Delegating Learning*

Juan F. Escobar† Qiaoxi Zhang‡

August 23, 2019

Abstract

Learning is crucial to organizational decision making but often needs be delegated. We examine a dynamic delegation problem where a principal decides on a project with uncertain profitability. A biased agent, who is initially as uninformed as the principal, privately learns the profitability over time and communicates to the principal. We formulate learning delegation as a dynamic mechanism design problem and characterize the optimal delegation scheme. We show that private learning gives rise to the tradeoff between how much information to acquire and how promptly it is reflected in the decision. We discuss implications on learning delegation for distinct organizations.

Keywords: private learning, delegation, delays, deadlines, commitment, cheap talk

*We are extremely grateful for comments from Nageeb Ali, Alessandro Bonatti, Rahul Deb, Elton Dusha, Federico Echenique, Andrew McClellan, Asher Wolinsky and participants of various conferences and seminars. We acknowledge financial support from the Institute for Research in Market Imperfections and Public Policy, MIPP, ICM IS130002. Zhang also acknowledges financial support from CONICYT-FONDECYT postdoctoral award (Project 3170783).

†Department of Industrial Engineering, University of Chile, Beauchef 851, Santiago, Chile. E-mail: jescobar@dii.uchile.cl.

‡Department of Industrial Engineering, University of Chile, Beauchef 851, Santiago, Chile. E-mail: jdzhangqx@gmail.com.

1 Introduction

Suppose that two people need to decide whether to invest in a project. If they invest, they could receive a gain or suffer a loss. If they do not invest, they wait, obtain new information, and may invest in the future. Now suppose that one of them, the principal, prefers to learn, while the other, the agent, prefers to invest. If learning has to be delegated to the agent and the principal cannot observe the learning outcome, can the agent convey it truthfully? If so, what should be the optimal delegation scheme? How does it change over time given what the agent has learned so far? How long should learning take place?

Delegating learning is a common occurrence. For example, suppose the board of directors of a company is deliberating whether to acquire another company. Apart from the financial value of the acquisition, its strategic value—e.g., its impact on the price and competition, the current employees, the bargaining power with suppliers—is also relevant. Since much of this information is hard to observe directly, the board needs to rely on learning by the manager, who has direct access to the parties involved. For another example, the Food and Drug Administration (FDA) relies on pharmaceutical companies to develop drugs and test their efficacy. Although the companies are required to submit clinical trial results, the trials themselves cannot be fully monitored, and therefore the results can be manipulated.[1]

We study how a principal should delegate an investment decision to an agent who privately learns about the investment over time. Our analysis extends the traditional static delegation approach (Holmstrom (1984); Melumad and Shibano (1991); Alonso and Matouschek (2008); Amador and Bagwell (2013)) to allow for evolving private information. We formulate delegation as a dynamic mechanism design problem and characterize its solution. Our results shed light on how organizations could incentivize learning and informative communication with an evolving delegation scheme and why distinct organizations should implement distinct learning delegation procedures.

In our model, a principal and an agent face a project that never expires. The project's quality, which can be good or bad, is initially uncertain. Players share the same belief about the project's quality. The principal needs to decide when, if ever, to invest in the project. The project generates a signal whose arrival time is random.

[1] Several authors have documented frauds in clinical trials (George and Buyse, 2015). Seife (2015) shows that the FDA has found substantive evidence of fraudulent data in biomedical research on humans. For a summary, see Seife, "Are your medications safe," Slate, February 9, 2015.

As long as no investment has happened, the agent privately observes the signal, or the absence thereof, without cost. Hence, investing and learning are two sides of the same coin in that as long as investment has not happened, learning continues. A signal perfectly reveals the project quality. At each point in time before investment happens, the agent sends a cheap talk message about the information learned so far to the principal. The principal commits to a delegation rule that specifies, for each point in time and each possible message history at that time, whether to invest or not. Once the investment happens, the game ends.

No one receives a payoff if no investment has happened. Once the investment happens, each player receives a time-discounted payoff determined by project quality and the player's identity. A good investment brings gains to both players, while a bad investment brings losses to both players. Consequently, each player will want to invest if they are optimistic enough about project quality and will want to wait for more information otherwise. However, the common initial belief is high enough for the agent such that he prefers to invest immediately and low enough for the principal such that she prefers to wait and invest only when a good signal arrives. The challenge for the principal is then to incentivize the agent to tell the truth when he has not received any signal, while trying to invest as soon as possible after a good signal arrives.

First, we note that investment must follow a good signal with delay. If a good signal triggers immediate investment, the agent would like to pretend to be informed when in fact no signal has arrived, and no learning would take place. To see how the delay should evolve over time to incentivize learning, we need to understand the driving forces behind learning. Learning benefits the agent because if a bad signal arrives, he would then learn that the project is bad and avoid the loss from investing. On the other hand, learning costs the agent in that it takes time. Suppose that no signal has arrived and the agent is still optimistic enough to prefer to invest right away. At this point, the cost of learning outweighs the benefit. To encourage learning, the principal needs to decrease the cost by making investment respond faster to the good signal throughout time. Suppose that the principal wants to encourage the agent to learn for one more day. If a claim of good signal leads the principal to invest immediately, then the cost of learning is one day's delay. However, if a good signal that arrives today leads to investment 5 days later while a good signal that arrives tomorrow leads to investment 4.5 days after tomorrow, the cost of learning is only a half-day's delay. Delays in investment that decrease in the arrival times of the good signal allow the principal to balance the cost and benefit of learning for the agent, hence his truthful revelation.

It is natural to think that if no good signal has arrived, investment should not happen. This is not always the case. In fact, as a result of the trade-off between the amount of information acquired and how effectively it is used, the principal may prefer to invest at a deadline even if no good signal has been claimed. Suppose that at some time T, even if no signal has arrived, learning stops and investment happens. Since no incentives for learning is required from T on, the delay decreases gradually to 0 at T. If instead the principal decides to incentivize learning after T, the delay at T must be positive. Accordingly, delays of investment for claims of the good signal at each time before T must also be increased. Therefore, the longer learning takes place, the better information the principal receives, the more accurate but less prompt her decision to invest is. If no investment happens unless a good signal arrives, learning could take place for an arbitrarily long period of time. Consequently, the decision to invest is 100% accurate because the principal only invests if she is completely sure that the project is good. However the downside is that she has to provide incentives for learning for a long time. The resulting long delays in investment when a good signal arrives can therefore be a prohibitive cost for the principal.

Our results have implications on learning delegation in organizations. First, although delays in FDA's approval process are much criticized,[2] we show that they have a key role in ensuring truthful revelation by the pharmaceutical companies. Second, our results speak to how distinct organizations should use distinct protocols to delegate learning. With drug approval, the loss from approving a damaging drug is substantial. FDA's optimal action is then to be prudent by establishing long revision processes so that a damaging drug will never get approved, despite the resulting delay. In contrast, if a manager's career concern is strong and he has a high gain-loss ratio, the optimal action for the board when it comes to acquisition decisions is to set short learning phases and then acquire as long as no negative news has arrived.

Our results are robust to a number of variations of the environment. First, we allow the principal to use random mechanisms. We show that as long as the principal

[2] See for example: Robert Pear, "Fast approval of AIDS drugs is urged," New York Times, August 16, 1990 and Daniel B. Klein, "Economists against the FDA," Independent Institute, September 1, 2000.

is at least as patient as the agent, the optimal deterministic mechanism remains so even when randomizations are allowed. Next, we allow the principal to use time-dependent transfers that are paid after a good signal is revealed. We show that the principal should use delays instead of transfers to motivate the agent. Again, the optimal contract from our baseline model remains unaltered when more general mechanisms are allowed. Lastly, we explore a model in which the agent needs to incur an unverifiable cost to acquire information. The unverifiable learning cost makes the incentive problem much more severe as the agent may not only lie but also shirk. In this general environment, a contract with delays and deadlines can motivate the agent to dynamically acquire information and to report his signal truthfully.

Our paper contributes to the delegation literature initiated by Holmstrom (1984) and extended by Melumad and Shibano (1991), Alonso and Matouschek (2008), Armstrong and Vickers (2010), Amador and Bagwell (2013), and Ambrus and Egorov (2017), among others. As in all these papers, in our model the principal may grant flexibility to the agent so that he can use his information, but granting too much flexibility may open up room for opportunistic behavior. However, these papers study static models and do not address the issue of how to provide incentives to an agent with evolving private information.[3] In particular, our work emphasizes how the dynamic provision of incentives determines how information is used and for how long learning takes place.

Grenadier, Malenko, and Malenko (2016) and Guo (2016) explore delegation models in dynamic contexts. In Grenadier, Malenko, and Malenko (2016), a timing decision needs to be made and an agent who is informed at time 0 communicates with the principal throughout time. Whereas Grenadier, Malenko, and Malenko (2016) explore how the value of commitment for the principal depends on the sign of the agent's bias, we take commitment for granted but explore how to delegate with evolving private information. As Grenadier, Malenko, and Malenko (2016) point out, their full commitment case is similar to standard static delegation problems and, as a result, interval delegation is optimal. In Guo (2016), the principal delegates the decision to experiment over time to an agent who has private information about its profitability

[3]While in our model the principal dynamically screens the agent's information, we depart from the growing dynamic mechanism design literature (Pavan, Segal, and Toikka, 2014; Bergemann and Valimaki, forthcoming; Madsen, 2018) by assuming transfers are infeasible. We explore the role of transfers in Section 5 and Section E.2 of Supplementary Material.

at time 0.[4] Once experimentation starts, however, all signals are public. A comparison between our paper and Guo (2016) highlights the differences between private and public learning, which have important implications for the design of incentive schemes. In her model with a continuum of types, since signals are public, once a good signal arrives, the risky project is publicly known to be optimal and is fully implemented. In our model, however, investment decisions commonly known to be optimal are nonetheless delayed. This is the principal's response to the problem of providing incentives to an agent with evolving private information.

Our paper is also related to the study of optimal delegation decisions when information acquisition is endogenous. In Aghion and Tirole (1997), Szalay (2005) and Deimen and Szalay (forthcoming), information acquisition is a one-time decision, therefore the trade-off between extracting information and using information efficiently is different from ours. Lewis and Ottaviani (2008) study a setting where the agent searches for the best alternative over time and money transfers are used, which we rule out.

Frankel (2016), Li, Matouschek, and Powell (2017), Lipnowski and Ramos (2017), Guo and Hörner (2018), and Chen (2018) study repeated delegation models in which parties face a stream of decisions. In these models, incentives can be provided by linking the different decisions. In contrast, we study situations in which a single, irreversible decision is to be made and therefore linking decisions is infeasible.[5]

Finally, our work is related to dynamic persuasion models, particularly, McClellan (2017), Henry and Ottaviani (forthcoming), and Orlov, Skrzypacz, and Zryumov (forthcoming). These papers explore how to design approval rules when learning is costly, signals are public, and incentives are misaligned ex-post. In contrast, we mainly focus on the case where learning is costless, signals are private, and incentives are misaligned ex-ante. In Section 5.3, we allow costly learning.[6]

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 formalizes the dynamic delegation problem. Section 4 presents our main results. Section 5 discusses some extensions. Section 6 concludes.

---

[4] Guo (2016) focuses on the full commitment case, but she also shows that the sign of the agent's bias determines the value of commitment.

[5] Another difference is that, with the exception of Guo and Hörner (2018), the repeated delegation literature has focused on serially uncorrelated incomplete information.

