signal arrivals are private information of the agent. The assumption that this signal is private reflects the notion that it requires expertise to recognize and interpret. Moreover, the signals are not verifiable and, hence, can be arbitrarily fabricated or misrepresented by the agent at no cost. The assumption that effort is private reflects the notion that the principal cannot easily monitor the agent's activity.

2.2 Contracting

We will focus on a menu representation of dynamic contracts for incentivizing the agent to exert costly effort.[12] Specifically, the designer offers the agent a menu of rewards

$$\mathcal{M} := \{\mathcal{M}_t \subseteq [0,1]^2\}_{t=1}^{T}.$$

At any time $t \leq T$, if the agent picks a reward profile $r_t = (r_{t,0}, r_{t,1}) \in \mathcal{M}_t$, the agent receives a reward of $r_{t,\theta}$ after time T when the state θ is publicly revealed. As mentioned in the introduction, we can also interpret the reward as the probability of receiving an endorsement that positively influences the forecaster's reputation. Such reward bounds are also widely assumed in the literature on evaluating forecasters (e.g., Deb et al., 2018), as well as in the literature on information elicitation more generally.[13] Note that a crucial feature of this menu representation is that the agent only makes one irrevocable choice for the reward vectors. This is without loss of generality since we have assumed that the Poisson signals can arrive at most once over the entire time horizon.

We assume that the agent does not discount the future.[14] That is, by exerting effort for Z periods and receiving a reward of $r \in [0,1]$ in the end, the utility of the agent is

$$r - cZ.$$

[12] A more general format based on arbitrary communication and its equivalence to our menu representation is provided in Section OA 2.2.

[13] While we interpret the reward as non-monetary, our model also applies to applications where the designer faces explicit budget constraints set by a third party. Anthony et al. (2007) discusses firms imposing budget constraints on different divisions; National meteorological agencies, such as the U.S. National Weather Service (NWS) receive public funding for operations; the Congressional Budget Office (CBO) operates under a fixed annual budget and provides forecasts for Congress. In such cases, designers are restricted to using this budget to incentivize information acquisition.

[14] We discuss the extensions with discounted utilities in Section 7.2.

Given a menu $\mathcal{M}$, let $Z_{\mathcal{M}}$ denote the number of periods the agent would exert effort in the absence of a Poisson signal.[15] In our Poisson learning environment, the induced experiment (the distribution over posteriors) is uniquely determined by $Z_{\mathcal{M}}$. Moreover, a larger $Z_{\mathcal{M}}$ yields Blackwell more informative signals. Therefore, the designer's problem is

$$\mathcal{M}^* \in \arg\max_{\mathcal{M}} Z_{\mathcal{M}}.$$

This objective reflects the designer's goal of providing the strongest possible incentives for learning.

Contracts must address both hidden effort and hidden information: the principal does not observe whether effort was exerted, when a Poisson signal arrived, or which signal was observed. These frictions play distinct but interacting roles. Hidden effort creates the need to provide incentives for information acquisition. Private and nonverifiable learning constrains how those incentives can be provided, because rewards designed to encourage effort must also preserve the agent's incentives to select the intended report. Private signal arrival further allows the agent to conceal when information was obtained, making the timing of available rewards strategically relevant. The comparison between static and dynamic elicitation therefore asks whether time-contingent menus can strengthen effort incentives while maintaining reporting incentives as beliefs evolve. Section 3.2 outlines the resulting incentive compatibility constraints that arise given the class of contracts we consider.

2.3 Static Implementation

Since the agent's signal is private, the designer cannot compel early selection: at any date t, the agent can guarantee himself access to any reward that will appear in the future simply by waiting. This observation immediately implies the following Lemma:

Lemma 1 (Shrinking Menus are Without Loss)**.** *For any menu profile $\mathcal{M} = (\mathcal{M}_t)_{t=1}^T$, define*

$$\widetilde{\mathcal{M}}_t = \bigcup_{t' \geq t} \mathcal{M}_{t'}.$$

Then $\widetilde{\mathcal{M}}$ is weakly shrinking and induces the same feasible reward choices, and hence the same agent behavior, as $\mathcal{M}$.

The weakly shrinking property follows immediately from the definition of $\widetilde{\mathcal{M}}_t$, i.e.:

[15] We do not obtain a closed form for $Z_{\mathcal{M}}$, as the best response strategy of the agent can potentially be complex given an arbitrary menu $\mathcal{M}$. We will provide a more transparent description in Section 3.1 when we simplify the best response strategies of the agent.

$$\widetilde{\mathcal{M}}_t \supseteq \widetilde{\mathcal{M}}_{t'} \qquad \text{for all } 1 \leq t \leq t' \leq T. \tag{1}$$

In light of this nested (weakly shrinking) structure of dynamic contracts, we now ask whether time variation is necessary at all, or if a single, time-invariant menu maximizes effort. The next definition formalizes this possibility, which we call a *static implementation*.

Definition 1 (Static Implementation)**.**
An optimal menu $\mathcal{M}$ has a static implementation if there exists $\widehat{\mathcal{M}}$ such that $Z_{\widehat{\mathcal{M}}} = Z_{\mathcal{M}}$ and

$$\widehat{\mathcal{M}}_t = \widehat{\mathcal{M}}_{t'}, \qquad \textit{for any } 1 \leq t \leq t' \leq T.$$

Under a static implementation, the designer does not need to monitor the exact time at which the agent observes the Poisson signal to incentivize maximum effort. Given a menu with a static implementation, the agent can defer the choice to time T and select the reward based on the aggregated information acquired over the entire process. Under a static implementation, advice is summarized at the end of the interaction.

Using terminology from the information elicitation literature, a static implementation essentially offers a *scoring rule* at time T: Specifically, a scoring rule

$$P : \Delta(\Theta) \times \Theta \to [0, 1]$$

maps the agent's report—which is required to take the form of a posterior belief—and the realized state to a reward. Scoring rules are far simpler than arbitrary dynamic menus: they require only a summary (in the form of a belief report) rather than real-time monitoring and time-varying rewards. In contrast, richer dynamic contracts may require closer integration within the decision-maker's organization. Our goal is to characterize whether unrestricted dynamic contracting is necessary to incentivize maximum information acquisition:

Main Question: *Does the effort-maximizing menu have a static implementation?*

3 Structures of Dynamic Incentives

Before answering our main question, we provide some useful results that help us formulate the approach we take. Section 3.4 provides some additional discussion and intuition.

3.1 Simplifying Effort Strategies to Stopping Strategies

The agent's information acquisition strategy can be arbitrary and complex in dynamic environments. But in our model, it is without loss of generality for the agent to front-load all effort and adopt a *stopping strategy*—i.e., a strategy in which the agent simply decides *when* to stop working in the absence of a Poisson signal arrival (exerting effort until then).

More precisely, let σ denote the (random) time of the first Poisson signal (with $\sigma = \infty$ if no such signal arrives). A stopping strategy with stopping time $\tau \leq T$ prescribes effort at any time period $t \leq \min\{\tau, \sigma\}$ and no effort thereafter: effort ceases after the first occurrence of a Poisson signal or $\{t = \tau\}$. Slightly abusing notation, we also use τ to represent the stopping strategy with a stopping time τ. Let $\tau_{\mathcal{M}}$ be the stopping strategy implemented under the menu $\mathcal{M}$.[16]

Lemma 2 (Stopping Strategies are Without Loss)**.**
Given any menu of rewards $\mathcal{M}$ and any best response of the agent with maximum effort length $Z_{\mathcal{M}}$ conditional on not receiving any Poisson signal,[17] *there exists a stopping strategy $\tau_{\mathcal{M}}$ with $\tau_{\mathcal{M}} \geq Z_{\mathcal{M}}$ that is also optimal for the agent.*

The proof of Lemma 2 is simple. Since the agent only learns about the state when exerting effort, and since effort can always be concealed until a later date, any prescribed effort strategy can be modified to one where effort is "front-loaded," with all effort choices moved earlier without increasing its cost or reducing the feasible reporting options. A best response can thus be chosen to exert effort consecutively until either a Poisson signal arrives or the relevant stopping time is reached.

Given Lemma 2, the designer's problem reduces to

$$\mathcal{M}^* \in \arg\max_{\mathcal{M}} \tau_{\mathcal{M}}.$$

3.2 The Incentive Structure of Dynamic Contracts

3.2.1 Information Revelation Incentives

Since it is without loss of generality to consider menus with a shrinking choice set (see Equation (1)), it is in the agent's best interest to select a reward from the available menu immediately once he hits his stopping time in a stopping strategy, i.e., when he receives a Poisson signal or decides not to exert effort anymore.

[16] We break ties by maximizing the stopping time if there are multiple stopping strategies.
[17] If the agent randomizes, we let $Z_{\mathcal{M}}$ denote the maximum effort length among all realizations.

Given any menu, the designer can infer the posterior belief of the agent based on his choice of rewards. Moreover, only rewards that maximize the on-path beliefs can be chosen. Specifically, at any time t, for any Poisson signal $s \in S$, let $r_t^s \in \mathcal{M}_t$ be the reward vector chosen by the agent at time t when his belief is μ_t^s. Moreover, let $r_{\tau_\mathcal{M}}^N$ be the reward vector chosen by the agent at stopping time $\tau_\mathcal{M}$ when his belief is $\mu_{\tau_\mathcal{M}}^N$. Any rewards beyond the collection of $\{r_t^s\}_{t\leq\tau_\mathcal{M},s\in S}\cup\{r_{\tau_\mathcal{M}}^N\}$ will never be chosen by the agent. Let $u(\mu, r) \triangleq \mathbf{E}_{\theta\sim\mu}[r(\theta)]$ denote the agent's expected utility with belief μ under the reward function r.

Lemma 3 (Simplified Menu Representation)**.**
Given any menu $\mathcal{M}$ implementing stopping time $\tau_\mathcal{M}$, let $\{r_t^s\}_{t\leq\tau_\mathcal{M},s\in S}\cup\{r_{\tau_\mathcal{M}}^N\}$ be the collection of rewards chosen by the agent at some on-path history. At any $t \leq \tau_\mathcal{M}$, it is equivalent to offer a menu $\mathcal{M}_t \triangleq \{r_{t'}^s\}_{t\leq t'\leq\tau_\mathcal{M},s\in S} \cup \{r_{\tau_\mathcal{M}}^N\}$ to the agent such that

$$r_t^s \in \arg\max_{r\in\mathcal{M}_t} u(\mu_t^s, r), \qquad \forall s \in S. \tag{IC-Reporting}$$

In essence, Lemma 3 is the relevant version of the taxation principle for our environment.

For any time $t \leq \tau_\mathcal{M}$ and any Poisson signal $s \in S$, let $u_t^s(\mathcal{M}) \triangleq u(\mu_t^s, r_t^s)$ be the utility of the agent when his belief is μ_t^s. Similarly, for any $t \leq \tau_\mathcal{M}$, let

$$r_t^N \in \arg\max_{r\in\mathcal{M}_t} u(\mu_t^N, r), \qquad u_t^N(\mathcal{M}) \triangleq u(\mu_t^N, r_t^N)$$

with the convention that μ_0^N is the prior D and $\mathcal{M}_0 = \mathcal{M}_1$. We omit the dependence on $\mathcal{M}$ whenever there is no risk of confusion.

3.2.2 Effort Incentives

We introduce a notion that allows us to describe the incentives for exerting effort at a given time. Specifically, we decompose the dynamic problem into a sequence of static problems. Each static problem, in turn, considers the relevant continuation incentives of the agent under the dynamic contract.

Each static game in our decomposition is indexed by a pair of times, $t \leq t' \leq T$, as well as a menu of rewards $\mathcal{M}$. We define the *continuation game between t and t'*, denoted $\mathcal{G}_{t,t'}^\mathcal{M}$, as the *static* decision problem with prior belief μ_{t-1}^N where the agent makes a one-time binary effort choice at time t: (a) Exert no effort or (b) Exert effort up to and including time t' or until a Poisson signal arrives. We let $C(t, t', c)$ denote the cost of option (b) in the continuation game between t and t' when the cost of effort in each period is c. Letting $f_t^s(t')$ be the probability of receiving Poisson signal s at time t' conditional on not receiving Poisson

signals before time t, and letting $F_t^s(t')$ be the corresponding cumulative probability:

$$C(t,t',c) = \left(1 - \sum_{s\in S} F_t^s(t')\right) \cdot c \cdot (t'-t+1) + \sum_{\hat{t}=t}^{t'} \sum_{s\in S} f_t^s(\hat{t}) \cdot c \cdot (\hat{t}-t+1).$$

Indeed, this formula is simply the expected total cost of exerting effort from time t until the earlier of t' or the first Poisson signal arrival.

An important special case is when the time t' is itself the stopping time for the menu $\mathcal{M}$, with $\tau_{\mathcal{M}} \leq T$. In this case, we say $\mathcal{G}^{\mathcal{M}}_{t,\tau_{\mathcal{M}}}$ is the continuation game at time t for menu $\mathcal{M}$. In what follows, we omit the superscript of $\mathcal{M}$ when it is clear from context. Note that given any menu of rewards $\mathcal{M}$ with stopping time $\tau_{\mathcal{M}}$, at any time $t \leq T$, it is straightforward to show that the following are equivalent:

1. $t \leq \tau_{\mathcal{M}}$, i.e., the agent has an incentive to exert effort at time t in this dynamic environment;

2. there exists $t' \geq t$ such that the agent has an incentive to exert effort in the continuation game $\mathcal{G}_{t,t'}$ (corresponding to some feasible continuation under $\mathcal{M}$).

When $t \leq \tau_{\mathcal{M}}$, a feasible choice of $t' \geq t$ that ensures sufficient effort incentives in the continuation game is $t' = \tau_{\mathcal{M}}$, as this is consistent with the agent playing a best response.

The bottom line is that to design a menu of rewards that implements any $\tau \leq T$, it suffices to ensure that the agent has an incentive to exert effort in continuation games $\mathcal{G}_{t,\tau}$ for all $t \leq \tau$. Now, fixing c and the desired stopping time τ, there typically will exist multiple effort-maximizing contracts, as incentives could be slack given the cost c. An alternative is to consider an auxiliary problem of *maximizing effort incentives*. We describe formally what maximizing effort incentives means: given any continuation game $\mathcal{G}_{t,t'}$, let $\Delta(\mathcal{G}_{t,t'})$ be the difference in expected reward between exerting effort and not exerting effort. That is,

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

The last term $u(\mu_{t-1}^N, r_{t-1}^N)$ is the agent's utility for not exerting effort in the continuation game $\mathcal{G}_{t,t'}$. The index in this term is $t-1$ since the prior belief in this continuation game is μ_{t-1}^N, and if the agent decides not to exert effort from t onward, the agent could actually make the report at time $t-1$ to get his favorite reward option r_{t-1}^N. Thus, the agent has an incentive to exert effort in the continuation game $\mathcal{G}$ if and only if $\Delta(\mathcal{G}_{t,t'}) \geq C(t,t',c)$.