[6] It is not possible to extract and use any information in a model with private learning in which incentives are misaligned ex-post (and not just ex-ante as in our baseline model).

2 The Model

We consider an infinite-horizon continuous-time game played by a principal and an agent. There is an initially unknown state $\theta \in \{0, 1\}$. We call $\theta = 1$ the good state and $\theta = 0$ the bad state. At time 0, the agent and the principal are symmetrically uninformed about the state θ, with $\mathbb{P}[\theta = 1] = p_0$ being the initial prior.

The agent privately learns about the state without cost.[7] A signal is generated according to an exponential distribution with arrival rate λ^θ, which depends on θ. Specifically, conditional on θ, over an interval $[t, t + dt]$, a signal $s_t = \theta$ is realized with probability $\lambda^\theta dt$. The arrival of the signal is privately observed by the agent. Thus, the arrival of a signal perfectly reveals the state to the agent. We say that the agent is uninformed if he has not observed a signal. The agent's private history up to period t is denoted h^t. We use $\varnothing$ to denote the history with no signal.

The private belief process $p_t = \mathbb{P}[\theta = 1 \mid h^t]$ is formed according to the initial prior p_0 and the agent's private history h^t up to period t. The law of motion for the agent's private belief p_t can be derived as follows. If a signal $s_t = 1$ arrives during the interval $[t, t + dt)$, the belief jumps to 1; if a signal $s_t = 0$ arrives during the interval, the belief jumps to 0. If no signal arrives, Bayes's rule can be used to deduce that the posterior at the end of $t + dt$ is

$$p_t + dp_t = \frac{p_t(1 - \lambda^1 dt)}{(1 - p_t)(1 - \lambda^0 dt) + p_t(1 - \lambda^1 dt)}.$$

That is, when no signal arrives, the evolution of the belief is governed by the differential equation[8]

$$\frac{dp_t}{dt} = -(\lambda^1 - \lambda^0)p_t(1 - p_t).$$

We assume that $\lambda^0 < \lambda^1$ and thus no news is bad news. In other words, the belief decreases in the absence of a signal. We show that our results extend to the case $\lambda^0 \geq \lambda^1$ in Section E.3.

The principal chooses $y_t \in \{0, 1\}$ at each $t \geq 0$. $y_t = 1$ means to invest and $y_t = 0$ means not to invest. The decision to invest is irreversible: if $y_t = 1$ for some t, then $y_\tau = 1$ for all $\tau > t$ and the interaction ends.

Players' preferences over investment coincide conditional on θ. During each inter-

[7]In Section 5, we extend our model and results to incorporate costly learning.
[8]See Liptser and Shiryaev (2013) for details.

val $[t, t+dt)$ for which $y_t = 0$, both players receive zero payoff. Conditional on θ, if the principal invests at time t, she gets total discounted payoffs equal to

$$e^{-Rt}V \text{ if } \theta = 1 \text{ and } e^{-Rt}(-\nu) \text{ if } \theta = 0,$$

whereas the agent gets total discounted payoffs equal to

$$e^{-rt}W \text{ if } \theta = 1 \text{ and } e^{-rt}(-\omega) \text{ if } \theta = 0,$$

where $R, r > 0$ are the discount rates for the principal and the agent, respectively, and V, ν, W, and ω are strictly positive.

To state our assumption on the conflict of interest, it is useful to describe the one-person benchmark. Suppose that the agent not only perfectly observes the arrival of the signal but also has the right to invest. The optimal policy for the agent is characterized by a cutoff $p^* := \frac{(\lambda^1+r)\omega}{rW+(\lambda^1+r)\omega}$ (Keller, Rady, and Cripps, 2005). The agent finds it optimal to invest given the current belief p if and only if he is optimistic enough about the state; that is, $p \geq p^*$. Intuitively, the optimal policy must be a cutoff policy because if the uninformed agent does not find it attractive to invest at t, then neither does the uninformed agent at $t + dt$ who is more pessimistic about the value of the investment than at t. Similarly, suppose that the principal not only controls decisions but also observes the signal. Given the current belief p, the principal would find it optimal to invest iff $p \geq q^* = \frac{(\lambda^1+R)\nu}{RV+(\lambda^1+R)\nu}$.

We can now state the assumption on the conflict of interest, which is maintained throughout the paper.

Assumption 1 *$p^* < p_0 < q^*$.*

This assumption implies that at $t = 0$, the agent wants to invest immediately whereas the principal wants to invest only after observing a good signal. An equivalent formulation for Assumption 1 is

$$\frac{W}{\omega}\frac{r}{\lambda^1+r}\frac{p_0}{1-p_0} > 1 > \frac{V}{\nu}\frac{R}{\lambda^1+R}\frac{p_0}{1-p_0}.$$

One can interpret this to mean that the gain-loss ratio for the agent, W/ω, is sufficiently high while the gain-loss ratio for the principal, V/ν, is sufficiently low. Alternatively, one can think of the agent being sufficiently impatient (r large) and the principal being sufficiently patient (R small). Note that when Assumption 1 does not hold, the principal can easily align the agent's incentives.[9]

Since $\lambda^1 > \lambda^0$, as time goes on and no signal is received, the agent gets more pessimistic. At some point, the agent would prefer to wait and invest only after observing the good signal. Let t^* be the time at which the agent becomes indifferent between investing and waiting for a good signal. Formally, for $\lambda^1 > \lambda^0$,

$$t^* = \frac{1}{\lambda^1 - \lambda^0} \ln\Big(\frac{p_0}{1-p_0}\frac{W}{\omega}\frac{r}{(\lambda^1 + r)}\Big).$$

Figure 1 illustrates $(p_t)_{t\geq 0}$ and t^*. For $t < t^*$, the principal's and the agent's interests are not aligned when no signal has arrived. For $t > t^*$, the principal's and the agent's interests coincide for all private histories. We can thus interpret t^* as a measure of how long it takes for the incentives to be aligned. t^* increases as the agent becomes more willing to invest without any information (i.e. when W/ω becomes larger) and as the absence of signal becomes less informative (i.e. when $\lambda^1 - \lambda^0$ becomes smaller so that learning becomes slower).

3 The Dynamic Delegation Problem

We set up the principal's problem of eliciting the agent's evolving private information to maximize her expected profits. Following the delegation literature (Holmstrom, 1984), we focus on incentive provision through the design of control rights in the absence of transfers. To do this when learning is private, we formulate a dynamic mechanism design problem with commitment. At each $t \in [0, \infty)$ the agent sends a costless message $m_t \in \{0, 1, \emptyset\}$ given the private history h^t.[10] The principal commits to an action $y_t \in \{0, 1\}$ as a function of the message history up to t, $m^t \equiv \{m_\tau\}_{0\leq\tau<t}$.

A contract is a tuple $\langle T, \tau\rangle$, with $T \in \Re_+ \cup \{\infty\}$ and $\tau\colon [0, T] \to \Re_+$ if $T < \infty$ while $\tau\colon \Re_+ \to \Re_+$ if $T = \infty$. We use $\mathrm{dom}(\tau)$ to denote the domain of τ. If the agent has reported $m_t = \emptyset$ for all $t \in \mathrm{dom}(\tau)$, the principal invests at time T. The

[9] To see this, note that if $q^* < p_0$, then the principal would like to invest at $t = 0$ and would not need the agent. If $p_0 < p^*$ and $p_0 < q^*$, both the principal and the agent would like to invest only after observing a good signal. In this case, both parties' preferences are perfectly aligned throughout the game and the first best can be achieved even without commitment.

[10] Given our single-agent setting, it is without loss to restrict to direct mechanisms here as well as in the costly learning case in Section 5.3. See Sugaya and Wolitzky (2017) for details.

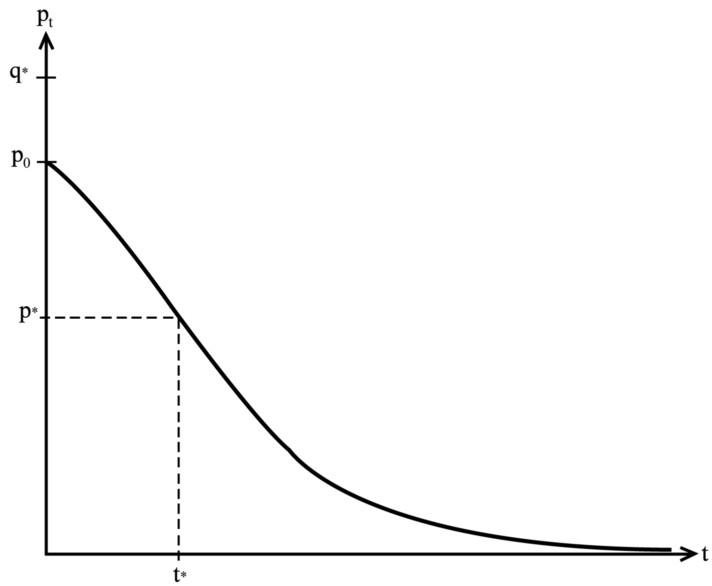

Figure 1: Evolution of p_t when no signal arrives. Preferences are aligned after time t^*. Parameter values: $\lambda^1 = 5$, $\lambda^0 = 4.8$, $R = r = 0.03$, $p^0 = 0.8$, $w = 10$, $W = 1800$, $t^* = 7.28$.

function τ is the time at which the investment is made when the agent reports that he has received the good signal at t ($m_t = 1$). The principal never invests after the agent reports a bad signal. In Section D of Supplementary Material, we define the general class of contracts and show that our restriction to contracts of the form $\langle T, \tau \rangle$ is without loss of generality. We also allow for random contracts in Section 5.[11]

We now describe the feasibility and incentive constraints. Since time is irreversible, $\tau(t) \geq t$ for all $t \in \text{dom}(\tau)$. To ensure the agent truthfully reveals when he is informed that the state is good at t instead of delaying the report, it must be that $\tau(t)$ is non-decreasing. Otherwise, take $\tau(t_1) > \tau(t_2)$ with $t_1 < t_2$ and note that the agent who receives the good signal at t_1 could wait and report the good signal at $t_2 > t_1$. The principal also needs to ensure that the informed agent at t reveals truthfully instead of pretending to be uninformed during the rest of the game. Formally, $\tau(t) \leq T$ for all $t \in \text{dom}(\tau)$.

A key incentive constraint is to ensure the uninformed agent at t does not want to

[11] Here the agent has no outside option. In Section E.1 of Supplementary Material, we consider the case where the agent is allowed to withdraw after claiming a good signal and receive payoff 0 before investment occurs and show that the optimal deterministic contract features a fixed investment time.

claim that he is informed and has received a good signal. To ensure truthful revelation of the uninformed agent at t, $\langle T, \tau \rangle$ must satisfy

$$\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}} e^{-r\tau(s)} W ds + \Big(p_t \frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} e^{-rT} W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}} e^{-rT}(-\omega)\Big) \geq \max\Big\{e^{-r\tau(t)}\big(-\omega + p_t(W+\omega)\big), 0\Big\}$$

for all $t \in \mathrm{dom}(\tau)$. Note that the agent can always claim that the state is bad and ensure a payoff equal to 0. The right-hand side is the maximum between 0 and the expected payoff of an uninformed agent at t (who has belief p_t) if he claims the state is good and induces investment at $\tau(t)$. The left-hand side is the agent's expected payoff if he claims to be uninformed and his continuation policy is to report truthfully. In this case, he could receive the good signal at $s < T$ and get the payoff $e^{-r\tau(s)}W$ with conditional probability $p_t\lambda^1 e^{-\lambda^1(s-t)}ds$, or receive no signal before T and induce an uninformed investment decision at T.[12]

The dynamic delegation problem can be formulated as:

$$\max_{T\in\Re_+\cup\{\infty\},\tau(\cdot)} \int_0^T p_0\lambda^1 e^{-\lambda^1 s} e^{-R\tau(s)} V ds + \Big(p_0 e^{-\lambda^1 T} e^{-RT} V + (1-p_0) e^{-\lambda^0 T} e^{-RT}(-\nu)\Big) \tag{1}$$

subject to

$$\tau(t) \geq t \quad \forall t \in \mathrm{dom}(\tau) \tag{2}$$

$$\tau \text{ is non-decreasing} \tag{3}$$

$$\tau(t) \leq T \quad \forall t \in \mathrm{dom}(\tau) \tag{4}$$

$$\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}} e^{-r\tau(s)} W ds + \Big(p_t \frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} e^{-rT} W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}} e^{-rT}(-\omega)\Big)$$

$$\geq \max\Big\{e^{-r\tau(t)}\big(-\omega + p_t(W+\omega)\big), 0\Big\} \quad t \in \mathrm{dom}(\tau). \tag{5}$$

[12] This incentive constraint could be considered insufficient as the agent could find it optimal to be truthful in some interval $[t, t+\epsilon]$ and lie after $t+\epsilon$. As we show in Section D of Supplementary Material, this is not the case.

This problem maximizes the principal's expected payoffs (1) over all contracts subject to the feasibility constraint (2) and the dynamic incentive constraints (3)-(5). The dynamic incentive constraints ensure that at any private history, the agent has incentives to truthfully reveal his information. Since private information changes over time, our dynamic delegation problem contrasts with most related problems in the literature that study either static models (Holmstrom, 1984; Melumad and Shibano, 1991; Alonso and Matouschek, 2008; Amador and Bagwell, 2013) or dynamic models in which information asymmetry is present at the beginning of the relationship (Guo, 2016; Grenadier, Malenko, and Malenko, 2016).

4 Analysis

In this section, we characterize the solution to the dynamic delegation problem.

4.1 Delays

This subsection characterizes the delay with which an investment commonly known to be profitable is implemented. The proofs are relegated to Appendix A. Our first result shows that optimal investments are delayed in any contract that satisfies the dynamic incentive constraints.

Lemma 1 *Let $\langle T, \tau \rangle$ satisfy (2) and (5). Then, $\tau(t) > t$, for all $t < \min\{t^*, T\}$.*

Conditional on the project being revealed profitable at $t < \min\{t^*, T\}$, the implementation time is inefficient (from both the principal's and the agent's perspectives). This distortion arises precisely due to the fact that learning is private: if the implementation time were not distorted and $\tau(t) = t$ for some $t < \min\{t^*, T\}$, the uninformed agent at t would claim he learned that the state is good in order to induce immediate investment.

In order to solve our dynamic delegation problem, it will be useful to find a solution τ to (1) keeping $T \in \Re_+ \cup \{\infty\}$ fixed. Solving the dynamic delegation problem for a fixed T is analytically useful and allows us to illustrate the tradeoffs involved when delegating to an agent who privately learns over time. The dynamic delegation problem keeping T fixed can be analyzed by finding solutions to the following *relaxed problem* (6). It is obtained by ignoring constraints (3)-(4) and by imposing the feasibility constraint (2) and the dynamic incentive constraint (5) over subsets of dom(τ).

$$\max_{\tau(\cdot)} \int_0^T p_0\lambda^1 e^{-\lambda^1 s}e^{-R\tau(s)}V ds + \Big(p_0 e^{-\lambda^1 T}e^{-RT}V + (1-p_0)e^{-\lambda^0 T}e^{-RT}(-\nu)\Big) \tag{6}$$

subject to

$$\tau(t) \geq t \quad \forall t \geq \min\{t^*, T\}, \tag{7}$$

$$\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}}e^{-r\tau(s)}W ds + \Big(p_t \frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}}e^{-rT}W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}e^{-rT}(-\omega)\Big)$$

$$\geq \max\Big\{e^{-r\tau(t)}\big(-\omega + p_t(W+\omega)\big), 0\Big\} \quad \forall t \leq \min\{t^*, T\}. \tag{8}$$

The following result establishes a necessary and sufficient optimality condition for the relaxed problem (6).

Lemma 2 *Let τ satisfy (7) and (8).*

(a) Suppose $T \leq t^*$. Then, τ solves the relaxed problem iff (8) binds for almost every $t \in [0, T]$.

(b) Suppose $T > t^*$. Then, τ solves the relaxed problem (6) iff (7) binds for almost every $t \geq t^*$ and (8) binds for almost every $t \leq t^*$.

In the optimal solution to the relaxed problem, the uninformed agent is indifferent between truthful revelation and claiming to know that the state is good for almost all $t \leq \min\{t^*, T\}$. To see this, suppose that τ is optimal and there is a set $A \subseteq [0, \min\{t^*, T\}]$ of positive measure such that for any $t' \in A$, the uninformed agent strictly prefers to reveal the truth. The principal could construct a new function τ' that coincides with τ outside of A but is slightly smaller than τ inside A. τ' results in higher expected payoffs for the principal than τ, and it satisfies (7) and (8). Thus, τ cannot be optimal. Moreover, Lemma 2 also shows that for $t > \min\{t^*, T\}$, there is no need to distort investment. Since after t^* the incentives are aligned, delaying investments only makes it harder to provide incentives before t^*.

We now further explore an important consequence of the binding incentive constraint (8) over $[0, \min\{t^*, T\})$.

Lemma 3 *Fix T and $\tau(\cdot)$ such that (8) binds for all $t < \min\{t^*, T\}$. Then, the derivative of τ with respect to t is given by*

$$\dot{\tau}(t) = (\frac{\lambda^0}{r})\frac{\omega}{W\frac{p_t}{1-p_t} - \omega}$$

for all $t < \min\{t^*, T\}$. In particular, over $t < \min\{t^*, T\}$, τ is strictly increasing and convex, and its slope is strictly less than 1.

This lemma characterizes the slope of a timing policy τ when (8) is binding. It can be intuitively derived as follows. Since (8) is binding everywhere in $[0, \min\{t^*, T\})$, the uninformed agent at t is indifferent between claiming he has received the good signal and truth-telling for all $t' \geq t$. The expected payoff the agent gets from truth-telling for $t' \geq t$ can be decomposed into the current and continuation payoffs. Current payoffs are 0 as by declaring truthfully no investment is made at t. For continuation payoffs, note that since the incentive constraint (8) is also binding at $t + dt$, the uninformed agent at $t + dt$ gets the same expected payoff from truth-telling for all $t' \geq t + dt$ and from pretending to have observed the good signal at $t + dt$. Combining these two remarks, the payoff the uninformed agent gets at t from being truthful for all $t' \geq t$ is the same as what he gets from truth-telling at t and lying at $t + dt$. As a result, the uninformed agent at t is indifferent between (i) claiming to have observed the good signal at t (Lie at t), and (ii) being truthful at t but lying at $t + dt$ if he is still uninformed (Lie at $t + dt$). Table 1 shows the agent's payoffs from both policies for all possible outcomes.

Outcomes	$s^t = 1$	$s^t = 0$	$s^t = \emptyset, \theta = 1$	$s^t = \emptyset, \theta = 0$
Lie at t	$e^{-r\tau(t)}W$	$e^{-r\tau(t)}(-\omega)$	$e^{-r\tau(t)}W$	$e^{-r\tau(t)}(-\omega)$
Lie at $t + dt$	$e^{-r\tau(t+dt)}W$	0	$e^{-r\tau(t+dt)}W$	$e^{-r\tau(t+dt)}(-\omega)$
Probabilities	$p_t\lambda^1 dt$	$(1 - p_t)\lambda^0 dt$	$p_t(1 - \lambda^1 dt)$	$(1 - p_t)(1 - \lambda^0 dt)$

Table 1: Payoffs from two different policies. Under the first policy (Lie at t), the uninformed agent claims that the state is good at t. Under the second policy (Lie at $t + dt$), the uninformed agent claim to be uninformed at t but lies at $t + dt$ if he remains uninformed.

Since the expected payoffs from both policies coincide,

$$e^{-r\tau(t)}\Big(p_tW+(1-p_t)(-\omega)\Big)=e^{-r\tau(t+dt)}\Big(p_tW+(1-p_t)(1-\lambda^0dt)(-\omega)\\ +0\cdot(1-p_t)\lambda^0dt\Big).$$

Equivalently,

$$(1-p_t)\lambda^0dt\omega e^{-r\tau(t)}=(e^{-r\tau(t)}-e^{-r(\tau(t+dt))})\Big(p_tW+(1-p_t)(1-\lambda^0dt)(-\omega)\Big).$$

Rearranging terms, dividing by dt and taking $dt\to 0$, we deduce that

$$(1-p_t)\lambda^0\omega=r\dot{\tau}(t)\Big(p_tW+(1-p_t)(-\omega)\Big), \tag{9}$$

which provides the characterization in Lemma 3.

Equation (9) illustrates how τ balances the costs and benefits of learning for the agent. The left hand side in (9) is the benefit from learning as the agent could avoid investment when the project is bad. The right hand side in (9) is the cost of learning as when no signal arrives the investment is just delayed. An important implication from this characterization is that $\dot{\tau}(t)<1$ and thus the delay with which investment decisions are made, $\tau(t)-t$, is decreasing in t. Intuitively, to motivate the agent to learn, the agent's cost of learning has to be lower than that in the single-player benchmark for the agent and therefore the principal sets $\dot{\tau}(t)<1$. In Section E.3 of Supplementary Material, we show that this feature of decreasing delays is robust when $\lambda^0\geq\lambda^1$.

4.2 Optimal Dynamic Delegation

This subsection characterizes the solutions to the optimal delegation problem and establishes the tradeoff between the amount of information acquired and how effectively it is used.

We first find a solution τ^T to the relaxed problem when $T\leq t^*$. We impose (8) binding everywhere in $[0,T]$. By Lemma 3, (8) binding in $[0,T)$ gives

$$\dot{\tau}^T(t)=(\frac{\lambda^0}{r})\frac{\omega}{W\frac{p_t}{1-p_t}-\omega},\quad t<T. \tag{10}$$

(8) binding at T gives

$$\tau^T(T)=T. \tag{11}$$

These two together give us

$$\tau^T(t) = T - \frac{\lambda^0}{r} \int_t^T \frac{1}{\frac{W}{\omega} \frac{p_s}{1-p_s} - 1} ds, \quad t \leq T. \tag{12}$$

Figure 2 illustrates the solution.

Since $\tau^T(\cdot)$ satisfies the conditions in Lemma 2, it solves the relaxed problem. We now verify that it actually solves the original dynamic delegation problem (1) for a given T. First note that τ^T satisfies (2). Indeed, $\tau^T(t) = \tau^T(T) - \int_t^T \dot{\tau}^T(s)ds$ and, since $\tau^T(T) = T$ and $\dot{\tau}^T(t) < 1$, $\tau^T(t) \geq \tau^T(T) - (T-t) = t$ for all $t \in [0, T]$. Second, τ^T satisfies (5) because it holds with equality over $t \in [0, T]$. Finally, since $\tau^T(t)$ is increasing over $[0, T]$ and $\tau^T(T) = T$, τ^T also satisfies the incentive constraints (3)-(4). As a result, τ^T indeed solves the dynamic delegation problem (1) for a given T. As can be seen, the incentive constraint for the uninformed agent is the key to pinning down the optimal contract when $T \leq t^*$.