By the stopping-strategy reduction, the effort component of implementing $\tau_{\mathcal{M}}$ is sum-

marized by the following constraints:

$$\Delta(\mathcal{G}_{t,\tau_{\mathcal{M}}}) \geq C(t, \tau_{\mathcal{M}}, c) \qquad \forall t \leq \tau_{\mathcal{M}}. \tag{IC-Effort}$$

This reduces the relevant effort constraints to (IC-Effort). Moreover, effort incentives are stronger when the difference in expected rewards between exerting and not exerting effort is larger.

3.3 The Structure of Effort-Maximizing Dynamic Contracts

We can now present a characterization of effort-maximizing contracts that arise in light of the simplifications above. To summarize, constraints (IC-Reporting) and (IC-Effort) capture the two dimensions of the agent's incentives: The former requires the agent to select the intended reward given the information acquired, while the latter requires him to prefer the prescribed continuation of experimentation to stopping. Notice that these constraints *also* accommodate joint deviations of effort and reporting. If the agent stops exerting effort after the no-information history at time $t-1$, he may select whichever reward in $\mathcal{M}_{t-1}$ is most attractive at belief μ^N_{t-1}. By definition, his payoff from the best such report is $u(\mu^N_{t-1}, r^N_{t-1})$, which is the outside option subtracted in $\Delta(\mathcal{G}_{t,\tau_{\mathcal{M}}})$. Thus, IC-Effort compares the prescribed continuation with stopping followed by an optimal report, rather than only with truthful stopping. Together with the stopping-strategy reduction, imposing these constraints at every no-information history accounts for deviations involving fewer periods of effort, while IC-Reporting governs reporting deviations following a Poisson signal.

For our characterization, it is useful to classify signals according to whether they move beliefs to the left or right of the no-information posterior. We say that a signal s is *left-biased* if $\mu^s_t \leq \mu^N_t$ for all t, and *right-biased* otherwise. Given any time $t \leq \tau_{\mathcal{M}}$, define

$$G_t(\mu) \triangleq \sup_{r:\Theta\to[0,1]} \left\{ u(\mu, r) : u(\mu^N_{t'}, r) \leq u^N_{t'},\ \forall t' \in \{0, \ldots, t\} \right\}. \tag{2}$$

Since each reward induces an affine function of μ, G_t is the greatest *convex minorant* of the no-information points $\{(\mu^N_{t'}, u^N_{t'})\}_{t'=0}^{t}$ generated by feasible reward vectors bounded in $[0,1]^2$.

Theorem 1 (Effort-Maximizing Dynamic Contracts)**.**
For any prior D and any signal arrival rates λ, there exists an effort-maximizing menu $\mathcal{M}$ with stopping time $\tau_{\mathcal{M}}$ and an associated reward sequence $\{r^N_t, r^s_t\}_{t\leq\tau_{\mathcal{M}}, s\in S}$ such that:[18]

[18]Even though our menu representation (Lemma 3) does not require the specification of r^N_t for

1. ***Decreasing no-information rewards:*** *for all* $1 \leq t \leq t' \leq \tau_{\mathcal{M}}$ *and* $\theta \in \Theta$,

$$r^N_{t',\theta} \leq r^N_{t,\theta}.$$

Moreover, $r^N_{t,0} = 1$ *whenever* $r^N_{t,1} > 0$. *Equivalently, the sequence* r^N_t *first decreases in the reward for state* 1, *and once that reward reaches* 0, *it decreases in the reward for state* 0.

2. ***Left-biased signals:*** *if* s *is left-biased, then one can choose*

$$r^s_t = r^N_t \qquad \text{for all } t \leq \tau_{\mathcal{M}}.$$

3. ***Right-biased signals:*** *if* s *is right-biased, then letting* G_t *be defined by* (2),

$$u(\mu^s_t, r^s_t) = G_t(\mu^s_t),$$

and the affine function $\mu \mapsto u(\mu, r^s_t)$ *is a supporting line to* G_t *at* μ^s_t.

Compared to the menu representation in Lemma 3, Theorem 1 reduces the relevant structure to the path of no-information rewards $\{r^N_t\}_{t \leq \tau_{\mathcal{M}}}$. Part 1 gives a one-switch structure: the reward first decreases in state 1, and once that reward reaches 0, it decreases in state 0. Parts 2 and 3 then pin down all signal-contingent rewards from that path. Because a reward vector has only two coordinates, any right-biased reward is pinned down by at most two binding no-information constraints, together, possibly, with one boundary constraint from the box $[0,1]^2$.

A useful way to understand Parts 2 and 3 is to fix the no-information path $\{r^N_t\}_{t \leq \tau_{\mathcal{M}}}$ from Part 1. Because menus are shrinking, any reward selected at date t must also be feasible for every earlier no-information type. Thus, conditional on this path, the reward assigned after signal s at date t solves the one-shot problem

$$\max_{r:\Theta \to [0,1]} u(\mu^s_t, r) \qquad \text{s.t.} \qquad u(\mu^N_{t'}, r) \leq u^N_{t'} \quad \forall t' \in \{0, \ldots, t\}.$$

This auxiliary problem chooses the reward that is best for the signal type while not raising the stopping utility of any earlier no-information type. When s is left-biased, the canonical no-information reward r^N_t solves this problem. When s is right-biased, the value of the problem is exactly $G_t(\mu^s_t)$, so the optimal reward is a supporting line to the convex minorant.

When we later specialize to the binary-signal case $S = \{L, R\}$, with L left-biased and

$t < \tau_{\mathcal{M}}$, we include them here to highlight the structure of the optimal contract.

R right-biased, Theorem 1 implies that

$$r_t^L = r_t^N \qquad \text{for all } t \leq \tau_{\mathcal{M}},$$

while r_t^R is given by the supporting line to G_t at μ_t^R.

The intuition for the decreasing reward structure in Part 1 is the same as in the introduction: if the agent expects weakly higher rewards later, he may shirk immediately and imitate a later report or defer reporting an early signal in order to enjoy the later reward. Decreasing the no-information reward can instead strengthen incentives. After a null signal, the posterior drifts toward state 0, making rewards tilted toward state 0 more attractive and reducing the value of additional information. Lowering those rewards over time restores dispersion in continuation payoffs and thereby encourages continued effort.

3.4 Discussions and Intuitions

An auxiliary perspective on effort incentives The design of the optimal menu can be reduced to the problem of maximizing the effort incentive for all continuation games before the stopping time. More concretely, for any stopping time $\tau \leq T$, we can solve the following linear program:

$$\begin{aligned} c_\tau := \max_{\mathcal{M},\tilde{c}} \quad & \tilde{c} && \text{(Effort Maximization)} \\ \text{s.t.} \quad & \Delta(\mathcal{G}_{t,\tau}^{\mathcal{M}}) \geq C(t,\tau,\tilde{c}), \quad \forall t \leq \tau. \\ & \mathcal{M} \text{ satisfies (IC-Reporting).} \end{aligned}$$

Because $C(t,\tau,c)$ is increasing in the per-period cost c, the set of costs at which a fixed stopping time τ is implementable is downward closed (i.e., the same stopping time is implementable even when costs are lower). Consequently, c_τ is the *cutoff cost* for implementing τ. For any primitive effort cost c, the designer's original problem is therefore equivalently:

$$\tau^*(c) = \max\{\tau \leq T : c_\tau \geq c\}.$$

To find the optimal menu, we can enumerate all $\tau \leq T$ and identify the largest τ^* such that $c_{\tau^*} \geq c$. The reward menu implementing τ^* is an optimal menu.

Optimal menu for continuation games We first consider the design of the optimal menu that maximizes the reward difference in the continuation game given a pair of time $t \leq \tau$. In such static environments, this is essentially the optimization of the scoring rules,

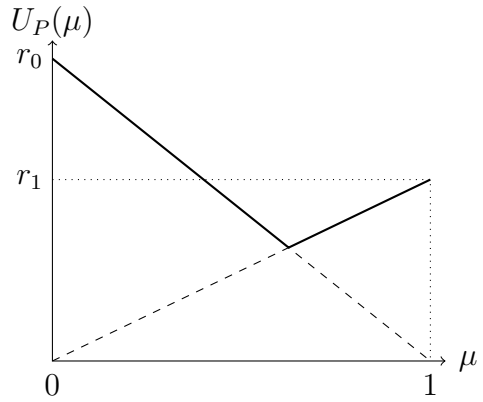

Figure 1: Expected score $U_P(\mu)$ of a V-shaped scoring rule P with parameters r_0, r_1.

which is fully characterized in Li et al. (2022).

Definition 2 (V-shaped Scoring Rules)**.**
A scoring rule P is V-shaped with parameters r_0, r_1 if

$$P(\mu, \theta) = \begin{cases} r_1 & \mu \geq \frac{r_0}{r_1 + r_0} \text{ and } \theta = 1 \\ r_0 & \mu < \frac{r_0}{r_1 + r_0} \text{ and } \theta = 0 \\ 0 & \text{otherwise.} \end{cases}$$

We say P is a V-shaped scoring rule with a kink at $\hat{\mu} \in [0, 1]$ if the parameters r_0, r_1 satisfy $r_0 = 1, r_1 = \frac{1-\hat{\mu}}{\hat{\mu}}$ if $\hat{\mu} \geq \frac{1}{2}$ and $r_0 = \frac{\hat{\mu}}{1-\hat{\mu}}, r_1 = 1$ if $\hat{\mu} < \frac{1}{2}$.

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 1. Furthermore, given any V-shaped scoring rule P with a kink at D, the agent with prior belief D is indifferent between reward options $(r_0, 0)$ and $(0, r_1)$.

Proposition 1 (Li et al., 2022)**.** *For any $t \leq \tau$, the reward menu offering the V-shaped scoring rule with kink at μ_{t-1}^N in all periods maximizes the reward difference in continuation game $\mathcal{G}_{t,\tau}$.*

Intuitively, V-shaped scoring rules maximize the expected reward at all posteriors, subject to (a) the constraint that the indirect utility is convex (as a consequence of incentive compatibility) and (b) the expected reward at the prior being 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 1 would only decrease the agent's expected utility under that information structure, diminishing effort incentives. And moving the kink increases the payoff from not exerting effort by more than the payoff from exerting effort.

Dynamic effort incentives Proposition 1 illustrates the tensions involved in designing the optimal menu in dynamic environments. As the agent's posterior evolves over time, the priors of the continuation games at different time periods vary, leading to inconsistencies in the scoring rules that maximize incentives for effort across time. In particular, to implement maximum effort in dynamic environments, the moral hazard constraints often bind at *both* time 1 and the stopping time τ when the signals are perfectly revealing. A V-shaped scoring rule with a kink at μ_τ^N would yield insufficient incentives for the agent to exert effort at time 0, preventing the agent from working in the first place. 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. 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 discuss, the optimal scoring rule need not always take a V-shaped form to provide balanced incentives in dynamic environments.

Comparison to contracts for experimentation In standard experimentation models with publicly observed outputs (e.g., Halac et al., 2016), the principal can condition rewards on outputs, potentially increasing rewards over time to offset rising effort costs. In our setting, this approach is infeasible: an agent with favorable early "outputs" (i.e., Poisson signals) could strategically withhold information to claim higher rewards later. Instead, to strengthen incentives for continued effort, the optimal dynamic contract decreases rewards over time. This reduction in rewards is beneficial because it minimizes the agent's utility from stopping early, thereby encouraging continued effort. On top of this, the arrival of "output" yields private information for the agent. This additional endogenous private information creates an extra complication in our model.

4 Numerical Illustrations

This section presents numerical computations of optimal contracts for an important special case of our model which we will revisit later, namely when the set of Poisson signals is binary:

$$S = \{L, R\}.$$

Furthermore, we assume that $\mu_t^L < \mu_t^N < \mu_t^R$—thus, the left-biased signal L moves beliefs toward 0, and the right-biased signal R moves beliefs toward 1. Specializing to this case, Theorem 1 implies that

$$r_t^L = r_t^N \qquad \text{for all } t \leq \tau_{\mathcal{M}},$$

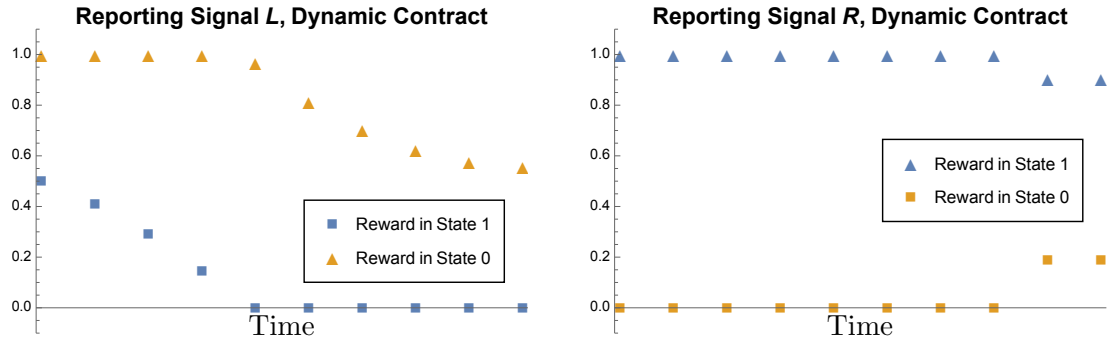

Figure 2: Solution to the linear program for the effort-maximizing contract with 10 periods, allowing for reporting at any time—i.e., unrestricted advice. The parameters are chosen as $\lambda_1^R = 1/3, \lambda_1^L = \lambda_0^R = 1/10, \lambda_0^L = 1/5, D = 2/3$.

while r_t^R is given by the supporting line to G_t at μ_t^R.

Note that since the optimization program (Effort Maximization) is a linear program for any $\tau \leq T$, it suggests a general method for how to numerically find an optimal menu by applying standard linear programming algorithms. The solution to the linear program not only tells us how high c can be, but also identifies a particular contract that can implement efforts for τ periods given this c.

Figure 2 presents a solution for certain representative parameter values. We compute the optimal menu for implementing a stopping time $\tau = 10$. The left side of Figure 2 illustrates the decreasing reward structure in the optimal menu characterized in Theorem 1: rewards following L signals are first decreasing for state 1, then decreasing for state 0 after their rewards for state 1 hit 0. Moreover, the realized rewards following R signals are not necessarily decreasing over time in all states; in particular, rewards in state 0 may increase over time, although due to incentive compatibility, such an increase in rewards in state 0 implies a decrease in rewards in state 1.