The following result provides a key insight for solving for the optimal $T < t^*$.

Proposition 1 *Let $t < T < \hat{T} < t^*$. Then, $\tau^T(t) < \tau^{\hat{T}}(t)$.*

Figure 3 illustrates Proposition 1. Increasing the deadline is beneficial for the principal in that more information is acquired and thus investment in the bad state is less likely to happen. Proposition 1 shows that more learning imposes a nontrivial incentive cost on the principal because when T increases, $\tau^T(t)$ must increase too. This means that when T increases, investments are delayed more when the good signal is received.

Formally, Proposition 1 follows immediately from Equation (12). To better understand Proposition 1, take $t < T < \hat{T}$ and assume for the moment that t is close to T. When the uninformed agent at t faces the contract $\langle T, \tau^T \rangle$, he knows that by declaring truthfully, the investment will be made at T (unless a bad signal is received in the meanwhile). Now, when the uninformed agent at t faces the contract $\langle \hat{T}, \tau^{\hat{T}} \rangle$, the earliest time at which the investment could be made is $\tau^{\hat{T}}(T) > T$. As a result, the expected continuation payoff that the uninformed agent gets at t by being truthful is lower when he faces $\langle \hat{T}, \tau^{\hat{T}} \rangle$ than when he faces $\langle T, \tau^T \rangle$. Therefore, to provide incentives for truthful revelation at t, contract $\langle \hat{T}, \tau^{\hat{T}} \rangle$ must punish the agent even more when he claims a good signal. In other words, $\tau^{\hat{T}}(t) > \tau^T(t)$. This intuition can be iteratively applied backwards to render this property for all $t < T$.

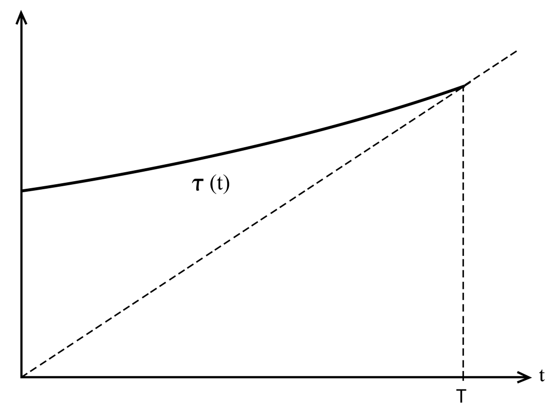

Figure 2: The dark line shows the time at which the investment is made as a function of the time at which the good signal is received. For $t < T$, the investment decision is delayed and the delay, $\tau^T(t) - t$, is decreasing. Parameter values: $\lambda^1 = 5$, $\lambda^0 = 4.8$, $R = r = 0.03$, $p^0 = 0.8$, $v = 120$, $w = 10$, $V = 480$, $W = 1800$, $T = 2.4$.

We now solve the relaxed problem given $T > t^*$ by imposing (8) binding everywhere in $[0, t^*)$ and (7) binding everywhere in $[t^*, T]$. By Lemma 3, (8) binding in $[0, t^*)$ implies

$$\dot{\tau}^T(t) = (\frac{\lambda^0}{r})\frac{\omega}{W\frac{p_t}{1-p_t} - \omega} \quad t \in [0, t^*).$$

Combined with (7) binding for $t \geq t^*$, we have

$$\tau^T(t) = \begin{cases} t^* - \frac{\lambda^0}{r}\int_t^{t^*} \frac{1}{\frac{W}{\omega}\frac{p_s}{1-p_s} - 1} ds & t \leq t^*, \\ t & t > t^*. \end{cases}$$

Lastly, to make sure that τ^T satisfies (8) at t^* and therefore solves the relaxed problem, we need T to be infinity. To see this, notice that at t^*, by being truthful that he has not received a signal, the agent receives the payoff from the policy "invest as soon as a good signal arrives before T and invest at T if no signal arrives before T," which is weakly less preferred to the policy "invest as soon as a good signal arrives and do not invest if no signal arrives." Since at t^* the agent is indifferent between the latter policy and the policy "invest right away," we need $T = \infty$ to ensure incentive compatibility. Therefore a solution to the relaxed problem is

$$\tau^{\infty}(t) = \begin{cases} t^* - \frac{\lambda^0}{r} \int_t^{t^*} \frac{1}{\frac{W}{\omega}\frac{p_s}{1-p_s} - 1} ds & t \leq t^*, \\ t & t > t^*. \end{cases}$$

Since τ^{∞} is increasing and (5) is satisfied everywhere in $[0, \infty)$, it solves the original dynamic delegation problem (1) given that $T > t^*$.

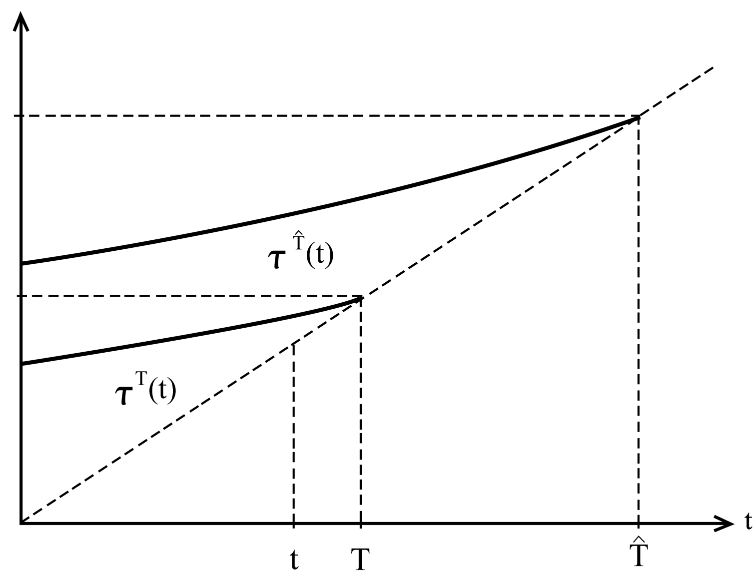

Figure 3: For $t \leq T < \hat{T}$, $\tau^T(t) < \tau^{\hat{T}}(t)$. Parameter values: $\lambda^1 = 5$, $\lambda^0 = 4.8$, $R = r = 0.03$, $p^0 = 0.8$, $v = 120$, $w = 10$, $V = 480$, $W = 1800$, $T = 2.4$, $\hat{T} = 4.3$.

The following theorem summarizes our characterization.

Theorem 1 *The optimal contract takes one of the following two forms:*

(a) There is a deadline $T < t^*$. If a good signal arrives before T, investment happens with a delay. If no signal arrives before T, investment happens at T.

(b) If a good signal arrives before t^*, investment happens with a delay. If a good signal arrives after t^*, investment happens with no delay.

To find the optimal contract $\langle T^*, \tau^{T^*} \rangle$, it suffices to compare the optimal solution when $T \in [0, t^*]$ to the case in which $T = \infty$. It is thus enough to compare the expected payoff for the principal from the optimal τ^T when $T \leq t^*$ to that from τ^{∞}.

The optimal contract can be implemented by setting time-dependent delegation sets illustrated in Figure 4. At any $t < \min\{t^*, T^*\}$, the agent is allowed to commit to invest in $[\tau^{T^*}(t), \infty)$ or just wait and commit later. For $t \geq \min\{t^*, T^*\}$, the agent is granted full freedom.

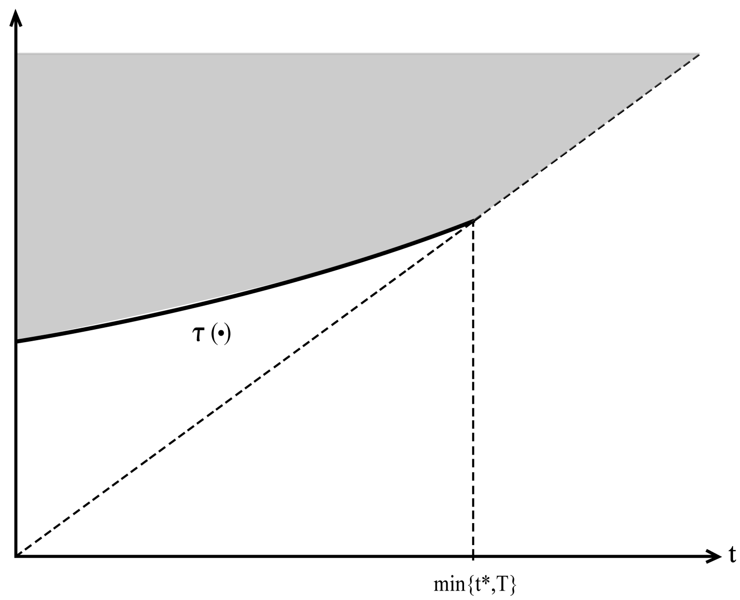

Figure 4: For $t < \min\{t^*, T^*\}$, the agent can choose to invest at any $t' \in [\tau(t), \infty)$. For $t \geq \min\{t^*, T^*\}$, the agent can invest at any $t' > t$. Parameter values: $\lambda^1 = 5$, $\lambda^0 = 4.8$, $R = r = 0.03$, $p^0 = 0.8$, $v = 120$, $w = 10$, $V = 480$, $W = 1800$, $T = 4.3$, $t^* = 7.2$.

4.3 Comparative Statics

We now derive some comparative statics results. These results assume that parameters satisfy Assumption 1, that is, $\frac{W}{\omega} \frac{r}{\lambda^1 + r} \frac{p_0}{1 - p_0} > 1 > \frac{V}{\nu} \frac{R}{\lambda^1 + R} \frac{p_0}{1 - p_0}$.

Proposition 2

(a) Fix all parameters except W and ω. There exist cutoffs $0 < \underline{\kappa} < \bar{\kappa}$ such that for all $W/\omega < \underline{\kappa}$, the optimal contract sets no deadline, whereas for $W/\omega > \bar{\kappa}$ the optimal contract sets a deadline $T^* < t^*$.

(b) Fix all parameters except V and ν. There exist cutoffs $0 < \underline{\eta} < \bar{\eta}$ such that for all $V/\nu < \underline{\eta}$, the optimal contract sets no deadline, whereas for $V/\nu > \bar{\eta}$ the optimal contract sets a deadline $T^* < t^*$.

Part (a) shows that when W/ω is sufficiently small, it is optimal to invest only after the principal has perfectly learned that $\theta = 1$. In this case, t^* is small, so the incentives will become aligned rapidly, and there is no need to significantly delay investments for $t < t^*$. In contrast, when W/ω is large, t^* is large, and the conflict of interest is severe. So decisions to invest need to be significantly distorted for $t < \min\{T^*, t^*\}$. To economize on distortions early in the game, the principal commits to invest at a deadline $T^* < t^*$ even when this means learning stops early.

Part (b) characterizes the solutions as we vary the principal's payoffs V and ν. When V/ν is small, V is small compared to ν and it is relatively costly for the principal to invest when the state is bad. To avoid the costs of a failure, the principal prefers to perfectly learn the state even when this entails significant delays for $t < t^*$. In contrast, when V/ν is large, the cost of a failure is small, and the principal fixes a deadline $T^* < t^*$ that stops learning and reduces distortions for $t < T^*$.