We now describe how to modify the linear program to restrict it to *static* contracts. Notice that the optimal dynamic menu allows the principal to discriminate over time: An agent who observes signal s at time t cannot choose the reward function that would have been selected had that signal arrived at time $t' < t$. Such deviations *are* possible under single-elicitation mechanisms. Therefore, the incentive compatibility constraints in the static mechanisms would instead require that for any $t \leq \tau_{\mathcal{M}}$,

$$r_t^s \in \arg\max_{r \in \mathcal{M}_0} u(\mu_t^s, r), \qquad \forall s \in S. \tag{Static-IC}$$

By replacing the (IC-Reporting) with (Static-IC) in (Effort Maximization), we obtain the

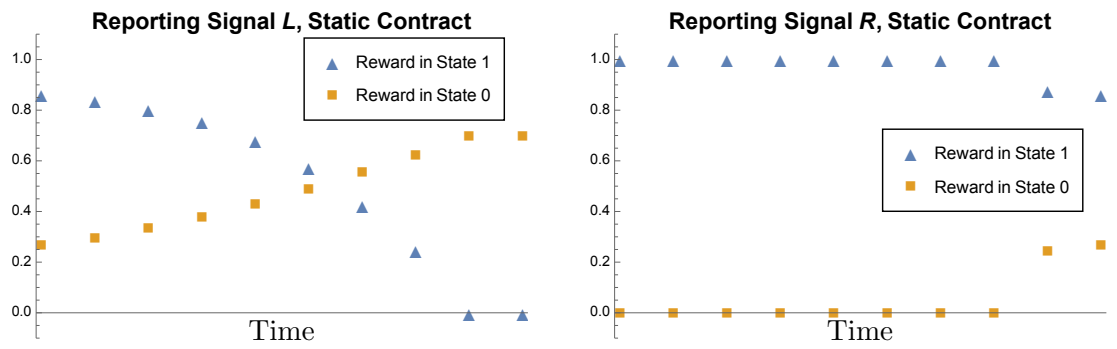

Figure 3: Solution to the linear program for the effort-maximizing contract with 10 periods, assuming that the contract involves a single elicitation at time 10—i.e., summarized advice. Note that while time is illustrated, any reward can be selected at any time, and the rewards presented are those selected consistent with incentive compatibility. The parameters are chosen as $\lambda_1^R = 1/3, \lambda_1^L = \lambda_0^R = 1/10, \lambda_0^L = 1/5, D = 2/3$.

linear program for optimization over scoring rules in this dynamic environment.

We illustrate the solution to this linear program in Figure 3. Here, the different rewards correspond to what the agent would optimally select from the *same* menu if the signal were to arrive at that time, even though all reports are made after time 10. A notable feature is that the rewards in state 0 following signal L are increasing over time, rather than decreasing as in the dynamic case. Intuitively, the dynamic contract in Figure 2 uses a relatively high early reward in state 0 after L to encourage the agent to begin working, and then lowers it over time; however, under static contracts, this front-loading would violate incentive compatibility because an agent who worked longer would prefer to "backdate" his report and pick that generous early L reward. To prevent such profitable backdating, the optimal static menu instead makes various offers at the same time, and the agent receiving a L signal at a later time would be more certain of the state being 0, selecting a reward option with a higher reward for state 0. Dynamic elicitation avoids this tension because the contract available at time t is not available at later times, so the principal can start with high incentives and taper them without creating incentives to mimic earlier types.

These computations show that a contract implementing maximum effort among *static* contracts can be qualitatively quite different from those that can resort to dynamic elicitation. The linear programs further identify a range of costs such that the principal can implement effort up to time 10 under a dynamic contract, but not using static elicitation (specifically, whenever $0.041 \leq c \leq 0.049$). For this range of cost parameters, dynamic elicitation enables the principal to incentivize *strictly more* effort. Our main results show that no gap arises in the single-signal, stationary, or perfectly revealing environments. Crucially, the parameters for these solutions do not fit under any of these environments.

These illustrations show that some condition is necessary for static contracting to implement maximum effort. One natural way to try to prove the sufficiency of summarized advice would be to show that the constraints added to (Static-IC) do not bind under these assumptions. The limitation is that it is not immediately clear how to determine which constraints will bind at which time given the properties of the linear program alone. Hence, this exercise motivates different techniques to determine the *qualitative properties* of such contracts. The message we deliver is that the existence of a gap in this numerical example is attributable to the existence of noisy L realizations.

5 Implementing Maximum Effort via Scoring Rules

We now show that scoring rules implement maximum effort in the three environments highlighted in Section 2.1: stationary, perfect-learning, and single-signal.

The central challenge is that different beliefs require different reward options to maximize incentives for exerting effort. As the agent works without receiving a signal, his no-information belief μ_t^N drifts toward 0. When μ_t^N is close to 1, strong rewards for reporting $\theta = 1$ weaken effort incentives—so the agent would exert effort only if c is small. When μ_t^N approaches 0, the situation reverses: strong rewards for reporting $\theta = 1$ are necessary to encourage the agent to keep working. Effort-maximizing contracts balance the incentives across different periods as the posterior belief drifts from close to 1 to close to 0.

We demonstrate the sufficiency of static elicitation constructively: starting from any dynamic contract, we identify a static contract such that the agent's incentives for exerting effort are preserved in all continuation games. Anticipating Section 6, the "problematic" beliefs in our failure result arise when the agent obtains an imperfectly informative Poisson signal that moves beliefs toward state 0. Section 5.2 illustrates why the principal should never withhold rewards if the Poisson signal is *conclusive* that the state is 0; in particular, maximizing rewards at every time implies that these rewards should be made constant. Section 5.3 shows that the principal can convexify the agent's no-information utility to establish a static implementation when only one Poisson signal can arrive. While details differ across environments, the unifying principle is that, under these learning technologies, the structure of beliefs is sufficiently constrained to allow rewards to be adjusted in a way that renders the contract effectively static, thus managing the tension between providing incentives in early and late periods.

5.1 Stationary Environment

In the stationary environment, the agent's belief never changes in the absence of Poisson signals: $\mu_t^N = D$ for all $t \leq T$. This removes any intertemporal tension.

Theorem 2 (Stationary Environment)**.**
In the stationary environment, a V-shaped scoring rule with kink at D is effort-maximizing.

Proposition 1 implies that for any continuation game, the effort-maximizing reward depends only on the current belief. Since this belief is constant across time, the same V-shaped rule simultaneously maximizes incentives everywhere. There is no conflict between the incentives at date 0 and those at date T (or any date in between).

The more subtle cases arise when beliefs drift as the agent exerts effort. In those cases, the optimal scoring rule may not take the form of a V-shape with a kink at the prior, reflecting a compromise between early-period incentives (where beliefs are near D) and late-period incentives (where beliefs approach μ_τ^N).

5.2 Perfect-learning Environment

When signals fully reveal the state, the sufficiency of scoring rules is less immediate.[19] Note that when signals fully reveal the state, it is without loss to focus on binary signals $\{L, R\}$.

Theorem 3 (Perfect-learning Environment)**.**
In the perfect-learning environment, there exists $r_1 \in (0, 1]$ such that a V-shaped scoring rule with parameters $r_0 = 1$ and r_1 (and hence, a kink at $\frac{1}{1+r_1} \in [1/2, 1)$) implements maximum effort.

A key feature of perfect learning is that once a signal arrives, the problem ends: the agent becomes certain of the state. Thus, the only interior beliefs are those at which the agent has not observed any informative signal. This greatly simplifies the structure, as offering rewards in both states can only weaken incentives: it is without loss to give rewards only in the state the agent believes is more likely.

The proof follows two steps, each replacing part of an arbitrary dynamic contract with another contract that (weakly) strengthens effort incentives. Our discussion uses the notation for binary signals introduced in Section 4.

Step One: Maximizing rewards in state 0. The first step is to show that the principal should always maximize the reward for the state realization of 0 when a left-biased signal L is observed or no Poisson signal ever arrives.

[19] This setting generalizes perfect-learning technologies from past work by allowing signals to potentially reveal *either* state.

To see why, fix an optimal contract implementing stopping time $\tau_{\mathcal{M}}$, and consider the agent's decision in the period $\tau_{\mathcal{M}}$. Since the agent will stop working in the next period whether or not a Poisson signal is observed, the agent's final payoff depends only on the signal observed in that period. His payoff from exerting effort at time $\tau_{\mathcal{M}}$ is therefore:

$$-c + (1 - \mu^N_{\tau_{\mathcal{M}}-1})(\lambda^L_0 r^L_{\tau_{\mathcal{M}},0} + (1 - \lambda^L_0) r^N_{\tau_{\mathcal{M}},0}) + \mu^N_{\tau_{\mathcal{M}}-1} \lambda^R_1 r^R_{\tau_{\mathcal{M}},1}$$

Since rewards are decreasing over time by Theorem 1, and since the agent can always select the largest available reward, this expression is equal to:

$$-c + (1 - \mu^N_{\tau_{\mathcal{M}}-1}) r^L_{\tau_{\mathcal{M}},0} + \mu^N_{\tau_{\mathcal{M}}-1} \lambda^R_1 r^R_{\tau_{\mathcal{M}},1} \tag{3}$$

But stopping immediately and reporting 0, without exerting any effort at all, yields a payoff:

$$(1 - \mu^N_{\tau_{\mathcal{M}}-1}) r^L_{\tau_{\mathcal{M}}-1,0} \tag{4}$$

Taking the difference between (3) and (4) reveals that exerting effort is more favorable when $r^L_{\tau_{\mathcal{M}},0} - r^L_{\tau_{\mathcal{M}}-1,0}$ is larger. Since $r^L_{\tau_{\mathcal{M}},0} \leq r^L_{\tau_{\mathcal{M}}-1,0}$, raising both rewards—up to 1—can only increase this difference and therefore the gain from exerting effort.

Now consider the agent's incentives at time t more generally if the principal offers the full reward for a report of left-biased signal L. Adding this reward will always improve the continuation payoff from exerting effort by *at least* $(1 - \mu^N_{t-1})(1 - r^L_{t,0})$—following the previous logic, this change would exactly be the increase in payoffs if rewards following signal L were constant after time t. If, instead, they are strictly decreasing, then the change yields a strictly higher gain from effort. However, the gain from stopping effort is always *at most* $(1 - \mu^N_{t-1})(1 - r^L_{t-1,0})$—again, it would be exactly this amount if the agent were to report signal L when shirking; otherwise, there would be no change at all. Since the gain from exerting effort is always at least as large as the gain from shirking, this modification always (at least weakly) strengthens the incentives to exert effort.

Step Two: Collapsing the rewards after signal R into a single payment The second step shows that it suffices to offer just one menu option in case signal R is observed. The idea is to replace all rewards following signal R with the *single* menu option $r^R_{\hat{t}}$, where $\hat{t}$ is selected as the minimum time (or equivalently the highest belief $\mu^N_{\hat{t}}$) at which the agent would prefer rewards $(1, 0)$ to $r^R_{\hat{t}}$ in the absence of a Poisson signal arrival. Figure 4 illustrates this replacement: the red line represents the agent's expected payoff if choosing the reward $(1, 0)$, while the thick black line represents the agent's payoff if stopping effort and pretending to have obtained the R signal at a given on-path belief. The replacement

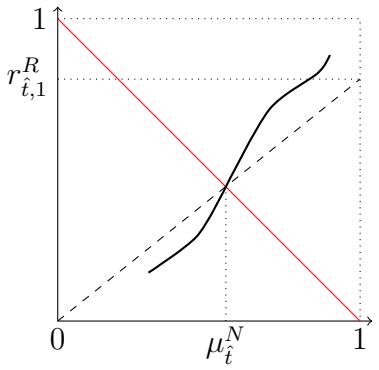

Figure 4: Illustration of the second step of the proof of Theorem 3

we identify is illustrated by the dashed black line.

Why does this modification weakly strengthen the incentives to exert effort? When $\mu^N_{t-1} \leq \mu^N_{\hat{t}}$ or equivalently $t \geq \hat{t}+1$, the agent's payoff from not exerting effort in the continuation game $\mathcal{G}_t$ remains unchanged under the new contract: the replacement is chosen precisely so that the reward option $(1,0)$ is selected at any such no information belief. But since rewards decrease over time, this modification ensures a larger reward following a right-biased signal R, strengthening the incentives to exert effort. The case of $\mu^N_{t-1} > \mu^N_{\hat{t}}$ or equivalently $t \leq \hat{t}$ is more subtle: Here, the payoff from both exerting effort *and* stopping in the continuation game $\mathcal{G}_t$ *decreases*. However, it turns out that the payoff from stopping decreases by more. Indeed, the decrease in the expected utility from stopping at no information belief μ^N_{t-1} is $\mu^N_{t-1}(r^R_{t-1,1} - r^R_{\hat{t},1})$. The decrease in expected utility from exerting effort is at most $\mu^N_{t-1}(r^R_{t,1} - r^R_{\hat{t},1})$ since there is a chance that an R signal is never observed, which implies that the probability of this change in rewards being relevant when exerting effort is less than μ^N_{t-1}. Thus, the payoff from exerting effort decreases by less than the payoff from stopping due to the decreasing rewards structure. Therefore, the agent is still willing to exert effort under this replacement. See Section B.2 for complete formal details.

5.3 Single-signal Environment

We now consider the environment where absent Poisson signals, μ^N_t drifts toward 0, while a single Poisson signal causes beliefs to jump upward.

Theorem 4 (Single-signal Environment)**.**
In the single-signal environment, there exists a scoring rule implementing maximum effort.

Theorem 4 is more subtle than Theorem 3 because imperfect signals might require positive rewards in both states to implement optimal dynamic efforts. This creates additional

complexity: we cannot simply "flatten" rewards as in Theorem 3 without carefully tracking how these changes affect incentives across all beliefs.

Step One: Convexifying the no-information utility Our first step is to replace the arbitrary dynamic contract with one where the no-information utility is "convexified"—i.e., replacing the contract with one where u_t^N is convex in μ_t^N. While this property holds for any (static) scoring rule, this property need not hold for dynamic contracts.

We describe the argument. Define $\bar{t}$ to be the earliest time such that the no-information utility exhibits a nonconvexity. A nonconvexity in the no-information utility implies that $\bar{t} \leq \tau_{\mathcal{M}}$. One of our observations is that the dynamic incentives imply that the effort constraint cannot be binding at $\bar{t}+1$ under non-convexity.[20] Intuitively, the no-information utility at time t cannot be larger than the convex combination between (1) the reward of receiving a signal R immediately at time $t+1$; and (2) the continuation payoff at time $t+1$ for not receiving any Poisson signal at time t. If the effort constraint is binding at time $\bar{t}+1$, the latter equals the no information utility at time $\bar{t}+1$, which leads to a contradiction due to the non-convexity of the no information utility.

On the other hand, if the agent's incentive to exert effort is slack at time $\bar{t}+1$, we may raise some rewards in the menu option $r_{\bar{t}}^N$ to smooth the non-convex piece without violating earlier IC constraints. Repeating this argument pushes convexity forward through time until the entire u_t^N becomes convex.

Step Two: Identifying a static implementation of the scoring rule Since convex functions are equal to the upper envelope of the linear functions below them, a natural conjecture in light of Step One is that the principal could offer a static contract providing the set of reward functions that are tangent to u_t^N at some belief. This contract provides the same payoff if effort is not exerted; moreover, no higher rewards could be provided if Poisson signals arrive without increasing the no-information utility above u_t^N.

However, this argument does not work since the constructed rewards may lie outside of $[0,1]$ for an arbitrary no-information utility function; for instance, consider $u_t^N = (\mu_t^N)^2$. A dynamic contract that implements this no-information utility is 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 8, 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.

Intuitively, this kind of a violation of the constraint that rewards lie in $[0,1]$ arises because the no-information utility is too convex. In this case, we can flatten the no-information utility by decreasing the reward at earlier times—analogous to Step Two of Theorem 3. Un-

[20] See Figure 7 in Section B.3 for an illustration.