Proposition 2 sheds light on how different organizations should provide incentives for learning. For example, the FDA incurs significant costs when approving bad drugs (so that ν is large). Our results suggest that the FDA should set lengthy revision processes to ensure pharmaceutical companies learn the value of the drugs even if this entails substantial delays between the drugs' discovery and the FDA's final approval. In contrast, the board of a company that is contemplating a partially reversible acquisition (so ν is small) or that cannot align the manager's career incentives (so W/ω is large) should set a deadline $T^* < t^*$ that facilitates truthful communication even at the possible cost of an incorrect decision.

5 Extensions

5.1 Random Mechanisms

To investigate the robustness of the optimal mechanism identified in Section 3, we now allow the principal to choose a random timing of investment. In particular, apart from delaying the decision to invest, now the principal can also commit to never invest with some probability. By varying this probability with the time at which the agent claims a good news, the principal can use the threat of no investment to incentivize learning. We show that as long as the principal is at least as patient as the agent, the optimal deterministic mechanism remains optimal among the family.

Following Pavan, Segal, and Toikka (2014), we focus on random mechanisms which do not allow the agent to update his belief about the outcomes of the randomization until the game ends.[13] In particular, we study the following family of contracts: let $T \in \Re_+ \cup \{\infty\}$ be deterministic. At any $t \leq T$, if the agent announces $m_t = 1$, the principal chooses the investment time according to $q_\tau(\cdot \mid t)$, which is the probability measure from the space $([t, \infty], \mathcal{B}([t, \infty]), q_\tau(\cdot \mid t))$. In addition, if $T < \infty$ and the agent never announces $m_t = 1$ or $m_t = 0$ for any $t \leq T$, the principal chooses the investment time according to q_T, which is the probability measure from the space $([T, \infty], \mathcal{B}([T, \infty]), q_T(\cdot))$. Once the investment time has been determined by q_τ or q_T, the game ends.

This family of random mechanisms allows us to illustrate the connection between deterministic and random mechanisms. As shown in the following proposition, the discount factor $e^{-r\tau(t)}$ resulting from the delay function in a deterministic contract can act as probabilities in a deterministic contract. As a result, the optimal random contract (q_t, q_T, T) can be characterized essentially in the same way. When comparing the principal's payoff from the optimal random contract with that from the optimal deterministic contract, what matters is the relative patience of the players. Since the agent gets the same discounted payoff from the optimal random and deterministic contract, the principal, who is strictly more patient, gets strictly higher payoff from the optimal deterministic contract. In other words, the agent's payoff is discounted equally under the optimal deterministic and random contracts while the principal's payoff is less discounted under the deterministic contract. Therefore when the agent is strictly less patient, delaying the investment is relatively more effective than randomization when it comes to providing incentives. The opposite is true when the agent is strictly more patient.

Proposition 3 *The optimal deterministic contract is weakly better than all random contracts with no leakage iff $R \geq r$.*

5.2 Transfers and Limited Liability

In this subsection, we explore the role of transfers. We show that when the agent faces limited liability, a principal who can transfer money to the agent after a good signal strictly prefers to motivate learning by setting delays instead of transfers. Thus, even

[13] See page 619, "Definition 4 - No Leakage" in Pavan, Segal, and Toikka (2014) for details.

when the principal could incentivize the agent by transfering money, it is optimal not to do it.[14]

A *contract with transfers* is a tuple $\langle T, \tau, q\rangle$ where $T \in \Re_+ \cup \{\infty\}$ and $\tau(\cdot)$ are as in Section 3, while $q\colon [0, T] \to \Re_+$. Contract $\langle T, \tau, q\rangle$ works as described in Section 2, but whenever the agent declares a good signal at $t \leq T$, the principal not only commits to make a decision at $\tau(t)$ but also transfers $q(t) \geq 0$ to the agent at t.[15] This family of contracts allows the principal to subsidize the agent for decisions that are made later in the game. Intuitively, the principal now has delays and transfers to motivate the uninformed agent to learn over $[0, T]$.

The incentive constraints are

$$\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}}\Big(e^{-r\tau(s)}W + e^{-rs}q(s)\Big)ds + \Big(p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}}e^{-rT}W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}e^{-rT}(-\omega)\Big) \geq e^{-r\tau(t)}\Big(p_tW + (1-p_t)(-\omega)\Big) + e^{-rt}q(t) \tag{13}$$

$$e^{-r\tau(t)}W + e^{-rt}q(t) \text{ non increasing} \tag{14}$$

$$e^{-r\tau(t)}W + e^{-rt}q(t) \geq e^{-rT}W \tag{15}$$

Thus, the principal solves

$$\max_{\langle T,\tau,q\rangle} \int_0^T p_0\lambda^1 e^{-\lambda^1 t}(e^{-R\tau(t)}V - e^{-Rt}q(t))dt + p_0e^{-\lambda^1 T}e^{-rT}W + (1-p_0)e^{-\lambda^0 T}e^{-RT}(-\omega) \tag{16}$$

subject to (2)-(13)-(14)-(15). We will proceed as before by fixing $T < t^*$ and exploring the solutions $\langle \tilde{\tau}^T, \tilde{q}^T\rangle$ to the dynamic delegation problem with transfers (16).

Proposition 4 *Take $\tilde{q} \equiv 0$ and $\tilde{\tau} \equiv \tau^T$, where τ^T is the solution to the baseline dynamic delegation problem characterized in (12). Then, $\langle \tilde{\tau}, \tilde{q}\rangle$ is optimal for (16).*

This proposition says that using delays to motivate the agent is a more efficient way than the use of transfers. As the delegation literature has long pointed out,

[14] In Section E of Supplementary Material, we show that when the agent is not subject to limited liability, transfers allow the principal to achieve the first best and extract the full surplus.

[15] We restrict attention to transfers that reward good signals only. The principal could also use transfers that reward no information, or even bad signals. The class of transfers we consider seems appropriate for applications and is a natural candidate to overthrow delays.

transfers are oftentimes unfeasible and therefore there is little room to use them as an incentive devise Holmstrom (1984). We strengthen this observation by showing that even when some forms of transfers may be feasible, a principal facing an agent that privately learns over time should distort decisions rather than using subsidies. In doing so, we provide an important robustness check for our main results in Section 4.

The economics behind Proposition 4 is simple. Take any contract $\langle \tilde{\tau}, \tilde{q} \rangle$ and suppose that there is a positive measure set $A \subseteq [0, T]$ such that $\tilde{q}(t) > 0$ for all $t \in A$. For each $t \in A$, on the (τ, q) plane we can draw the indifference curves crossing $(\tilde{\tau}(t), \tilde{q}(t))$ for both the truthful informed agent (who knows and claims that the state is 1) and for the lying uninformed agent (whose belief is p_t but claims $m^t = 1$) :

$$I_{info} : \{(\tau, q) \mid e^{-r(\tau - t)}W + q = e^{-r(\tilde{\tau}(t) - t)}W + \tilde{q}(t)\}$$

$$I_{uninfo} : \{(\tau, q) \mid e^{-r(\tau - t)}(p_t W + (1 - p_t)(-\omega)) + q = e^{-r(\tilde{\tau}(t) - t)}(p_t W + (1 - p_t)(-\omega)) + \tilde{q}(t)\}$$

As shown in Figure 5, a single crossing property holds and the indifference curve for the informed agent, I_{info}, is steeper than the indifference curve for the lying uninformed agent, I_{uninfo}.

Now, the principal would like to reduce delays and transfers but has to motivate truthful reporting from the uninformed agent. To do that she would like to reward the uninformed agent who decides to learn. Build a new contract $\langle \tau^*, q^* \rangle$ that is identical to $\langle \tilde{\tau}, \tilde{q} \rangle$, but for $t \in A$, $(\tau^*(t), q^*(t))$ is in the indifference curve for the lying uninformed agent, I_{uninfo}, to the left of $(\tilde{\tau}(t), \tilde{q}(t)))$, as in the picture. For the truthful informed agent at t, $(\tau^*(t), q^*(t))$ belongs to an indifference curve I^*_{info} strictly more attractive than I_{info}. This means that for $t' < t$, the new contract provides even stronger incentives to the uninformed agent as his value from following the truthful reporting strategy over $[t', T]$ has increased. The new contract is thus feasible and reduces delays and transfers. The principal therefore should not use transfers to motivate the agent.[16] Problem (16) reduces to our baseline dynamic delegation problem (42) and has τ^T as a solution.

[16] The proof of Proposition 4 builds on this idea but has to deal with the feasibility constraint $\tau(t) \geq t$. In the proof, we show that the feasibility constraint does not bind and therefore our graphical argument goes through.

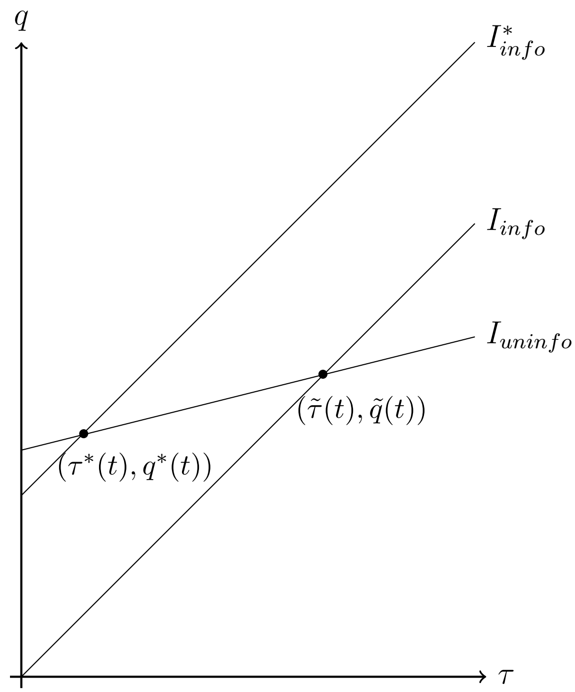

Figure 5: In the optimal contract, transfers are not used.

5.3 Costly Learning Effort

We extend our insights to a model with costly learning. Costly learning captures the idea that the agent needs to incur an unverifiable effort cost to obtain signals. From the principal's perspective, the agent may decide not to make effort and lie about his information. The agency relationship thus becomes a dynamic moral hazard model in which incomplete information evolves over time and transfers are unavailable. This subsection shows how delays and deadlines can induce both truthful revelation and costly learning in this more general environment.

We now assume that to generate signals at any time t, the agent needs to incur a cost $c \cdot dt$ per unit of time. The agent could also decide not to incur the cost, in which case he receives no information. In the baseline model in Section 2, we assumed that $c = 0$. We now explore the case $c > 0$.

The one-person problem for the principal remains unchanged as her payoffs have not been modified. When the agent has decision rights, the solution to his optimization problem is characterized by cutoffs $\bar{p} > \underline{p}$ such that, given the current belief p_t, the agent invests if $p_t \geq \bar{p}$, does not invest but incurs the learning cost if $\bar{p} > p_t \geq \underline{p}$, and does not invest or learn if $p_t < \underline{p}$. We assume that $\bar{p} < p_0 < q^*$. This means that at $t = 0$ the principal would like to wait for information, while the agent would like to invest immediately.

A contract is a tuple $\langle T, \tau \rangle$ that suggests learning effort at any $t < T$ ans asks the agent for his information.[17] If the agent reports $m_t = \emptyset$ for all $t \in \text{dom}(\tau)$, the principal invests at T so $y_T = 1$. Similar to the baseline model, if the agent reports that he has received a good signal at t, then the principal invests at $\tau(t) \geq t$.

We simplify exposition by making two restrictions. First, we focus on strategies for the agent in which the decision to stop experimentation is irreversible. This restriction rules out strategies in which the agent shirks for a period of time but keeps the option to resume effort late in the game by claiming to be uninformed. After stating and discussing our main characterization result in Proposition 5, we show that under the optimal contract the agent would not find profitable to use more complicated strategies and thus our restriction is without loss.

Second, we restrict attention to contracts $\langle T, \tau \rangle$ in which the deadline T is not too large. Concretely, we assume that $T < t^*$ (where t^* was defined in Section 2). This restriction implies that the optimal one-person decision problem for the uninformed agent at T who has incurred the learning cost for all $t \in [0, T]$ (so that his belief is p_T) is to invest at T. When this restriction does not hold, the agent would eventually prefer to incur the learning cost and thus the incentives of the principal and the agent become aligned. We therefore focus on the more interesting case in which along the path of play, interests are misaligned.

Fix a contract $\langle T, \tau \rangle$. Note that if the agent stops learning at $S < T$ when he is uninformed, his optimal response is to declare immediately that he observed the good signal. This is a consequence of the restriction to irreversible stopping decisions, but we relax this restriction after Proposition 5. We thus define $u^{\langle T, \tau \rangle}(t, S)$ as the total expected utility for the agent who is uninformed at t, learns in $[0, S]$, reports truthfully at any given $t' \in [0, S)$, and claims a good signal at $t' = S$:

$$
\begin{aligned}
u^{\langle T,\tau\rangle}(t,S) = & \int_t^S p_t \lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}} \Big(e^{-r\tau(s)} W - c\frac{1-e^{-rs}}{r} \Big) ds \\
& + \int_t^S (1-p_t)\lambda^0 \frac{e^{-\lambda^0 s}}{e^{-\lambda^0 t}} (-c) \frac{1-e^{-rs}}{r} ds + p_t \frac{e^{-\lambda^1 S}}{e^{-\lambda^1 t}} e^{-r\tau(S)} W \\
& + (1-p_t)\frac{e^{-\lambda^0 S}}{e^{-\lambda^0 t}} e^{-r\tau(S)}(-\omega) - c\frac{1-e^{-rS}}{r}\Big(p_t \frac{e^{-\lambda^1 S}}{e^{-\lambda^1 t}} + (1-p_t)\frac{e^{-\lambda^0 S}}{e^{-\lambda^0 t}} \Big).
\end{aligned}
$$

[17] In particular, the message space for the agent is the same as in Section 2.

A key incentive constraint is to ensure that the agent finds it optimal to make the costly effort at all times $t \in \text{dom}(\tau)$. In other words,

$$T \in \arg\max_{S \in \text{dom}(\tau)} u^{\langle T,\tau \rangle}(t, S) \tag{17}$$

for all $t \in \text{dom}(\tau)$. This constraint ensures that at any t, the agent finds attractive to keep effort and truthful reporting over the entire game. The *dynamic delegation problem* is therefore formulated by maximizing (1) over all contracts $\langle T, \tau \rangle$ subject to (2)-(3)-(4)-(17).

We proceed by solving a relaxed delegation problem. In particular, to solve our dynamic delegation problem, we fix T and maximize (1) over $\tau\colon [0,T] \to \Re_+$ subject to the constraint

$$u^{\langle T,\tau \rangle}(t, T) \geq u^{\langle T,\tau \rangle}(t, t) \tag{18}$$

for all $t \in \text{dom}(\tau)$. This constraint relaxes (17) by imposing that the uniformed agent finds optimal to incur the learning effort over $[t, T]$ rather than lying at t to get the investment immediately. The *relaxed problem* maximizes (1) subject to (18) for fixed T.

The following proposition shows that our main qualitative results are preserved when the cost c is low enough.

Proposition 5 *The following hold:*

(a) There exists $\bar{c} = \bar{c}(T) > 0$ such that for all $c < \bar{c}$ the unique solution $\tau^T\colon [0,T] \to \Re_+$ to

$$\tau(T) = T, \quad \dot{\tau}(t) = \frac{(1-p_t)\lambda^0\omega - ce^{r(\tau(t)-t)}}{r(p_t W + (1-p_t)(-\omega))} \quad \forall t \leq T \tag{19}$$

solves the relaxed problem given T.

(b) For all $c < \bar{c}$, $\tau^T\colon [0,T] \to \Re_+$ to (19) solves the dynamic delegation problem for given T and $0 < \dot{\tau}^T(t) < 1$.

(c) Let $t < T < \hat{T}$ and $c < \min\{\bar{c}(T), \bar{c}(\hat{T})\}$. Then, $\tau^T(t) < \tau^{\hat{T}}(t)$.

This proposition shows that the main qualitative properties of the solution in Section 4 extend to the model with costly effort. In particular, part (b) shows that the delay with which investments are made, $\tau^T(t) - t > 0$, is decreasing in t. Finally, part (c) shows that more learning imposes a nontrivial incentive cost on the principal because when T increases, $\tau^T(t)$ must increase too and investments are delayed more when the good signal is received.

Proposition 5 follows from two important insights. First, at any t the uninformed agent should be indifferent between lying to get the investment and stopping learning at any future time. Indeed, in the Appendix we prove that $u^{\langle T,\tau\rangle}(t,S)$ is constant in S. Similar to the intuition behind Lemma 3, this indifference condition results in the ordinary differential equation (19) by noting that the benefit of learning should equal its cost:

$$\underbrace{(1-p_t)\lambda^0\omega e^{-r\tau(t)}}_{\text{Learning benefit}}dt = \underbrace{e^{-r\tau(t)}r\dot{\tau}(t)dt(p_tW+(1-p_t)(-\omega))}_{\text{Delay cost}} + \underbrace{ce^{-rt}dt}_{\text{Effort cost}}. \tag{20}$$

Second, there are two opposing forces in the model with costly learning. One the one hand, to ensure the uninformed agent incurs the costly effort and declares truthfully, the decisions must be delayed. When c is large, it is hard to motivate the agent to make the effort. Indeed, as (20) shows, when c is large the delay cost should be small and eventually negative. To ensure the delay cost is negative, the principal has to commit to $\dot{\tau} < 0$. On the other hand, the principal also needs to ensure the informed agent at t declares truthfully and thus $\dot{\tau} \geq 0$. When c is large enough, it is not feasible to provide incentives to the agent regardless of his information at any t. The restriction $c < \bar{c}$ in part (b) solves this conflict by ensuring that the delay cost need not be negative to ensure incentives for the uninformed agent.

We have assumed the decision to stop learning is irreversible for the agent. When $c < \bar{c}$, the optimal contract actually provides incentives for costly effort and truthful revelation over $[0,T]$ even when the agent can use arbitrary strategies (and is not restricted to irreversible stopping decisions as we have assumed so far). In this version of the model, at any t at which the message history is null (so the agent has declared to be uninformed in the past), the agent can make any effort decision and declare any message.

Proposition 6 *Let $c < \bar{c}$ and $\tau = \tau^T$ be the solution from Proposition 5. When the agent faces mechanism $\langle T,\tau\rangle$ and is allowed to use arbitrary strategies, the agent's optimal strategy is to make effort and declare truthfully at every $t \in [0,T]$ on the path of play.*

This result says that our restriction to strategies in which the decision to stop costly effort is irreversible is actually without loss. To intuitively see this result, suppose that the agent has been on-path over $[0,t]$, has belief p_t at t and is contemplating to declare he is uninformed but not to incur the learning cost at t. After the deviation, the agent will come to period $t+dt$ with belief p_t and has to decide what to do. Had he not deviated, he would have belief p_{t+dt} and would be indifferent between lying and incurring the effort at $t+dt$. Since he deviated, his belief at $t+dt$ is $p_t > p_{t+dt}$ so he now finds strictly optimal to lie and get the investment at $\tau(t+dt)$. If the agent deviates at t, he thus gets payoffs $e^{-r(\tau(t+dt)-t)}(p_tW+(1-p_t)(-\omega))$. But this payoffs is strictly less than the payoff $e^{-r(\tau(t)-t)}(p_tW+(1-p_t)(-\omega))$ that the agent would get by lying at t. Thus, the deviation cannot be optimal.

The proof of Proposition 6 fixes the mechanism $\langle T,\tau\rangle$ and formulate the agent's optimization problem using dynamic programming. The state variables for the agent's optimization problem are the time $t\in[0,T]$ and belief $p\in\{0,1\}\cup[p_t,p_0]$. Let $v(p,t)$ be the optimal value function for the agent before the principal has committed to a decision. For $p\geq p_t$, the value function $v(p,t)$ satisfies

$$
\begin{aligned}
&v(p,t) = \\
&\quad \max\Big\{ -cdt + p\lambda^1 dte^{-r(\tau(t)-t)}W + \Big(1-p\lambda^1dt-(1-p)\lambda^0dt\Big)(1-rdt)v(p+dp,t+dt); \\
&\quad -cdt + p\lambda^1 dte^{-r(\tau(t)-t)}W + (1-p\lambda^1dt-(1-p)\lambda^0dt)e^{-r\tau(t)-t}((p+dp)W+(1-p-dp)(-\omega)); \\
&\qquad e^{-r(\tau(t)-t)}(pW+(1-p)(-\omega)); (1-rdt)v(p,t+dt); 0\Big\}.
\end{aligned}
$$

The first term is the payoff from effort and truthful revelation at t, the second term is the payoff after costly effort but lying if uninformed at the end of round t,[18] the third term is the payoff from no effort and declaring a good signal, the fourth term is the payoff from no effort and declaring being uninformed, and the fifth term is the payoff from no effort and declaring a bad signal. Rearranging terms, we derive the

[18] Obviously, the agent would not want to lie if he receives a good or a bad signal. Note that this term converges to the payoff the agent gets by not incurring effort and declaring a good signal, $e^{-r(\tau(t)-t)}(pW+(1-p)(-\omega))$. This means that the second term in this dynamic programming equation can be ignored.

following Hamilton Jacobi Bellman (HJB) equation: For $p \geq p_t$

$$
\begin{aligned}
&0 = \max\{\\
&-c+p\lambda^1 e^{-r(\tau(t)-t)}W - (r + p\lambda^1 + (1-p)\lambda^0)v(p,t) + v_p(p,t)(\lambda^0 - \lambda^1)p(1-p) + v_t(p,t);\\
&e^{-r(\tau(t)-t)}(pW + (1-p)(-\omega)) - v(p,t); v_t(p,t) - rv(p,t); -v(p,t)\}
\end{aligned}
$$

with boundary conditions $v(p,T) = \max\{0, pW + (1-p)(-\omega)\}$ for all $p \geq p_t$ and $v(0,t) = 0$ for all $t \leq T$. This HJB is a partial differential equation whose solution yields the optimal policy and is characterized in the following lemma.