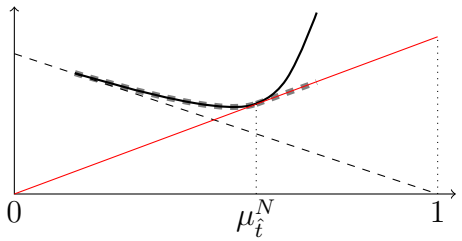

Figure 5: The black curve is the convex no-information utility of the agent, and $\hat{t}$ is the minimum time with a bounded tangent line (red line). The thick dashed line is the no-information utility of the agent given a feasible scoring rule that offers a menu option that corresponds to the red line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

der this flattening, the no-information utility decreases weakly more than the continuation payoff in all continuation games, yielding stronger incentives to exert effort. Figure 5 illustrates this idea, with details provided in Section B.3.

6 Insufficiency under Noisy Learning and Slow Drift

We now turn to environments with noisy signals and slow drift. We specialize to the binary-signal environment from Section 4. Noisy signals satisfy:

$$\lambda^R_1 > \lambda^R_0 > 0 \text{ and } 0 < \lambda^L_1 < \lambda^L_0,$$

i.e., learning is imperfect, and either signal may arise under either state. Slow drift means:

$$\lambda^R_0 + \lambda^L_0 \lessapprox \lambda^R_1 + \lambda^L_1$$

so that the belief moves gradually towards 0 when no Poisson signal arrives. Importantly, the single-signal assumption implies that, for a fixed arrival rate of R, beliefs drift at the maximal rate conditional on non-arrival. In this sense, the single-signal environment represents an extremal benchmark opposite to settings with slow drift.

Intuitively, increasing a late reward now affects rewards for exerting effort *less* than shirking incentives under noisy signals: e.g., if the reward for the state being 0 is increased, the agent may not enjoy this increase when exerting effort if he receives a right-biased signal R while the state is 0. Thus, the logic of Theorem 3 fails. Furthermore, since beliefs drift toward state 0, rewards for receiving a left-biased signal L can be lowered over time without incentivizing the agent to shirk and lie. As described in Section 3.4, these modifications are in the direction of those that maximize the incentives for the agent to exert effort.

To state our results, we first provide a formal bound on how long the horizon should be

so that the time horizon is not a binding constraint. We define this time as $T_{\lambda,D,c}$:

Lemma 4. *Fix any signal arrival rates satisfying $\lambda_1^R > \lambda_0^R$. Let $\mu_{\lambda,c} \triangleq \min\{\frac{1}{2}, \frac{c}{\lambda_1^R - \lambda_0^R}\}$ and let $T_{\lambda,D,c}$ be the maximum time such that $\mu^N_{T_{\lambda,D,c}-1} \geq \mu_{\lambda,c}$. The stopping time $\tau_{\mathcal{M}}$ satisfies $\tau_{\mathcal{M}} \leq T_{\lambda,D,c}$, given any prior $D \in (0,1)$, cost of effort c, and contract $\mathcal{M}$ with rewards belonging to $[0,1]$.*

Intuitively, $T_{\lambda,D,c}$ is the maximum calendar time the agent can be incentivized to exert effort in any contract when the prior is D. The following result provides sufficient conditions under which more complex dynamic structures are necessary in an effort-maximizing contract:

Theorem 5 (Strictly Less Effort Under Scoring Rules)**.**
Fix any prior $D \in (0, \frac{1}{2})$, any cost of effort c, and any constant $\kappa_0 > 0, \frac{1}{4} \geq \bar{\kappa}_1 > \underline{\kappa}_1 > 0$. There exists $\epsilon > 0$ such that for any λ satisfying:

- $\lambda_1^R - \lambda_0^R \geq \frac{1}{D}(c + \kappa_0)$; *(sufficient-incentive)*
- $\lambda_1^L, \lambda_0^L, \lambda_0^R, \lambda_1^R \in [\underline{\kappa}_1, \bar{\kappa}_1]$; *(noisy-signal)*
- $\lambda_1^R + \lambda_1^L \in (\lambda_0^R + \lambda_0^L, \lambda_0^R + \lambda_0^L + \epsilon)$, *(slow-drift)*
- *and $T \geq T_{\lambda,D,c}$,* *(sufficient-horizon)*

any static scoring rule implements effort strictly less than the maximum.

The first and last assumptions prevent trivial horizon or zero-incentive constraints. The substantive assumptions are noisy signals and slow drift. The former implies the technology is far enough from the case considered in Theorem 3, while the latter reflects the technology is far enough from the case considered in Theorem 4 as described above.

Our proof considers a particular contract that can outperform any static contract:

Definition 3 (Myopic-incentive Contract)**.**
When prior $D \in (0, \frac{1}{2})$, a contract $\mathcal{M}$ with menu options $\{r_t^s\}_{t \leq \tau_{\mathcal{M}}, s \in S} \cup \{r^N_{\tau_{\mathcal{M}}}\}$ is a myopic-incentive contract *if $r_t^R = (0,1)$ and $r_t^L = (\frac{\mu^N_{t-1}}{1-\mu^N_{t-1}}, 0)$ for any $t \geq 1$, and $r^N_{\tau_{\mathcal{M}}} = r^L_{\tau_{\mathcal{M}}}$.*

Intuitively, this contract offers the full reward when guessing state 1 and $\frac{\mu^N_{t-1}}{1-\mu^N_{t-1}}$ when guessing state 0. The rewards are carefully chosen such that an agent with belief μ^N_{t-1} is indifferent between these two reward options. Note that since μ^N_{t-1} decreases over time, so does the reward when the agent guesses state 0. This property is necessary for the constructed contract to be incentive-compatible.

The terminology 'myopic-incentive' reflects that this contract reoptimizes rewards period-by-period. The rewards $r_t^R = (0,1)$ and $r_t^L = (\frac{\mu_{t-1}^N}{1-\mu_{t-1}^N}, 0)$ form a V-shaped scoring rule with a kink at belief μ_{t-1}^N, which is the optimal scoring rule for the continuation game with the prior being μ_{t-1}^N. We mention that the myopic-incentive contract typically does *not* provide maximal incentives to exert effort in all continuation games—after all, rewards decrease over time, so an agent who exerts effort and receives a Poisson signal in later periods may receive a lower reward. However, Lemma 6 in the Online Appendix shows that these contracts nevertheless provide incentives *close* to the effort-maximizing contract under slow-drift. Intuitively, slow drift ensures that the aforementioned decreases in the rewards over time have a minor impact on the agent's incentives to exert effort at any belief.

Theorem 5 follows from noting that the same cannot be said for the effort-maximizing *static* contract. Letting $\mathcal{M}^*$ denote the effort-maximizing dynamic contract, for a contract to perform only slightly worse than $\mathcal{M}^*$, it is still necessary to provide incentives to exert effort at times close to $\tau_{\mathcal{M}^*}$. Section OA 4 illustrates that when signals are noisy, such static contracts are too weak to incentivize the agent to exert effort at time 0, leading to a contradiction. Simply put, dynamic contracts can adapt incentives to both the initial beliefs as well as the later beliefs; static contracts cannot.

We mention that the conditions in Theorem 5 facilitate a comparison in the amount of effort implemented by the myopic-incentive contract and an arbitrary (static) scoring rule. In particular, these conditions avoid the need to provide more structure to make this comparison sharp—for instance, to avoid integer issues which would suggest the increased strength from dynamics are insufficient to induce another period of effort. Nevertheless, the theorem delivers our main qualitative message: a sufficiently strong joint violation of the conditions supporting static elicitation can make dynamic elicitation necessary.

Along these lines, an assumption in Theorem 5, but one that appears relatively harmless, is that $D \in (0, \frac{1}{2})$. This restriction ensures the menus in the myopic-incentive contract remain within $[0,1]$. Combined with downward drift, this ensures feasibility and compatibility with truthful reporting in myopic incentive contracts.

7 Additional Observations and Final Thoughts

We conclude with some additional discussion, presenting our final thoughts and proposing several open questions in Section 7.4.

7.1 Properties of Optimal Scoring Rules Under Dynamics

The results in Section 5 provide conditions such that the optimal contract has a static implementation as scoring rules. However, the structure of the optimal scoring rules remains elusive. In this section, we show that for perfect-learning and single-signal environments, the dynamics in information acquisition still play a crucial role in determining the optimal scoring rules. In particular, we show that the optimal scoring rules in those two dynamic environments differ from those in static environments.

Perfect-learning In Theorem 3, we show that a V-shaped scoring rule implements the maximum effort. However, the effort-maximizing choice of the parameter r_1 is not provided. In this section, we show that the optimal choice of r_1 would not lead to a V-shaped scoring rule with a kink at the prior, resulting in a difference compared to the static environments.

To illustrate the idea, we consider an environment where T is sufficiently large so that the time horizon would not be a binding constraint for exerting costly effort. Given a V-shaped scoring rule P with parameters $r_0 = 1$ and $r_1 \in [0, 1]$, recall that u_t^N 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 μ_{t-1}^N: 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 μ_{t-1}^N, 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_t^N.[21] The agent has incentives to exert effort at time t given scoring rule P if and only if $\mu_t^N \in [\underline{\mu}(r_1), \bar{\mu}(r_1)]$. Figure 6 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 5. *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 5 implies that by increasing the reward r_1 for predicting 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). Therefore, the optimal choice of r_1 is

$$r_1^* = \max\{r_1 : \bar{\mu}(r_1) \geq D\}.$$

[21] More formally: Assuming that $U_t^+ \geq u_{t-1}^N$ for some $t \geq 0$, we define $\underline{\mu}(r_1) = \min\{\mu_t^N, t \geq 0 : U_t^+ \geq u_{t-1}^N\}$ and $\bar{\mu}(r_1) = \max\{\mu_t^N, t \geq 0 : U_t^+ \geq u_{t-1}^N\}$. If $U_t^+ < u_{t-1}^N$ holds for all t, then the agent does not work and we take $\underline{\mu}(r_1) = \overline{\mu}(r_1)$ to be undefined.

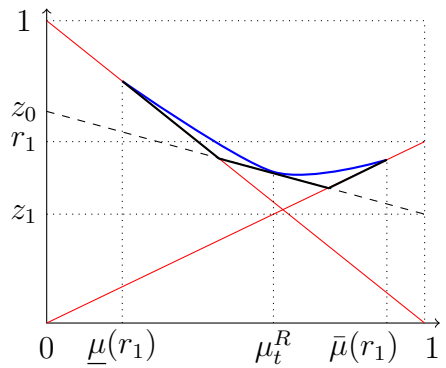

Figure 6: The illustration for both perfect-learning and single-signal environments. The red lines are the agent's utility u_{t-1}^N for not exerting effort and the blue curve is the agent's utility U_t^+ from exerting effort in at least one period, both as a function of the no information belief μ_{t-1}^N. 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).

Naturally, the kink of the V-shaped scoring rule induced by r_1^* would not be located at the prior D, as illustrated in Figure 6. Our analysis also highlights the economic intuition behind this difference. In dynamic environments, the principal needs to balance the incentives between earlier and later periods. By skewing the kink of the V-shaped scoring rule away from the prior, the agent faces a stronger incentive to exert effort in future periods.

Finally, our observation here also illustrates an interesting dynamic effect in our model: 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.

Single-signal A notable feature in the single-signal environment is that, 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 decreases the bonus reward for the agent since total rewards are constrained to the unit interval. A smaller bonus may subsequently lower the agent's incentive to exert effort.

However, 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 addition of the positive base reward leads the agent to anticipate higher rewards from exerting effort in cases where the terminal belief is in an intermediate range. 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 $t \in (0, \tau_{\mathcal{M}})$. As illustrated in Figure 6, 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 $\tau_{\mathcal{M}}$ with belief $\mu^N_{\tau_{\mathcal{M}}} = \underline{\mu}(r_1)$. In this case, since the signals are not perfectly revealing, there may exist a time t such that the agent's belief for receiving a right biased signal is $\mu^R_t \in (\underline{\mu}(r_1), \bar{\mu}(r_1))$. 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^R_t (e.g., the additional menu option (z_0, z_1) illustrated in Figure 6), 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 $\underline{\mu}(r_1)$.

7.2 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, 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 5—in the sense that given any dynamic contract $\mathcal{M}$ and any best response of the agent, there exists another contract $\widehat{\mathcal{M}}$ that can be implemented as a scoring rule, and the agent's best response in contract $\widehat{\mathcal{M}}$ first order stochastic dominates his best response in contract $\mathcal{M}$.[22] Therefore, the information acquired under contract $\widehat{\mathcal{M}}$ is always weakly Blackwell more informative compared to the information acquired under contract $\mathcal{M}$.

An important implication of extending our results to non-invariant costs is that they also apply to settings where the agent discounts future payoffs. Essentially, when the agent discounts future payoffs, it corresponds exactly to exponentially decreasing cost functions, as payments are only made after the state realization after T. Consequently, all our main results, such as the optimality of scoring rules in perfect-learning or single-signal environ-

[22] Let $z_t, \hat{z}_t \in \{0, 1\}$ be the effort decision of the agent given contracts $\mathcal{M}, \widehat{\mathcal{M}}$ respectively conditional on not receiving any Poisson signal before t. We can show that given any time t, $\sum_{i \leq t} z_i \leq \sum_{i \leq t} \hat{z}_i$.

ments, extend naturally to this setting.

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 right-biased signal that arrives in state θ with probability $\lambda^R_{\theta,\tilde{t}}$, and a left-biased signal with probability $\lambda^L_{\theta,\tilde{t}}$ (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 do 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 4 requires μ^R_t to be weakly monotone as a function of time (a property that holds when the arrival rate is constant). Otherwise, as long as the parameters stay within each environment articulated in Section 5, the proofs of these results extend unchanged.

7.3 Value of Dynamic Contracts

Our analysis compares contracts in terms of Blackwell informativeness. This is useful because it separates the incentive problem from any particular downstream application. At the same time, the payoff value of the additional informativeness generated by unrestricted advice depends on the decision environment. A small informational improvement may have little value in one application and a large value in another.

This point is particularly stark under the conditions of Theorem 5. Let τ^{S} denote the largest stopping time implementable by a scoring rule. By Theorem 5, unrestricted advice can induce strictly more effort, so in particular the agent can be induced to work through time $\tau^{\mathrm{S}} + 1$. Since beliefs drift toward state 0 absent signal arrival and L is left-biased, the posterior μ^L_t is strictly decreasing in t. Hence $\mu^L_{\tau^{\mathrm{S}}+1} < \mu^L_{\tau^{\mathrm{S}}}$.

Now choose any cutoff

$$q \in \left(\mu^L_{\tau^{\mathrm{S}}+1}, \mu^L_{\tau^{\mathrm{S}}}\right),$$

and consider a binary decision problem with actions a_0 and a_1 such that a_0 is optimal if and only if the posterior belief that $\theta = 1$ is at most q. For example, one can take payoffs

$$v(a_1, 0) = v(a_1, 1) = 0, \qquad v(a_0, 0) = 1, \qquad v(a_0, 1) = -\frac{1-q}{q}.$$

Since $q < \mu^L_{\tau^{\mathrm{S}}} < D$, the decision maker chooses a_1 without additional information. Under any scoring rule, the posterior never falls below $\mu^L_{\tau^{\mathrm{S}}}$, so summarized advice never changes the

decision and has value 0. Under unrestricted advice, however, a left-biased signal arriving at time $\tau^{\mathrm{S}} + 1$ occurs with strictly positive probability, and on that event, the posterior falls below q. The decision maker then switches to a_0, so unrestricted advice has strictly positive value.

Thus, the additional value of dynamic contracts depends on the decision environment. Under slow drift, the gap between $\mu^L_{\tau^{\mathrm{S}}}$ and $\mu^L_{\tau^{\mathrm{S}}+1}$ can be made arbitrarily small. However, a slight increase in informativeness can change the value of information from 0 to strictly positive. In this sense, the multiplicative gain from unrestricted advice can be unbounded.

7.4 Final Remarks

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 forecasters seeking a particular sought-after piece of falsifiable evidence, and under a class of contracts in line with past work on forecaster incentive provision.

We have shown that whether the decision maker benefits from a contracting environment that facilitates dynamic reporting depends on the nature of the dynamic learning process. Along the way, we discussed how the relevant properties of the information acquisition technology have natural interpretations in various settings of practical interest. 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 under which maximum effort is implementable by a scoring rule and identified conditions under which dynamic reporting is strictly necessary.

There are many natural avenues for future work. Empirically, our results show how learning technologies influence the benefits of time variation to rewards. While we focus on a simple forecasting problem, such questions may be of interest when eliciting other kinds of information beyond a prediction of a future event—for example, if the learning process itself is privately known by the expert (Chambers and Lambert, 2021). 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.

References

Anthony, R. N., Govindarajan, V., Hartmann, F. G., Kraus, K., and Nilsson, G. (2007). *Management Control Systems*, volume 12. McGraw-Hill, Boston.

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.

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

Bloedel, A. and Segal, I. (2024). The costly wisdom of inattentive crowds. working paper.

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 under partial knowledge. *arXiv preprint arXiv:2107.07420*.

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. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 514–514.

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. (2026). Indirect persuasion. *forthcoming at Journal of Political Economy*.

Elliott, G. and Timmermann, A. (2016). Forecasting in economics and finance. *Annual Review of Economics*, 8:81–110.

Frongillo, R. and Waggoner, B. (2024). Recent trends in information elicitation. *ACM SIGecom Exchanges*, 22(1):122–134.

Fu, W. (2024). Optimal reward scheme in private exploration with conclusive news. Working paper.

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

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.

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.

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. *Working paper*.

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.

Marinovic, I., Ottaviani, M., and Sørensen, P. (2013). Forecasters' objectives and strategies. In Elliott, G., Granger, C., and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2, Part B, pages 690–720. Elsevier.

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.

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.

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

Strulovici, B. (2022). Can society function without ethical agents? an informational perspective. working paper.

Szalay, D. (2005). The economics of clear advice and extreme options. *The 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.

Zellner, M., Abbas, A. E., Budescu, D. V., and Galstyan, A. (2021). A survey of human judgement and quantitative forecasting methods. *Royal Society open science*, 8(2):201187.

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

A Effort-Maximizing Dynamic Contracts

Proof of Theorem 1. By Lemma 3, it is without loss to work with a shrinking menu representation. Because menus are shrinking, any reward used on path at date t is also available at every earlier date, so it must satisfy

$$u(\mu_{t'}^N, r) \leq u_{t'}^N, \qquad \forall t' \leq t. \tag{5}$$

For each t, let $\hat{r}_t^N$ maximize r_0 subject to

$$r = (r_0, r_1) \in [0,1]^2 \qquad \text{and} \qquad u(\mu_t^N, r) = u_t^N.$$

Then $\hat{r}_{t,0}^N = 1$ whenever $\hat{r}_{t,1}^N > 0$, so each canonical reward is either of the form $(1, b_t)$ or $(a_t, 0)$. Moreover, $\hat{r}_t^N$ is obtained from r_t^N by shifting reward toward state 0 while keeping utility at μ_t^N fixed, hence

$$u(\mu, \hat{r}_t^N) \leq u(\mu, r_t^N) \qquad \forall \mu \geq \mu_t^N.$$

Therefore, adding $\hat{r}_t^N$ to all earlier menus does not increase the stopping utility of any earlier no-information type. Doing this for every t preserves effort maximization, so without loss we may take the no-information rewards themselves to be canonical.

Now fix $t \leq t'$. Since $r_{t'}^N \in \mathcal{M}_t$,

$$u(\mu_t^N, r_t^N) \geq u(\mu_t^N, r_{t'}^N). \tag{A.1}$$

Because both rewards are canonical, (A.1) implies $r^N_{t'} \leq r^N_t$ coordinatewise: if $r^N_t = (1, b_t)$, then any later canonical reward is either $(1, b_{t'})$ with $b_{t'} \leq b_t$ or $(a_{t'}, 0)$; if $r^N_t = (a_t, 0)$, then (A.1) rules out any later reward of the form $(1, b)$, so necessarily $r^N_{t'} = (a_{t'}, 0)$ with $a_{t'} \leq a_t$. This proves Part 1.

Next fix $t \leq \tau_{\mathcal{M}}$ and $s \in S$. By (5), any reward used after signal s at date t must satisfy

$$u(\mu^N_{t'}, r) \leq u^N_{t'}, \qquad \forall t' \in \{0, \ldots, t\}.$$

Conversely, adding at date t any reward satisfying these inequalities does not increase the utility from stopping at any earlier no-information belief. Hence, without loss, r^s_t solves

$$\max_{r:\Theta\to[0,1]} u(\mu^s_t, r) \qquad \text{s.t.} \qquad u(\mu^N_{t'}, r) \leq u^N_{t'} \quad \forall t' \in \{0, \ldots, t\}. \tag{A.2}$$

If s is left-biased, then $\mu^s_t \leq \mu^N_t$. In the relaxed version of (A.2) that keeps only the constraint at t, the objective is maximized by shifting as much reward as possible toward state 0. By construction, this is exactly the canonical reward r^N_t. Since r^N_t also satisfies all earlier constraints, it solves (A.2), proving Part 2.

If s is right-biased, then by definition of G_t,

$$u(\mu^s_t, r^s_t) = G_t(\mu^s_t).$$

Each feasible reward induces an affine function $\mu \mapsto u(\mu, r)$ lying below the no-information points $\{(\mu^N_{t'}, u^N_{t'})\}^t_{t'=0}$. Hence G_t is the pointwise supremum of such affine functions, i.e., the greatest convex minorant generated by feasible rewards bounded in $[0, 1]^2$. Since the affine function induced by r^s_t attains this supremum at μ^s_t, it is a supporting line to G_t at that belief. This proves Part 3. □

B Effort-Maximizing Contracts as Scoring Rules

B.1 Stationary Environment

Proof of Theorem 2. For any contract $\mathcal{M}$ with stopping time $\tau_{\mathcal{M}} \in [0, T]$, to show that there exists a static scoring rule P such that the agent has incentive to exert effort at least until time $\tau_{\mathcal{M}}$ given static scoring rule P, it is sufficient to show that there exists a static scoring rule P such that the agent has an incentive to exert effort at any continuation game $\mathcal{G}_t$ for any $t \in [0, \tau_{\mathcal{M}}]$.

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 $t' \in [t, \tau_{\mathcal{M}}]$, we can allow the agent to pick any menu option from time t to $\tau_{\mathcal{M}}$. 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 a kink at prior D. Since the optimal scoring rule remains the same for all continuation games in stationary environments, this scoring rule implements the optimal dynamic efforts. □

B.2 Perfect-learning Environment

Proof of Theorem 3. As alluded to in the main text, the proof follows two steps:

Step One: We first show that, for any prior D and any signal arrival probabilities λ, there exists an effort-maximizing contract $\mathcal{M}$ with a sequence of menu options $\{r_t^s\}_{t\leq\tau_{\mathcal{M}},s\in S} \cup \{r_{\tau_{\mathcal{M}}}^N\}$ such that $r_t^L = r_{\tau_{\mathcal{M}}}^N = (1,0)$ for any $t \leq \tau_{\mathcal{M}}$.

For any contract $\mathcal{M}$, by applying the menu representation in Lemma 3, let $\{r_t^s\}_{t\leq\tau_{\mathcal{M}},s\in S}$ be the set of menu options for receiving Poisson signals and let $r_{\tau_{\mathcal{M}}}^N = (z_0, z_1)$ be the menu option for not receiving any Poisson signal before the stopping time $\tau_{\mathcal{M}}$. Note that in the perfect-learning environment, it is without loss to assume that $r_{t,1}^L = 0$ for any $t \leq \tau_{\mathcal{M}}$ since the posterior probability of state 1 is 0 after receiving a Poisson signal L. Now consider another contract $\widehat{\mathcal{M}}$ with menu options $\{\hat{r}_t^s\}_{t\leq\tau_{\mathcal{M}},s\in S}$ and $(\hat{z}_0, \hat{z}_1)$, where

$$(\hat{z}_0, \hat{z}_1) = \underset{z,z'\in[0,1]}{\arg\max}\ z \quad \text{s.t.} \quad \mu_{\tau_{\mathcal{M}}}^N z' + (1-\mu_{\tau_{\mathcal{M}}}^N)z = u_{\tau_{\mathcal{M}}}^N, \tag{6}$$

and for any time $t \leq \tau_{\mathcal{M}}$ and any signal $s \in S$,

$$\hat{r}_t^s = \begin{cases} r_t^s & u(\mu_t^s, r_t^s) \geq u(\mu_t^s, (\hat{z}_0, \hat{z}_1)) \\ (\hat{z}_0, \hat{z}_1) & \text{otherwise.} \end{cases}$$

Essentially, contract $\widehat{\mathcal{M}}$ adjusts the reward function for no information belief $\mu_{\tau_{\mathcal{M}}}^N$ 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 $t \leq \tau_{\mathcal{M}}$, contract $\widehat{\mathcal{M}}$ allows the agent to optionally choose the additional 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 $s \in S$ and any time $t \leq \tau_{\mathcal{M}}$, the expected utility of the agent for receiving an informative signal s is weakly higher, and hence, at any time $t \leq \tau_{\mathcal{M}}$, the continuation payoff of the agent for exerting effort until time $\tau_{\mathcal{M}}$ weakly

increases in contract $\widehat{\mathcal{M}}$. Moreover, at any time t, the expected reward of the agent for not exerting effort at time t with belief μ^N_{t-1} satisfies $\hat{u}^N_{t-1} \le u^N_{t-1}$. This is because, by our construction, at any time $t \le \tau_{\mathcal{M}}$, fewer options are available to the agent in contract $\widehat{\mathcal{M}}$, except for the additional option of $(\hat{z}_0, \hat{z}_1)$. However, $u(\mu^N_{t-1}, (z_0, z_1)) \ge u(\mu^N_{t-1}, (\hat{z}_0, \hat{z}_1))$ since $\mu^N_{t-1} \ge \mu^N_{\tau_{\mathcal{M}}}$ and both options (z_0, z_1) and $(\hat{z}_0, \hat{z}_1)$ give the same expected reward for the posterior belief of $\mu^N_{\tau_{\mathcal{M}}}$. Combining both observations, we have $\tau_{\widehat{\mathcal{M}}} \ge \tau_{\mathcal{M}}$ and $\widehat{\mathcal{M}}$ is also an effort-maximizing contract.

Note that in optimization program (6), 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^L_t, we must have $\hat{r}^L_{t,0} = 1$ as well. Therefore, the agent receives the maximum reward of 1 whenever he receives a left-biased signal. In this case, incentive compatibility implies that $\hat{z}_1 \le r^R_{t,1}$ for all $t \le \tau_{\mathcal{M}}$. Therefore, we can also decrease $\hat{z}_1$ and $r^R_{t,1}$ for all $t \le \tau_{\mathcal{M}}$ by $\hat{z}_1$, which does not affect the agent's incentive for effort and hence the effort-maximizing contract satisfies that $r^L_t = r^N_{\tau_{\mathcal{M}}} = (1, 0)$ for any $t \le \tau_{\mathcal{M}}$.

Next, we will focus on the case when $\hat{z}_0 < 1$ and hence $\hat{z}_1 = 0$. Now consider another contract $\bar{\mathcal{M}}$ with menu options $\{\bar{r}^s_t\}_{t \le \tau_{\widehat{\mathcal{M}}}, s \in S}$ and $(\bar{z}_0, \bar{z}_1)$, where $(\bar{z}_0, \bar{z}_1) = (1, 0)$ and for any time $t \le \tau_{\mathcal{M}}$, $\bar{r}^R_t = \hat{r}^R_t$ and $\bar{r}^L_t = (1, 0)$. We show that this weakly improves the agent's incentive to exert effort until time $\tau_{\widehat{\mathcal{M}}}$ for any $t \le \tau_{\widehat{\mathcal{M}}}$. Specifically, for any $t \le \tau_{\widehat{\mathcal{M}}}$, the increase in the no information payoff for the continuation game $\mathcal{G}_t$ is

$$\bar{u}^N_{t-1} - \hat{u}^N_{t-1} \le (1 - \mu^N_{t-1})(1 - \hat{r}^L_{t-1,0}).$$

This is because in contract $\bar{\mathcal{M}}$, either the agent prefers the menu option $r^R_{t'}$ for some $t' \ge t-1$, 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 $(1 - \mu^N_{t-1})(1 - \hat{r}^L_{t-1,0})$ since one feasible option for the agent in contract $\widehat{\mathcal{M}}$ is $\hat{r}^L_{t-1}$ with an expected reward at least $(1 - \mu^N_{t-1})\hat{r}^L_{t-1,0}$. Moreover, for any time $t \le \tau_{\widehat{\mathcal{M}}}$, the increase in continuation payoff for exerting effort from t until $\tau_{\widehat{\mathcal{M}}}$ is at least $(1 - \mu^N_{t-1})(1 - \hat{r}^L_{t,0})$. This is because incentive constraints imply that $\hat{r}^L_{t,0}$ must decrease as t increases, which is due to the fact that in the perfect-learning environment, the posterior belief μ^L_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 $\bar{\mathcal{M}}$ and the reward of the agent is at most $\hat{r}^L_{t,0}$ in contract $\widehat{\mathcal{M}}$, implying that the difference in expected reward is at least $(1 - \mu^N_{t-1})(1 - \hat{r}^L_{t,0})$. Combining the above observations and observing that $(1 - \mu^N_{t-1})(1 - \hat{r}^L_{t,0}) \ge (1 - \mu^N_{t-1})(1 - \hat{r}^L_{t-1,0})$, we have $\tau_{\bar{\mathcal{M}}} \ge \tau_{\widehat{\mathcal{M}}}$, and hence $\bar{\mathcal{M}}$ is also effort-maximizing.

Step Two: We now show that we can replace all other menu options except $(1, 0)$ with a

single menu option that the agent can select at any time. By the previous step, it suffices to focus on contract $\mathcal{M}$ with a sequence of menu options $\{r_t^s\}_{t\leq\tau_{\mathcal{M}},s\in S}\cup\{r_{\tau_{\mathcal{M}}}^N\}$ such that $r_t^L = r_{\tau_{\mathcal{M}}}^N = (1,0)$ for any $t \leq \tau_{\mathcal{M}}$. In addition, since signals are perfectly revealing, it is without loss to assume that $r_{t,0}^R = 0$ and the incentive constraints imply that $r_{t,1}^R$ is weakly decreasing in t.