Lemma 4 *The following hold:*

(a) $v(p,t) = \max\{e^{-r(\tau(t)-t)}(pW + (1-p)(-\omega)); 0\}$ *solves the HJB equation and is the value function for the agent.*

(b) The following strategy is optimal for the agent: When $p > p_t$*, declare a good signal; when* $p = p_t$*, make an effort and declare to be uninformed; when* $p = 0$ *declare a bad signal.*

To prove part (a), we verify that the value $v(p,t)$ in the lemma satisfies the HJB. Proposition 6 follows from part (b). To see part (b) note that since when $p = p_t$, the agent gets exactly the same payoff by being obedient and truthful than by lying. Indeed, τ^T is such that $u^{\langle T,\tau^T\rangle}(t,S)$ is constant in S and, as we prove in the Appendix,

$$
\begin{aligned}
v(p_t,t) = &\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}}\left(e^{-r(\tau(s)-t)}W - c\frac{1-e^{-r(s-t)}}{r}\right)ds\\
&+ \int_t^T (1-p_t)\lambda^0 \frac{e^{-\lambda^0 s}}{e^{-\lambda^0 t}}(-c)\frac{1-e^{-r(s-t)}}{r}ds\\
&+p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}}e^{-r(\tau(T)-t)}W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}e^{-r(\tau(T)-t)}(-\omega)\\
&- c\frac{1-e^{-r(T-t)}}{r}\left(p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}\right).
\end{aligned}
$$

This means that at belief $p = p_t$, $v(p_t,t)$ is the discounted expected payoff the agent gets by incurring effort cost and being truthful over $[t,T]$.

6 Concluding Remarks

This paper studies a dynamic delegation model in which learning is private. Evolving private information shapes the optimal contract in distinctive ways. Indeed, we show that to ensure truthful revelation from the agent, the principal needs to delay investments commonly known to be optimal. As time goes on, the principal grants more flexibility to the agent and, eventually, the agent is free to make any decision. Our analysis uncovers a new tradeoff between how much information is acquired in the relationship and how efficiently new information is used. In sum, our analysis brings out a number of new economic features arising in delegation models with evolving private information.

Our model is stylized. The learning process is assumed to be Poisson,[19] investment is irreversible, and the agent has little freedom to decide how to learn.[20] The model could also be extended to allow for money burning.[21] Our dynamic delegation model with private learning can also be used as a workhorse to explore applied issues in political economy, finance, and organizational economics. We leave this research projects for future work.

[19]Exploring a model with Brownian learning would be interesting, but evolving private information makes the problem hard to analyze. When learning is Brownian delays and deadlines are likely to play a role but the contract may have additional features.

[20]At the other extreme, the agent could decide any experiment that reveals information about the state.

[21]This means that the agent can spend resources that have no value for the principal.

Appendix

This Appendix consists of three parts. Appendix A provides proofs for Section 4.1. Appendix B provides proofs for Section 4.3. Appendix C provides proofs for Section 5.

A Proofs for Section 4.1

Proof of Lemma 1. We prove that $\tau(t) > t$ for $t < \min\{t^*, T\}$. For simplicity, take $t = 0$. By contradiction, assume that $\tau(0) = 0$. The left hand side of (5) can be written as

$$\int_0^T \Big(p_0\lambda^1 \exp(-\lambda^1 s)\exp(-r\tau(s))W\Big)ds + \Big(p_0 \exp(-\lambda^1 T)e^{-rT}W + (1-p_0)\exp(-\lambda^0 T)e^{-rT}(-\omega)\Big) \leq \int_0^T \Big(\lambda^1 \exp(-\lambda^1 s)\exp(-rs)p_0 W\Big)ds + \Big(p_0 \exp(-\lambda^1 T)e^{-rT}W + (1-p_0)\exp(-\lambda^0 T)e^{-rT}(-\omega)\Big)$$

The inequality follows since $\tau(s) \geq s$. The term on the right hand side of the inequality above is the expected payoff that the agent would get following the policy of investing if any good signal is revealed before T and investing at T if no signal is revealed before T. Since $p_0 > p^*$, this policy must result in strictly lower payoffs than the expected payoff from investing at $t = 0$. So,

$$\int_0^T \Big(\lambda^1 \exp(-\lambda^1 s)\exp(-rs)p_0 W\Big)ds + \Big(p_0 \exp(-\lambda^1 T)e^{-rT}W + (1-p_0)\exp(-\lambda^0 T)e^{-rT}(-\omega)\Big) < p_0 W + (1-p_0)(-\omega).$$

Combining these inequalities we deduce that (5) is violated at $t = 0$ when $\tau(0) = 0$. It follows that $\tau(0) > 0$. ■

Proof of Lemma 2. Let τ^* solve the relaxed problem. By way of contradiction, assume that for some $A \subseteq [0, \min\{t^*, T\})$ with positive Lebesgue measure, and for all $t \in A$, the constraint (8) is slack. For $t \in \mathrm{dom}(\tau)$, define

$$\varphi_t = \int_t^T p_t \lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}} e^{-r\tau^*(s)} W ds + \Big(p_t \frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} e^{-rT}W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}} e^{-rT}(-\omega)\Big)$$

Now, define τ' as follows. For $t \notin A$, $\tau'(t) = \tau^*(t)$, while for $t \in A$,

$$e^{-r\tau'(t)}\big(-\omega + p_t(W+\omega)\big) = \varphi_t.$$

For $t \in A$, $e^{-r\tau'(t)} > e^{-r\tau^*(t)}$. Therefore, $\tau'(t) \le \tau^*(t)$ for all $t \in [0, \min\{t^*, T\}]$, with strict inequality for $t \in A$. We claim that τ' is feasible. To see this, note that for all $t \in [0, \min\{t^*, T\}]$

$$\begin{aligned}
&\int_t^T p_t \lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}} e^{-r\tau'(s)} W ds + \Big(p_t \frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} e^{-rT} W + (1-p_t) \frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}} e^{-rT}(-\omega)\Big) \\
&\quad \ge \int_t^T p_t \lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}} e^{-r\tau^*(s)} W ds + \Big(p_t \frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} e^{-rT} W + (1-p_t) \frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}} e^{-rT}(-\omega)\Big) \\
&\quad = \varphi_t \\
&\quad \ge e^{-r\tau'(t)}(-\omega + p_t(W+\omega)) \\
&\quad \ge \max\{0, e^{-r\tau'(t)}(-\omega + p_t(W+\omega))\}.
\end{aligned}$$

The first inequality follows since τ' is below τ^* and the equality is by definition of φ_t. The second inequality follows with equality when $t \in A$ (by definition of τ') and for $t \notin A$ follows since τ' and τ^* coincide and τ^* satisfies 8. The third inequality follows since $t < t^*$. It follows that τ' satisfies (7)-(8) and results in higher expected payoffs than τ^*. This contradicts the optimality of τ^* for the relaxed problem.

We now argue that when $T > t^*$ (case b in the statement of the Proposition), $\tau^*(t) = t$ for almost every $t \in [t^*, T]$. Otherwise, there is a set $A \subseteq [t^*, T]$ of positive measure such that for all $t \in A$, $\tau^*(t) > t$. Construct τ' that coincides with τ^* outside A, but $\tau'(t) = t$ for $t \in A$. It is clear that τ' satisfies (8) since for $t < t^*$, τ' does not change the payoff from lying but increased the payoff from truth-telling. It follows that τ' is feasible for the relaxed problem and results in higher expected payoffs for the principal than τ^*. This is a contradiction.

Now, to prove the converse, we assume that $T > t^*$. The proof of the converse when $T \le t^*$ is analogous. Take τ^* such that (7)-(8) bind almost everywhere. Take τ' that solves the relaxed problem (6). From the first part of this proof, τ' and τ^* coincide for almost every $t \in [t^*, T]$. The previous step also shows that τ' is such that (8) binds for almost every $t \in [0, \min\{t^*, T\}]$. Define

$$u(t) = \int_t^T e^{-\lambda^1 s}(e^{-r\tau^*(s)} - e^{-r\tau'(s)}) ds$$

for $t \in [0, \min\{t^*, T\}]$. Note that $u(t)$ is absolutely continuous and its derivative is defined almost everywhere and equals $-e^{-\lambda^1 t}(e^{-r\tau^*(t)} - e^{-r\tau'(t)})$. Now, using the fact that the constraint binds almost everywhere for both τ' and τ^*, we deduce that for almost every $t \in [0, \min\{t^*, T\}]$,

$$-p_t \lambda^1 W u(t) = u'(t)\Big(-\omega + p_t(W+\omega)\Big)$$

and $u(\min\{t^*, T\}) = 0$. It follows that for almost every $t \in [0, \min\{t^*, T\}]$, $\frac{d}{dt}\left(u(t)e^{\int_0^t H(s)ds}\right) = 0$ where H is a continuous function. Since $u(\min\{t^*, T\}) = 0$, $u(t) = 0$ for all $t \in [0, \min\{t^*, T\}]$. In particular, $0 = u'(t) = -e^{-\lambda^1 t}(e^{-r\tau^*(t)} - e^{-r\tau'(t)})$ almost everywhere and therefore τ' and τ^* coincide for almost every $t \in [0, \min\{t^*, T\}]$. Since τ^* satisfies (7)-(8), τ^* solves the relaxed problem. ■

Proof of Lemma 3. Since (5) is binding for all $t \in [0, \min\{t^*, T\})$,

$$\int_t^T \lambda^1 e^{-\lambda^1 s} e^{-r\tau(s)} W ds + \left(e^{-\lambda^1 T} e^{-rT} W + \frac{1-p_t}{p_t} e^{(\lambda^0 - \lambda^1)t} e^{-\lambda^0 T} e^{-rT}(-\omega)\right)$$
$$= e^{-\lambda^1 t} e^{-r\tau(t)} (W + \frac{1-p_t}{p_t}(-\omega))$$

where we use the fact that $t \leq t^*$. Since the left hand side of this equation and p_t are differentiable, so is τ. Taking derivatives and using the fact that $\frac{d}{dt}\left(\frac{1-p_t}{p_t}\right) = (\lambda^1 - \lambda^0)\frac{1-p_t}{p_t}$, we deduce that

$$-\lambda^1 e^{-\lambda^1 t} e^{-r\tau(t)} W = -(\lambda^1 + r\dot{\tau}(t)) e^{-\lambda^1 t - r\tau(t)} \left(W + \frac{1-p_t}{p_t}(-\omega)\right)$$
$$+ e^{-\lambda^1 t - r\tau(t)} (\lambda^1 - \lambda^0) \frac{1-p_t}{p_t}(-\omega).$$

Solving for $\dot{\tau}(t)$, we deduce that

$$\dot{\tau}(t) = (\frac{\lambda^0}{r}) \frac{\omega}{W\frac{p_t}{1-p_t} - \omega}.$$

The slope of τ is nonnegative. To see that τ is convex, note that $p_t/(1-p_t)$ is non-increasing and thus $\dot{\tau}$ is non-decreasing. To see that $\dot{\tau}$ is less than 1, note that

$$\dot{\tau} < 1 \quad \text{iff} \quad 1 < \frac{W}{\omega} \frac{r}{\lambda^0 + r} \frac{p_t}{1-p_t}.$$

To verify this last property, note that investing at t results in higher expected payoffs for the agent than learning at t and investing at $t + dt$ unless the bad state is revealed. That is,

$$p_t W + (1-p_t)(-\omega) \geq (1-p_t)(1 - \lambda^0 dt) e^{-rdt}(-\omega) + p_t e^{-rdt} W.$$

Reordering terms and taking $dt \to 0$, we deduce that $1 < \frac{W}{\omega} \frac{r}{\lambda^0 + r} \frac{p_t}{1-p_t}$. ■

B Proofs for Section 4.3

Proof of Proposition 2. We first prove part (a). Note that the principal's expected payoff from setting $T = \infty$ equals

$$\varphi(\frac{W}{\omega}) = \int_0^{t^*} p_0 \lambda^1 e^{-\lambda^1 s} e^{-R\tau^{t^*}_{\frac{W}{\omega}}(s)} V ds + \int_{t^*}^{\infty} p_0 \lambda^1 e^{-\lambda^1 s} e^{-Rs} V ds$$

where $\tau^{t^*}_{\frac{W}{\omega}}(s) = t^* - (\frac{\lambda^0}{r}) \int_s^X \frac{1}{\frac{W}{\omega}\frac{p_x}{1-p_x} - 1} dx$. We claim that for all $\epsilon > 0$, there exists L such that for all $\frac{W}{\omega} > L$, $\varphi(\frac{W}{\omega}) < \epsilon$.