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

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

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

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

Proof of Lemma 5. 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 $0 \leq z' < z \leq 1$. 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 μ_{t-1}^N, the agent's utility

for not exerting effort given menu option $(0, r_1)$ decreases by $\mu_{t-1}^N(z - z')$, while the agent's utility for exerting effort in at least one period decreases by at most $\mu_{t-1}^N(z - z')$ since the reward decrease can only occur when the state is 1. Therefore, $\bar{\mu}(r_1)$ also weakly increases. □

B.3 Single-signal Environment

Proof of Theorem 4. As mentioned in the main text, the proof proceeds in two steps:

Step One: We first show that an effort-maximizing contract exists with the no-information utility u_t^N convex in μ_t^N. For any contract $\mathcal{M}$, let $\underline{u}_t(\mu)$ be the convex hull of the no information payoff $u_{t'}^N$ for $t' \leq t$ by viewing $u_{t'}^N$ as a function of $\mu_{t'}^N$. Consider an effort-maximizing contract $\mathcal{M}$ with the following selection:

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

Note that the exponential weight e^{-i} is purely for the tie breaking selection and ensures a finite sum of the no information payoff. The existence of an effort-maximizing contract given such a 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 $\tau_{\mathcal{M}}$ be the stopping time of the agent for contract $\mathcal{M}$. We will show that $\bar{t} = \tau_{\mathcal{M}} + 1$.

First, we observe that it cannot be the case that $\bar{t} = \tau_{\mathcal{M}}$. This is because the agent does not have incentives to exert effort after time $\tau_{\mathcal{M}}$. By increasing the agent's utility $u_{\bar{t}}^N$ in this case, we can restore the convexity of u_t^N at $t = \bar{t}$ without violating the effort incentives.

Now, suppose by contradiction we have $\bar{t} \leq \tau_{\mathcal{M}} - 1$. At any time $t \leq \tau_{\mathcal{M}}$, recall that $\mathcal{G}_t$ is the continuation game at time t with prior belief μ_{t-1}^N such that the agent's utility for not exerting effort is u_{t-1}^N and the agent's utility for exerting effort in $\mathcal{G}_t$ is

$$U_t \triangleq \left(1 - \sum_{s \in S} F_t^s(\tau_{\mathcal{M}})\right) \cdot u(\mu_{\tau_{\mathcal{M}}}^N, r_{\tau_{\mathcal{M}}}^N) + \sum_{t'=t}^{\tau_{\mathcal{M}}} \sum_{s \in S} f_t^s(t') \cdot u(\mu_{t'}^s, r_{t'}^s).$$

Note that $U_t \geq u_{t-1}^N$ for any $t \leq \tau_{\mathcal{M}}$. We first show that the equality must hold at time $\bar{t} + 1$, i.e., $U_{\bar{t}+1} = u_{\bar{t}}^N$. Let $\bar{u}_t(\mu)$ be the upper bound on the expected reward at any belief

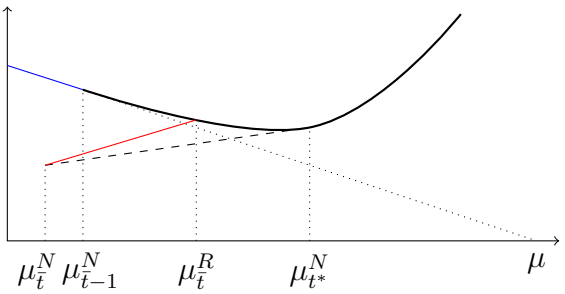

Figure 7: This figure illustrates the case when $\mu_{\bar{t}}^R \leq \mu_{t^*}^N$. The black curve is the function $\underline{u}_{\bar{t}}(\mu)$, blue line is the function $\bar{u}(\mu)$ and the red line is the function $y(\mu)$.

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 $t = \bar{t} - 1$ 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 $\bar{u}_{\bar{t}-1}(\mu_{\bar{t}}^N) > u_{\bar{t}}^N$. In this case, if $U_{\bar{t}+1} > u_{\bar{t}}^N$, by increasing $u_{\bar{t}}^N$ to $\min\{U_{\bar{t}+1}, \bar{u}_{\bar{t}-1}(\mu_{\bar{t}}^N)\}$, the agent's incentive 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 $\bar{t}+1$ is binding.

By the construction of $\mathcal{M}$, 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 its convex hull. See Figure 7 for an illustration. We consider two cases separately.

- $\mu_{t^*}^N \geq \mu_{\bar{t}}^R$. 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 $y(\mu_{\bar{t}}^R) = \underline{u}(\mu_{\bar{t}}^R)$. Function y is illustrated in Figure 7 as the red line. Note that in this case, we have $y(\mu_{\bar{t}-1}^N) < \underline{u}_{\bar{t}}(\mu_{\bar{t}-1}^N) = u_{\bar{t}-1}^N$. Moreover, $y(\mu_{\bar{t}-1}^N)$ is the maximum continuation payoff of the agent for exerting effort at time $\bar{t}$ given belief $\mu_{\bar{t}-1}^N$. This is because, by exerting effort, either the agent receives a Poisson signal R at time $\bar{t}$, which leads to posterior belief $\mu_{\bar{t}}^R$ with expected payoff $\underline{u}(\mu_{\bar{t}}^R) = y(\mu_{\bar{t}}^R)$, or the agent does not receive a Poisson signal, which leads to belief drift to $\mu_{\bar{t}}^N$, with the optimal continuation payoff being $U_{\bar{t}+1} = u_{\bar{t}}^N = y(\mu_{\bar{t}}^N)$. However, $y(\mu_{\bar{t}-1}^N) < u_{\bar{t}-1}^N$ implies that the agent has a strict incentive not to exert effort at time $\bar{t}$, a contradiction.

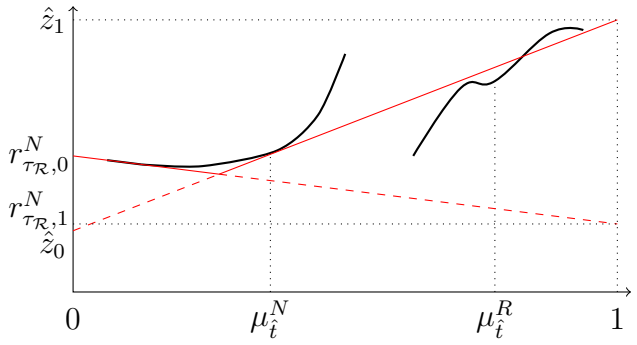

Figure 8: The black solid curves are the agent's expected utilities for not exerting effort as a function of his belief at any time t. The left curve is the expected utility for beliefs without receiving Poisson signals, and the right curve is the one receiving the Poisson signal R.

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

Therefore, we have $\bar{t} = \tau_{\mathcal{M}} + 1$ and the no information utility of the agent is a convex function.

Step Two: We now show that we can "flatten" rewards to strengthen the agent's incentives to exert effort in case the reward constraint is violated. By Step One, there exists a contract $\mathcal{M}$ with a sequence of menu options $\{r^s_t\}_{s\in S, t\leq\tau_{\mathcal{M}}} \cup \{r^N_{\tau_{\mathcal{M}}}\}$ in which the no information payoff is convex in the no information belief. If $\mu^N_t < u^N_t$ for all $t \leq \tau_{\mathcal{M}}$, 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 $t \leq \tau_{\mathcal{M}}$. 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 $t \leq \tau_{\mathcal{M}}$. Essentially, the straight line $(\hat{z}_0, \hat{z}_1)$ is tangent to the agent's utility curve for not receiving Poisson signals subject to the reward bound. Let $\hat{t}$ be the time corresponding to the rightmost tangent point. See Figure 8 for an illustration.

Let $\underline{u}(\mu)$ be the function that coincides with u^N_t for $\mu \leq \mu^N_{\hat{t}}$ and $\underline{u}(\mu) = (\hat{z}_1 - \hat{z}_0)\mu + \hat{z}_0$. Note that $\underline{u}$ is convex. Consider another contract $\widehat{\mathcal{M}}$ that is implemented by the 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 $\tau_{\widehat{\mathcal{M}}} \geq \tau_{\mathcal{M}}$ and hence contract $\widehat{\mathcal{M}}$ must also be effort-maximizing, which concludes the proof of Theorem 4.

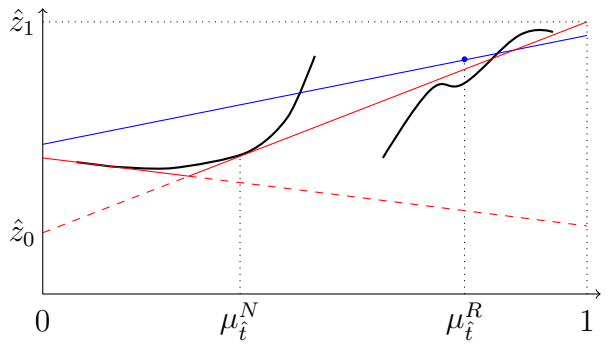

Figure 9: The blue line is the expected utility of the agent for choosing the menu option $(r_{t,0}^R, r_{t,1}^R)$ if the utility at belief $\mu_{\hat{t}}^R$ is higher than the red line.

In any continuation game $\mathcal{G}_t$, recall that u_{t-1}^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 $\mathcal{M}$. For any time $t \leq \tau_{\mathcal{M}}$, the agent has incentive to exert effort at time t given contract $\mathcal{M}$ implies that $u_{t-1}^N \leq U_t$. Given contract $\widehat{\mathcal{M}}$, we similarly define $\hat{u}_{t-1}^N$ and $\hat{U}_t$ and show that for any time $t \leq \tau_{\mathcal{M}}$, $U_t - \hat{U}_t \leq u_{t-1}^N - \hat{u}_{t-1}^N$. This immediately implies that the agent also has incentive to exert effort at any time $t \leq \tau_{\mathcal{M}}$ given contract $\widehat{\mathcal{M}}$ and hence $\tau_{\widehat{\mathcal{M}}} \geq \tau_{\mathcal{M}}$.

Our analysis for showing that $U_t - \hat{U}_t \leq u_{t-1}^N - \hat{u}_{t-1}^N$ is divided into two cases.

Case 1: $t \geq \hat{t}$. In this case, since $u_{t-1}^N \geq \hat{u}_{t-1}^N$ for any $t \leq \tau_{\mathcal{M}}$ by the construction of contract $\widehat{\mathcal{M}}$, it is sufficient to show that $\hat{U}_t \geq U_t$ for any $t \in [\hat{t}, \tau_{\mathcal{M}}]$. We first show that for any $t \in [\hat{t}, \tau_{\mathcal{M}}]$, if $\mu_t^R \leq \mu_{\hat{t}}^N$, we must have $\underline{u}(\mu_t^R) \geq u_t^R$ in order to satisfy the dynamic incentive constraint in contract $\mathcal{M}$. Next, we focus on the case where $\mu_t^R > \mu_{\hat{t}}^N$ and show that $\hat{u}_t^R = \mu_t^R \hat{z}_1 + (1 - \mu_t^R)\hat{z}_0 \geq u_t^R$. We prove this by contradiction. Suppose that $u_t^R > \mu_t^R \hat{z}_1 + (1 - \mu_t^R)\hat{z}_0$. Recall that $(r_{t,0}^R, r_{t,1}^R)$ are the options offered to the agent at time t that attain expected utility u_t^R under belief μ_t^R. Moreover, in our construction, either $\hat{z}_0 = 0$, or $\hat{z}_1 = 1$, or both equalities hold. Therefore, the bounded constraints $r_{t,0}^R, r_{t,1}^R \in [0, 1]$ and the fact that agent with belief μ_t^R prefers $(r_{t,0}^R, r_{t,1}^R)$ over $(\hat{z}_0, \hat{z}_1)$ imply that $r_{t,0}^R \geq \hat{z}_0$. See Figure 9 for an illustration. Since $\mu_{\hat{t}}^N < \mu_t^R$, this implies that the agent's utility at belief $\mu_{\hat{t}}^N$ given option $(r_{t,0}^R, r_{t,1}^R)$ is strictly larger than his utility under $(\hat{z}_0, \hat{z}_1)$, i.e.,

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

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

Finally, for $t \in [\hat{t}, \tau_{\mathcal{M}}]$, conditional on the event that the informative signal did not arrive at any time before t, since the agent expected utility given contract $\widehat{\mathcal{M}}$ is weakly higher compared to contract $\mathcal{M}$ 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 $\tau_{\mathcal{M}}$ 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 $u_t^N = \mu_t^N \hat{r}_t + (1 - \mu_t^N)\hat{z}_0$. Note that $\hat{r}_t \geq \hat{z}_1$ and it is possible that $\hat{r}_t \geq 1$. The construction of $\hat{r}_t$ is only used in the intermediate analysis, not in the constructed scoring rules. Let $\tilde{u}_t(\mu) \triangleq \mu \hat{r}_t + (1 - \mu)\hat{z}_0$ be the expected utility of the agent for choosing option $(\hat{z}_0, \hat{r}_t)$ given belief μ. This is illustrated in Fig. 10.

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

$$u_{t-1}^N - \hat{u}_{t-1}^N = \tilde{u}_{t-1}(\mu_{t-1}^N) - \hat{u}_{t-1}^N = \mu_{t-1}^N(\hat{r}_{t-1} - \hat{z}_1).$$

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

$$\begin{aligned} U_t - \hat{U}_t &= \sum_{t'=t}^{\tau_{\mathcal{M}}} \left(u_{t'}^R - \hat{u}_{t'}^R\right) \cdot f_t^R(t') \leq \sum_{t'=t}^{\tau_{\mathcal{M}}} \left(\tilde{u}_t(\mu_{t'}^R) - \hat{u}_{t'}^R\right) \cdot f_t^R(t') \\ &= (\hat{r}_t - \hat{z}_1) \cdot \sum_{t'=t}^{\tau_{\mathcal{M}}} \mu_{t'}^R \cdot f_t^R(t') \leq (\hat{r}_t - \hat{z}_1) \cdot \mu_{t-1}^N \leq (\hat{r}_{t-1} - \hat{z}_1) \cdot \mu_{t-1}^N \end{aligned}$$

where the second inequality holds by Bayesian plausibility and the last inequality holds since the no information belief drifts towards state 0. Combining the inequalities, we have $U_t - \hat{U}_t \leq u_{t-1}^N - \hat{u}_{t-1}^N$.

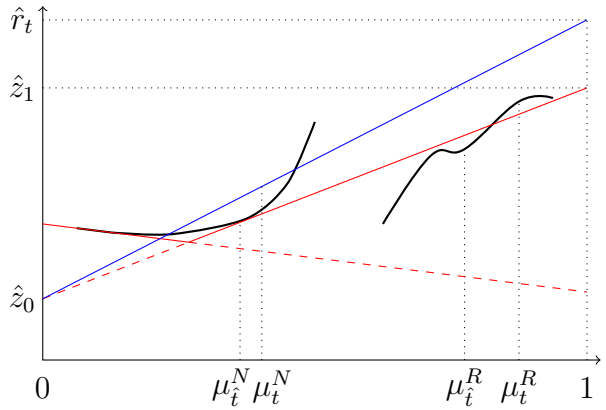

Figure 10: The blue line is the utility function $\tilde{u}_t$, which serves as an upper bound on the utility $u^R_{t'}$ for any $t' \geq t$.

Combining the above two cases, we have $U_t - \hat{U}_t \leq u^N_{t-1} - \hat{u}^N_{t-1}$ for any $t \leq \tau_{\mathcal{M}}$. Since the agent's optimal effort strategy is to stop at time $\tau_{\mathcal{M}}$ given contract $\mathcal{M}$, this implies that at any time $t \leq \tau_{\mathcal{M}}$, if the agent has not received any informative signal by time t, the agent also has incentive to exert effort until time $\tau_{\mathcal{M}}$ given contract $\widehat{\mathcal{M}}$ that can be implemented as a scoring rule. □

Online Appendix for "Incentivizing Forecasters to Learn: Summarized vs. Unrestricted Advice"

Yingkai Li and Jonathan Libgober

OA 1 Optimality of Dynamic Contracts

Proof of Lemma 4. 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 $\mathcal{M}$ such that $\tau_{\mathcal{M}} > T_{\lambda,D,c}$. The prior belief in the continuation game $\mathcal{G}_{\tau_{\mathcal{M}}}$ is $\mu^N_{\tau_{\mathcal{M}}-1} < \mu_{\lambda,c} \leq \frac{1}{2}$. By the definition of $\tau_{\mathcal{M}}$, the agent's optimal strategy is to exert effort for one period given contract $\mathcal{M}$. This implies that the agent has incentive to exert effort in continuation game $\mathcal{G}_{\tau_{\mathcal{M}}}$ given the effort-maximizing scoring rule for $\mathcal{G}_{\tau_{\mathcal{M}}}$. By Proposition 1, the effort-maximizing scoring rule for $\mathcal{G}_{\tau_{\mathcal{M}}}$ is the V-shaped scoring rule P with kink at $\mu^N_{\tau_{\mathcal{M}}-1}$. By simple algebraic calculation, the expected utility increase given scoring rule P for exerting effort in $\mathcal{G}_{\tau_{\mathcal{M}}}$ is $\mu^N_{\tau_{\mathcal{M}}-1}(\lambda^R_1 - \lambda^R_0)$, which must be at least the cost of effort c. However, this violates the assumption that $\mu^N_{\tau_{\mathcal{M}}-1} < \mu_{\lambda,c}$, a contradiction. □

Lemma 6 (Approximate Effort Maximization of Myopic-incentive Contracts)**.**
Given any prior $D \in (0, \frac{1}{2})$, any cost of effort c, any constant $\kappa_0 > 0$, and any $\eta > 0$, there exists $\epsilon > 0$ such that for any $T \geq T_{\lambda,D,c}$ and any λ satisfying that $\lambda^s_\theta \leq \frac{1}{4}$ for all $s \in \{L, R\}$ and $\theta \in \{0, 1\}$, and

- $\lambda^R_1 - \lambda^R_0 \geq \frac{1}{D}(c + \kappa_0)$; *(sufficient-incentive)*
- $\lambda^R_1 + \lambda^L_1 \leq \lambda^R_0 + \lambda^L_0 + \epsilon$, *(slow-drift)*

letting $\mathcal{M}$ be the myopic-incentive contract (Definition 3), we have $\mu^N_{\tau_{\mathcal{M}}} - \mu^N_{T_{\lambda,D,c}} \leq \eta$.

Proof of Lemma 6. For any time $t \geq 1$, given any information arrival probabilities λ such that $\lambda^R_1 + \lambda^L_1 \leq \lambda^R_0 + \lambda^L_0 + \epsilon$, we have

$$\begin{aligned}
\mu^N_{t-1} - \mu^N_t &= \mu^N_{t-1} - \frac{\mu^N_{t-1}(1 - \lambda^R_1 - \lambda^L_1)}{\mu^N_{t-1}(1 - \lambda^R_1 - \lambda^L_1) + (1 - \mu^N_{t-1})(1 - \lambda^R_0 - \lambda^L_0)} \\
&\leq \mu^N_{t-1}\left(1 - \frac{(1 - \lambda^R_1 - \lambda^L_1)}{(1 - \lambda^R_1 - \lambda^L_1) + (1 - \mu^N_{t-1})\epsilon}\right) \\
&\leq 2\mu^N_{t-1}(1 - \mu^N_{t-1})\epsilon \leq \frac{1}{2}\epsilon. \qquad (7)
\end{aligned}$$

The second inequality holds since $\lambda_1^R + \lambda_1^L \leq \frac{1}{2}$ and the last inequality holds since $\mu_{t-1}^N(1 - \mu_{t-1}^N) \leq \frac{1}{4}$.

Let $\Delta_\lambda \triangleq \lambda_1^R - \lambda_0^R$. By the sufficient-incentive assumption, $\Delta_\lambda \geq \frac{c+\kappa_0}{D}$ and hence $\frac{c}{\Delta_\lambda} \leq \frac{cD}{c+\kappa_0} < D < \frac{1}{2}$. Therefore,

$$\mu_{\lambda,c} = \min\left\{\frac{1}{2}, \frac{c}{\Delta_\lambda}\right\} = \frac{c}{\Delta_\lambda}.$$

Fix any $\eta > 0$, and let $\epsilon = \eta$. By (7), for every $t \geq 1$,

$$0 \leq \mu_{t-1}^N - \mu_t^N \leq \frac{\epsilon}{2}.$$

Since $T_{\lambda,D,c}$ is the maximum time such that $\mu_{T_{\lambda,D,c}-1}^N \geq \mu_{\lambda,c}$, we have $\mu_{T_{\lambda,D,c}}^N \geq \mu_{\lambda,c} - \frac{\epsilon}{2}$. Now consider any time t such that $\mu_{t-1}^N \geq \mu_{T_{\lambda,D,c}}^N + \eta$. Then

$$\mu_{t-1}^N \geq \mu_{\lambda,c} - \frac{\epsilon}{2} + \eta = \mu_{\lambda,c} + \frac{\eta}{2} \geq \mu_{\lambda,c}.$$

Under the myopic-incentive contract, the expected reward gain from exerting effort for one period at time t and then stopping is at least

$$\mu_{t-1}^N \lambda_1^R + (1 - \mu_{t-1}^N)(1 - \lambda_0^R) \cdot \frac{\mu_{t-1}^N}{1 - \mu_{t-1}^N} - \mu_{t-1}^N = \mu_{t-1}^N \Delta_\lambda \geq \mu_{\lambda,c} \Delta_\lambda = c.$$

Hence the agent has an incentive to exert effort at time t.

Therefore, it cannot be that $\mu_{\tau_\mathcal{M}}^N > \mu_{T_{\lambda,D,c}}^N + \eta$. Indeed, since μ_t^N is weakly decreasing in t, this inequality implies $\tau_\mathcal{M} < T_{\lambda,D,c}$, so the time $t = \tau_\mathcal{M} + 1$ is feasible and satisfies

$$\mu_{t-1}^N = \mu_{\tau_\mathcal{M}}^N > \mu_{T_{\lambda,D,c}}^N + \eta.$$

By the argument above, the agent would then still prefer to exert effort for one more period at time $\tau_\mathcal{M} + 1$, contradicting the definition of $\tau_\mathcal{M}$ (together with our tie-breaking rule toward the largest stopping time). Thus,

$$\mu_{\tau_\mathcal{M}}^N - \mu_{T_{\lambda,D,c}}^N \leq \eta,$$

and the claim holds. □

To prove Theorem 5, 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.

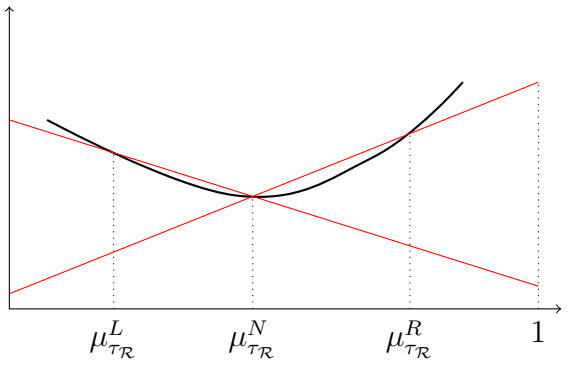

Figure 11: The black curve is the expected score function U_P. The red lines are linear functions $\underline{U}$ and $\overline{U}$ respectively.

Lemma 7. *For any bounded static scoring rule P with expected reward function $U_P(\mu)$ given posterior belief μ, we have*

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

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 $|U_P(\mu+\epsilon) - U_P(\mu)| \le \epsilon$ for any $\epsilon > 0$ and $\mu \in [0, 1-\epsilon]$. □

Proof of Theorem 5. By Lemma 6, 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 $\mathcal{M}$ that can be implemented as a scoring rule, we have $\mu^N_{\tau_\mathcal{M}} - \mu^N_{T_{\lambda,D,c}} > \eta$.

Suppose by contradiction there exists a contract $\mathcal{M}$ that can be implemented as a scoring rule and $\mu^N_{\tau_\mathcal{M}} - \mu^N_{T_{\lambda,D,c}} \le \eta$. Let P be the scoring rule that implements contract $\mathcal{M}$ 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 $\underline{U}(\mu^L_{\tau_\mathcal{M}}) = U_P(\mu^L_{\tau_\mathcal{M}})$ and $\underline{U}(\mu^N_{\tau_\mathcal{M}}) = U_P(\mu^N_{\tau_\mathcal{M}})$. Let $\overline{U}(\mu)$ be a linear function such that $\overline{U}(\mu^R_{\tau_\mathcal{M}}) = U_P(\mu^R_{\tau_\mathcal{M}})$ and $\overline{U}(\mu^N_{\tau_\mathcal{M}}) = U_P(\mu^N_{\tau_\mathcal{M}})$. See Figure 11 for an illustration. Let $f^s_t \triangleq \mu^N_t \lambda^s_1 + (1-\mu^N_t)\lambda^s_0$ for any $s \in \{L, R\}$. At time $\tau_\mathcal{M}$, the agent has incentive to exert effort, which implies that the cost of effort c is at most the utility increase for exerting effort

$$\begin{aligned}
&f^R_{\tau_\mathcal{M}-1} \cdot U_P(\mu^R_{\tau_\mathcal{M}}) + f^L_{\tau_\mathcal{M}-1} \cdot U_P(\mu^L_{\tau_\mathcal{M}}) + (1 - f^R_{\tau_\mathcal{M}-1} - f^L_{\tau_\mathcal{M}-1}) \cdot U_P(\mu^N_{\tau_\mathcal{M}}) - U_P(\mu^N_{\tau_\mathcal{M}-1}) \\
&= f^R_{\tau_\mathcal{M}-1} \cdot (U_P(\mu^R_{\tau_\mathcal{M}}) - \underline{U}(\mu^R_{\tau_\mathcal{M}})) + \underline{U}(\mu^N_{\tau_\mathcal{M}-1}) - U_P(\mu^N_{\tau_\mathcal{M}-1}) \le f^R_{\tau_\mathcal{M}-1} \cdot (U_P(\mu^R_{\tau_\mathcal{M}}) - \underline{U}(\mu^R_{\tau_\mathcal{M}}))
\end{aligned}$$

where the equality holds by linearity of expectation and the inequality holds by the convexity

of utility function U_P. Therefore, we have

$$\begin{aligned} U_P(\mu^R_{\tau_{\mathcal{M}}}) - \underline{U}(\mu^R_{\tau_{\mathcal{M}}}) &\geq \frac{c}{f^R_{\tau_{\mathcal{M}}-1}} = \frac{c}{\mu^N_{\tau_{\mathcal{M}}-1}\lambda^R_1 + (1-\mu^N_{\tau_{\mathcal{M}}-1})\lambda^R_0} \\ &\geq \frac{(\lambda^R_1 - \lambda^R_0)(1 - \frac{\eta}{\mu^N_{\tau_{\mathcal{M}}-1}})}{\lambda^R_1 + \frac{1}{\mu^N_{\tau_{\mathcal{M}}-1}}(1-\mu^N_{\tau_{\mathcal{M}}-1})\lambda^R_0} \geq \frac{(\lambda^R_1 - \lambda^R_0)}{\lambda^R_1 + \frac{1}{\mu^N_{\tau_{\mathcal{M}}-1}}(1-\mu^N_{\tau_{\mathcal{M}}-1})\lambda^R_0} - \frac{2\eta}{\underline{\kappa}_1} \end{aligned}$$

where the second inequality holds since $\mu^N_{\tau_{\mathcal{M}}} \leq \mu^N_{T_{\lambda,D,c}} + \eta \leq \mu_{\lambda,c} + \eta$, and the last inequality holds since $\lambda^R_0 \geq \underline{\kappa}_1$.

Fix any $\gamma > 0$ such that $\gamma < \frac{D\kappa_0}{2(c+\kappa_0)}$. For any sufficiently small η and ε, let time $t \leq \tau_{\mathcal{M}}$ be such that $\mu^N_{t-1} - \mu^N_{\tau_{\mathcal{M}}-1} > \gamma$. First note that the convexity of U_P and the constraint on rewards belonging to the unit interval at state 1 implies that

$$\overline{U}(\mu^N_{t-1}) - U_P(\mu^N_{t-1}) \leq \frac{2\eta}{\underline{\kappa}_1} \cdot \frac{1-\mu^N_{\tau_{\mathcal{M}}}}{1-\mu^R_{\tau_{\mathcal{M}}}} \cdot \frac{\mu^R_{\tau_{\mathcal{M}}} - \mu^N_{t-1}}{\mu^R_{\tau_{\mathcal{M}}} - \mu^N_{\tau_{\mathcal{M}}}} \leq \frac{2\eta(\underline{\kappa}_1 + \bar{\kappa}_1)}{\underline{\kappa}^2_1}$$

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

$$U_P(\mu^R_t) - \overline{U}(\mu^R_t) \leq \frac{2\eta}{\underline{\kappa}_1} \cdot \frac{1-\mu^N_{\tau_{\mathcal{M}}}}{\mu^R_{\tau_{\mathcal{M}}} - \mu^N_{\tau_{\mathcal{M}}}} \leq \frac{2\eta}{\underline{\kappa}_1} \cdot \frac{\lambda^R_1}{\mu^N_{\tau_{\mathcal{M}}}(\lambda^R_1 - \lambda^R_0)} \leq \frac{2\eta\bar{\kappa}_1}{\underline{\kappa}_1 c}.$$

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

$$\begin{aligned} &f^R_{t-1} \cdot U_P(\mu^R_t) + f^L_{t-1} \cdot U_P(\mu^L_t) + (1 - f^R_{t-1} - f^L_{t-1}) \cdot U_P(\mu^N_t) - U_P(\mu^N_{t-1}) \\ &\leq f^R_{t-1} \cdot \overline{U}(\mu^R_t) + f^L_{t-1} \cdot \underline{U}(\mu^L_t) - (f^R_{t-1} + f^L_{t-1}) \cdot \overline{U}(\mu^N_t) + \epsilon + \frac{2\eta\bar{\kappa}_1}{\underline{\kappa}_1 c} + \frac{2\eta(\underline{\kappa}_1 + \bar{\kappa}_1)}{\underline{\kappa}^2_1} \\ &\leq \mu^N_{\tau_{\mathcal{M}}}(\lambda^L_0 - \lambda^L_1) - \gamma\underline{\kappa}_1 + \epsilon + \frac{2\eta\bar{\kappa}_1}{\underline{\kappa}_1 c} + \frac{2\eta(\underline{\kappa}_1 + \bar{\kappa}_1)}{\underline{\kappa}^2_1}. \end{aligned}$$

Since $c = \mu_{\lambda,c}(\lambda^R_1 - \lambda^R_0) \geq (\mu^N_{\tau_{\mathcal{M}}} - \eta)(\lambda^L_0 - \lambda^L_1)$, the agent suffers a loss at least

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

for exerting effort in one period.

Now consider the utility increase for exerting effort in continuation game $\mathcal{G}_t$, starting from belief μ^N_{t-1}. Note that for any $\delta > 0$, by (7) the no-information belief can decrease by at most $\epsilon/2$ in one period. Hence before the no-information belief drifts by a total distance δ,

the agent must work for at least

$$N \triangleq \left\lceil \frac{2\delta}{\epsilon} \right\rceil$$

periods. Moreover, in each such period the total probability of receiving a Poisson signal is at least

$$f_{u-1}^R + f_{u-1}^L \geq 2\underline{\kappa}_1,$$

because each arrival rate lies in $[\underline{\kappa}_1, \bar{\kappa}_1]$. Therefore the probability of receiving at least one Poisson signal before the no-information belief drifts by δ is at least

$$1 - (1 - 2\underline{\kappa}_1)^N \geq 1 - \exp\left(-2\underline{\kappa}_1 N\right).$$

By continuity of the preceding one-period upper bound as a function of the current no-information belief, after possibly shrinking δ we may ensure that the loss is at least

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

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, and it can occur only if no Poisson signal arrives in the first N periods. The probability of this event is at most

$$(1 - 2\underline{\kappa}_1)^N \leq \exp\left(-2\underline{\kappa}_1 N\right).$$

Moreover, the probability that the agent reaches period $j+1$ without a Poisson signal in the first j periods is at most $(1 - 2\underline{\kappa}_1)^j$, for $j = 0, \ldots, N-1$. Therefore, by decomposing the continuation gain into the incremental gains from each successive exertion decision, the maximal continuation gain from exerting effort in continuation game $\mathcal{G}_t$, relative to stopping immediately, is at most

$$-m\sum_{j=0}^{N-1}(1 - 2\underline{\kappa}_1)^j + (1 - 2\underline{\kappa}_1)^N \leq -m\sum_{j=0}^{N-1}(1 - 2\underline{\kappa}_1)^j + \exp\left(-2\underline{\kappa}_1 N\right).$$

Since $m > 0$ when δ, η, ϵ are chosen sufficiently small compared to γ, we can then keep $\delta > 0$ fixed and choose ϵ sufficiently small so that $N = \lceil 2\delta/\epsilon \rceil$ is large enough for the right-hand side to be negative. Therefore, the agent's utility for exerting effort is smaller than not exerting effort in continuation game $\mathcal{G}_t$. This leads to a contradiction since the agent at time t will not choose to exert effort given scoring rule P. □

OA 2 General Dynamic Contracts

OA 2.1 Menu Representation

Proof of Lemma 3. Recall that by Lemma 2, it is without loss to assume that the agent uses a stopping strategy. Given any time t, let r_t^s be the menu option chosen by the agent if the agent receives a Poisson signal $s \in S$ at time t. Let $r_{\tau_{\mathcal{M}}}^N$ be the menu option the agent chooses if the agent does not receive any Poisson signal before stopping effort. By offering the menu options $\mathcal{M}_t = \{r_{t'}^s\}_{t' \geq t, s \in S} \cup \{r_{\tau_{\mathcal{M}}}^N\}$ at time t, the agent's utilities for stopping and exerting effort remain the same, and hence this simplified menu representation implements the same stopping strategy. □

OA 2.2 Communication-based Contracts

Let M_t be the message space of the agent at any time t. We denote the history at time t as $h_t = \{m_{t'}\}_{t' \leq t}$. Let $\mathcal{H}_t$ be the set of all possible histories at time t. Since we allow rewards to condition on the future realized outcome of interest (i.e., the state), we take the rewards to be within the unit interval. As mentioned in our model, we can interpret the reward as an endorsement that positively influences the forecaster's reputation. The agent thus receives a benefit according to the function:

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

where $\mathcal{R}(h_T, \theta)$ 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—or, alternatively, the probability that the agent receives the reward, so that $\mathcal{R}(h_T, \theta)$ can be interpreted as an endorsement probability.[23] If the agent has exerted effort in $\tilde{t}$ periods, his final payoff is $\mathcal{R}(h_T, \theta) - c\tilde{t}$.

This representation of the contract is equivalent to our menu representation. In particular, given any contract $\mathcal{R}$, at any time t, it is sufficient to provide the agent with menu options that will be implemented for the on-path beliefs μ_t^s for receiving a Poisson signal s, or a belief μ_t^N for not receiving any Poisson signal. Therefore, all our results extend naturally.

[23] While we allow randomization over the event that the agent receives the full reward, we do not allow stochastic messages from the principal to the agent. A discussion of how the possibility of such randomization influences the results is deferred to Section OA 2.3. Allowing stochastic messages would introduce a second wedge between scoring rules and arbitrary contracts, which we do not consider to maintain focus on the ability to elicit information over time.

OA 2.3 Randomized Contracts

Our goal has been to determine the maximum effort implementable within the class of contracts defined in Section 2.2 and OA 2.2. As histories only include messages sent by the agent, we implicitly rule out the use of stochastic messages sent from the designer to the agent (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 or how a mechanism designer could commit to implementing 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

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

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. Implementing such randomized contracts would require either a public randomization device or the principal to commit to a random effort recommendation policy.

It may no longer be without loss of generality to restrict attention to stopping strategies under a randomized contract. In particular, the agent may decide whether to work or not, depending on the past realizations of the public randomization. The agent may also strategically delay exerting effort to wait for the realization of the public randomization. As a result, simplifying the objective to maximizing the agent's stopping time is not appropriate for randomized contracts.

At the same time, our analysis provides some insights into why 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 in the absence of a Poisson signal arrival—so that the realized stopping time increases for *some* realization of the randomization.

We illustrate this intuition more formally assuming perfect "good-news" learning, i.e., $\lambda_1^R > 0$ and $\lambda_0^R = \lambda_1^L = \lambda_0^L = 0$, where learning is perfect and only right-biased signal R arrives with positive probability. We describe the resulting solution in Section OA 5. In this setting, our results imply that, under deterministic contracts, a V-shaped scoring rule with parameters $r_0 = 1$ and $r_1 \in [0, 1]$ implements the effort-maximizing contract $\mathcal{R}$; in particular, Section OA 5 characterizes when in fact effort-maximizing contracts require $r_1 < 1$. Recall that we denote $\tau_{\mathcal{M}}$ as the stopping time in the effort-maximizing deterministic contract and let $\mu^N_{\tau_{\mathcal{M}}}$ 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 $\frac{2\tau_{\mathcal{M}}}{3}$ 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 $\mu^N_{\frac{\tau_{\mathcal{M}}}{2}}$. Consider the randomized contract $\hat{\mathcal{R}}$ that provides menu options $(1, 0)$ and $(0, r_1 - \delta)$ from time 0 to $\frac{\tau_{\mathcal{M}}}{2}$, and after time $\frac{\tau_{\mathcal{M}}}{2}$, 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 $t \leq \frac{\tau_{\mathcal{M}}}{2}$. Moreover, after time $\frac{\tau_{\mathcal{M}}}{2}$, 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 $\tau_{\mathcal{M}}$ in the absence of a Poisson signal.[24]

OA 3 Additional Review of Scoring Rules

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

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

[24]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 $v(1, 0) = 1, v(1, 1) = -\frac{1-\mu^N_{\tau_{\mathcal{M}}}}{\mu^N_{\tau_{\mathcal{M}}}}$; plainly, the designer only seeks to change her action from 1 to 0 if the posterior belief is below $\mu^N_{\tau_{\mathcal{M}}}$. 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.

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

Lemma 8 (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.[25]

OA 4 Comparative Statics of Scoring Rules

As discussed in Section 3.4, 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 times, the agent is more uncertain about the states and hence it is easier to incentivize. In this appendix, we formalize this intuition using 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 $\{R, L\}$ 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, with $\sum_{s\in\{R,L\}} f_{\theta,s} = 1$ for any 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

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

Since $f_{0,L} > f_{1,L}$, we have $f_{0,L} + f_{1,R} = f_{0,L} + 1 - f_{1,L} > 1$. Therefore, the expected score increase is monotone increasing in prior D. That is, the closer the prior is to $\frac{1}{2}$, the easier

[25]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.

it is to incentivize the agent to exert effort.

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

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

Since $f_{1,R}<1$, the strength of the incentives the designer can provide is strictly decreasing in prior D'. Therefore, even though the 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 (assuming that the cost of effort is the same in both settings).

OA 5 Complete Solution under Perfect Good News Learning

OA 5.1 Effort-Maximizing Contracts

Here we describe our findings in the model where the time interval is Δ and take the $\Delta\to 0$ limit. We focus on the setting with a single perfectly revealing signal, i.e.,

$$\lambda_0^L=\lambda_1^L=\lambda_0^R=0,\quad \lambda_1^R>0.$$

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

For this learning environment, Theorem 3 implies 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 unlikely. In fact, one can show the agent is indifferent between continuing effort and selecting a reward when $\mu_\tau^N=\frac{c}{\lambda_1^R r_1}$ (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

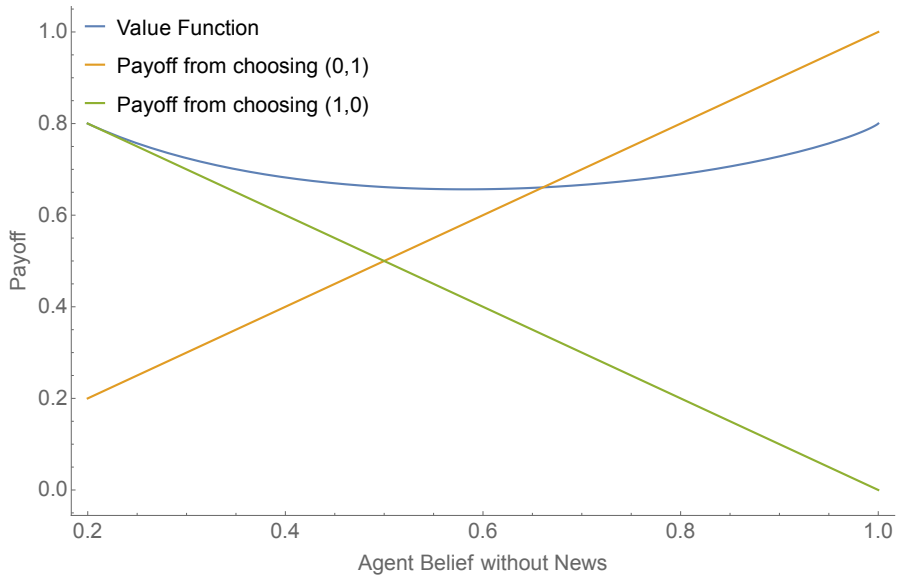

Figure 12: Value function with perfect learning; $r_1 = 1, c = .2, \lambda_1^R = 1$.

the agent's initial incentive constraints.

Figure 12 illustrates the agent's value function when $r_1 = 1$, assuming the agent works until $\mu_\tau^N = \frac{c}{\lambda_1^R}$, 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 $\mu_\tau^N = \frac{c}{\lambda_1^R r_1}$, so that the agent stops sooner when r_1 is lower. Now, when $\mu_0^N < c/\lambda_1^R$, it is impossible to induce the agent to acquire any information at all. Outside of this range, a pair of thresholds, μ^*, μ^{**} satisfying $c/\lambda_1^R < \mu^* < \mu^{**} < 1$ determine the form of the effort-maximizing scoring rule:

- For $c/\lambda_1^R \leq \mu_0^N \leq \mu^*$, 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 $c/(\lambda_1^R r_1)$ 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 observed a Poisson signal 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 5.3, 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 some of the intuition on how effort-maximizing rewards vary with the agent's beliefs.

OA 5.2 Details Behind the Calculation

Here we provide some additional details behind the calculations in Section OA 5. 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:

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

Imposing indifference between stopping and continuing yields:

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

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:

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

Algebraic manipulations show this coincides 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:

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

From this, we obtain the following differential equation:

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

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 $V(\mu_\tau^N) = (1 - \mu_\tau^N)$):

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

Note that $V''(\mu_t^N) > 0$ for this solution, as well as that $V'(c/(r_1\lambda_1^R)) = -1$, 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.

OA 6 Two Periods

In this section, we consider a simple two-period model and show that in the two-period environment, without any assumptions on the information acquisition process, the optimal contract always has a static implementation as a scoring rule.

Proposition 2. *In the two-period environments, there exists a scoring rule that implements maximum effort.*

Proof. First, note that if in the optimal contract, the agent only exerts effort for one period or does not exert effort at all, the contract for the second period is irrelevant; hence, it is immediate that there exists a scoring rule that implements maximum effort.

Now we focus on the case in which the agent can be incentivized to exert effort in two periods. Without loss of generality, we can assume that the principal offers a menu

$\mathcal{M}_2 = \{r_2^s\}_{s\in S} \cup r_2^N$ at time 2, and a menu $\mathcal{M}_1 = \{r_1^s\}_{s\in S} \cup \mathcal{M}_2$ at time 1. In this case, consider another menu $\widehat{\mathcal{M}}$ where $\widehat{\mathcal{M}}_1 = \widehat{\mathcal{M}}_2 = \mathcal{M}_1$. The agent's incentive to exert costly effort weakly increases in the first period as his continuation payoff weakly increases. Thus, the agent exerts effort for at least one period. Moreover, observe that if the agent knows that he will choose not to exert effort in the second period, he could pretend to receive a Poisson signal in the first period and obtain a payoff from menu $\mathcal{M}_1$ before entering the second period. Therefore, by including $\mathcal{M}_1$ in $\widehat{\mathcal{M}}_2$, the agent's payoff from not exerting effort remains the same. On the other hand, the agent has more options to choose from in the second period if he exerts effort. His incentive for exerting effort also weakly increases in the second period.

Combining all cases, there exists a scoring rule that implements maximum effort. □

Intuitively, the tension in dynamic incentives is that if we provide a higher reward to the agent at certain periods, the agent has a stronger incentive to exert effort in previous periods but a weaker incentive to exert effort in later periods. However, in the two-period model, the latter effect disappears since when we adjust the rewards in the second period, there won't be any future periods. Therefore, a static implementation, i.e., a scoring rule, implements the maximum effort.