First, notice that since $t^* \to \infty$ as $\frac{W}{\omega} \to \infty$, there exists L_1 such that for all $\frac{W}{\omega} > L_1$,

$$\int_{t^*}^{\infty} p_0 \lambda^1 e^{-\lambda^1 s} e^{-Rs} V ds < \epsilon/2.$$

Now we show that there exists L_2 such that for all $\frac{W}{\omega} > L_2$,

$$\int_0^{t^*} p_0 \lambda^1 e^{-\lambda^1 s} e^{-R\tau^{t^*}_{\frac{W}{\omega}}(s)} V ds < \frac{\epsilon}{2}.$$

To show this, we first show that for any δ, there exists L_3 such that for all $\frac{W}{\omega} > L_3$,

$$\frac{1}{\frac{W}{\omega}\frac{p_x}{1-p_x} - 1} < \delta, \forall x \in [0, t^*].$$

Since $\frac{p_x}{1-p_x}$ decreases in x, suffices to show

$$\frac{1}{\frac{W}{\omega}\frac{p_0}{1-p_0} - 1} < \delta.$$

This is done by letting $L_3 = \frac{\delta+1}{\delta}\frac{2(1-p_0)}{p_0}$.

Given this, we now show that for any η, there exists L_4 such that $\frac{W}{\omega} > L_4$ implies

$$e^{-R\tau(s)} < \eta, \forall s \in [0, t^*].$$

To show this, first we notice that $\tau(s)$ increases in s, so it suffices to show that there exists L_4 such that $\frac{W}{\omega} > L_4$ implies

$$e^{-R\left[t^* - \frac{\lambda^0}{r}\int_0^{t^*} \frac{1}{\frac{W}{\omega}\frac{p_x}{1-p_x} - 1} dx\right]} < \eta.$$

In other words,

$$t^* - \frac{\lambda^0}{r}\int_0^{t^*} \frac{1}{\frac{W}{\omega}\frac{p_s}{1-p_s} - 1}\, ds > \frac{\ln \eta}{-R}.$$

Given what we showed in the previous step, we can find L_3 such that

$$\frac{1}{\frac{W}{\omega}\frac{p_x}{1-p_x} - 1} < \frac{r}{2\lambda^0}, \forall x \in [0, t^*].$$

Therefore

$$\begin{aligned} t^* - \frac{\lambda^0}{r}\int_0^{t^*} \frac{1}{\frac{W}{\omega}\frac{p_x}{1-p_x} - 1}\, dx &> t^* - \frac{\lambda^0}{r}\int_0^{t^*} \frac{r}{2\lambda^0}\, dx \\ &= t^* - \frac{\lambda^0}{r}\frac{r}{2\lambda^0}t^* \\ &= \frac{t^*}{2} \to \infty \end{aligned}$$

as $\frac{W}{\omega} \to \infty$. We have therefore shown that there exists L_4 such that $\frac{W}{\omega} > L_4$ implies

$$e^{-R\tau(s)} < \eta, \forall s \in [0, t^*].$$

Now find L_4 such that

$$e^{-R\tau(s)} < \frac{\epsilon}{4Vp_0}, \forall s \in [0, t^*].$$

Therefore

$$\begin{aligned} \int_0^{t^*} p_0\lambda^1 e^{-\lambda^1 s} e^{-R\tau(s)} V\, ds &< \int_0^{t^*} p_0\lambda^1 e^{-\lambda^1 s}\frac{\epsilon}{4Vp_0} V\, ds \\ &= p_0\lambda^1 \frac{\epsilon}{4Vp_0} V \int_0^{t^*} e^{-\lambda^1 s}\, ds \\ &= p_0\lambda^1 \frac{\epsilon}{4Vp_0} V \cdot \frac{1}{\lambda^1}(1 - e^{-\lambda^1 t^*}) \\ &= p_0 \frac{\epsilon}{4Vp_0} V(1 - e^{-\lambda^1 t^*}) \\ &= \frac{\epsilon}{4}(1 - e^{-\lambda^1 t^*}) \\ &< \frac{\epsilon}{2}. \end{aligned}$$

Therefore, for $\frac{W}{\omega} > L_2 := \max\{L_3, L_4\}$, we have

$$\int_0^{t^*} p_0\lambda^1 e^{-\lambda^1 s} e^{-R\tau^{t^*}_{\frac{W}{\omega}}(s)} V ds < \frac{\epsilon}{2}.$$

Lastly, letting $L := \max\{L_1, L_2\}$, we then have

$$\varphi(\frac{W}{\omega}) < \epsilon$$

for $\frac{W}{\omega} > L$.

Now, note that by setting an optimal deadline $T \in [0, t^*]$, the principal's payoff equals

$$\Phi(\frac{W}{\omega}) = \max_{T\in[0,t^*]} \Phi(\frac{W}{\omega}, T)$$

where

$$\Phi(\frac{W}{\omega}, T) = \int_0^T p_0\lambda^1 e^{-\lambda^1 s} e^{-R\tau^T_{\frac{W}{\omega}}(s)} V ds + \Big(p_0 e^{-\lambda^1 T} e^{-RT} V + (1-p_0)e^{-\lambda^0 T} e^{-RT}(-\nu)\Big)$$

Note that

$$\begin{aligned}\Phi(\frac{W}{\omega}, T) &> p_0 e^{-RT}\Big(1 - e^{-\lambda^1 T}\Big)V + \Big(p_0 e^{-\lambda^1 T} e^{-RT} V + (1-p_0)e^{-\lambda^0 T} e^{-RT}(-\nu)\Big) \\ &= e^{-RT}\Big(p_0 V + (1-p_0)e^{-\lambda^0 T}(-\nu)\Big)\end{aligned}$$

Fix any T such that the expression above is strictly positive and equals $\eta > 0$. Let $\epsilon = \eta/2$ and take $W/\omega > L$ such that $\varphi(W/\omega) < \epsilon = \eta/2$ and $T < t^*$. In particular,

$$\Phi(\frac{W}{\omega}) \geq \eta > \eta/2 \geq \varphi(\frac{W}{\omega})$$

which proves that there exists some $\bar{\kappa}$ such that for all $\frac{W}{\omega} > \bar{\kappa}$, $T \in [0, t^*]$ results in higher payoffs than $T = \infty$.

To complete the proof of part (a), note that as W/ω goes to x where $x\frac{r}{\lambda^1+r}\frac{p_0}{1-p_0} = 1$, $t^* \to 0$. In particular,

$$\varphi(\frac{W}{\omega}) \to \int_0^\infty p_0\lambda^1 e^{-\lambda^1 s} e^{-Rs} V ds = p_0 V \frac{\lambda^1}{\lambda^1 + R}$$

whereas

$$\Phi(\frac{W}{\omega}) \to p_0 V + (1-p_0)(-\nu).$$

Since $p_0 V \frac{\lambda^1}{\lambda^1+R} > p_0 V + (1-p_0)(-\nu)$, there exists $\underline{\kappa}$ such that for all $W/\omega < \underline{\kappa}$, $\varphi(W/\omega) > \Phi(W/\omega)$.

To prove part (b), we normalize the principal's expected payoffs by ν and write

$$\varphi(\frac{V}{\nu}) = \int_0^{t^*} p_0\lambda^1 e^{-\lambda^1 s} e^{-R\tau^{t^*}_{\frac{W}{\omega}}(s)} \frac{V}{\nu} ds + \int_{t^*}^\infty p_0\lambda^1 e^{-\lambda^1 s} e^{-Rs} \frac{V}{\nu} ds$$

for the principal's payoff when $T = \infty$ and

$$\Phi(\frac{V}{\nu}, T) = \int_0^T p_0 \lambda^1 e^{-\lambda^1 s} e^{-R\tau^T_{\frac{W}{\omega}}(s)} \frac{V}{\nu} ds + \Big(p_0 e^{-\lambda^1 T} e^{-RT} \frac{V}{\nu} + (1-p_0) e^{-\lambda^0 T} e^{-RT}(-1) \Big)$$

for the principal payoff when setting $T < t^*$. Note that as $V/\nu \to 0$,

$$\varphi(\frac{V}{\nu}) \to 0, \quad \Phi(\frac{V}{\nu}, T) \to -(1-p_0) e^{-\lambda^0 T} e^{-RT}$$

Since Φ is continuous in $(\frac{V}{\nu}, T)$, there exists $\underline{\eta} > 0$ such that for all $V/\nu < \underline{\eta}$,

$$\varphi(\frac{V}{\nu}) > \max_{T \in [0, t^*]} \Phi(\frac{V}{\nu}, T)$$

and thus it is optimal for the principal to set $T = \infty$.

To complete part (b), define y such that $1 = y \frac{R}{\lambda^1 + R} \frac{p_0}{1-p_0}$. By definition,

$$\varphi(y) < \max_{T \in [0, t^*]} \Phi(y, T)$$

where the maximum on the right is attained at $T = 0$. By continuity, there exists $\bar{\eta} < y$ such that for all $V/\nu > \bar{\eta}$,

$$\varphi(\frac{V}{\nu}) < \max_{T \in [0, t^*]} \Phi(\frac{V}{\nu}, T)$$

and the principal sets a deadline $T < t^*$. ■

C Proofs for Section 5

Proof of Proposition 3. The principal's constrained maximization problem is as follows:

$$\max_{T \in \Re_+ \cup \{\infty\}, q_\tau(\cdot|t), q_T(\cdot)} \int_0^T p_0 \lambda^1 e^{-\lambda^1 s} \Big[\int_s^\infty e^{-R\tau} q_\tau(d\tau \mid s) \Big] V ds$$
$$+ \int_T^\infty e^{-R\tau} q_T(d\tau) \Big[p_0 e^{-\lambda^1 T} V - (1-p_0) e^{-\lambda^0 T} \nu \Big]$$

subject to

$$W \Big[\int_t^\infty e^{-r\tau} q_\tau(d\tau \mid t) \Big] \geq W \Big[\int_s^\infty e^{-r\tau} q_\tau(d\tau \mid s) \Big] \quad \forall s \geq t \tag{21}$$

$$W \Big[\int_t^\infty e^{-r\tau} q_\tau(d\tau \mid t) \Big] \geq W \int_T^\infty e^{-r\tau} q_T(d\tau), \forall t. \tag{22}$$

$$\begin{aligned}&\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}} W\Big[\int_s^\infty e^{-r\tau} q_\tau(d\tau \mid s)\Big] ds \\ &\quad + \int_T^\infty e^{-r\tau} q_T(d\tau)\Big[p_t e^{-\lambda^1(T-t)} W - (1-p_t)e^{-\lambda^0(T-t)}\omega\Big] \\ &\geq \max\left\{\Big[\int_t^\infty e^{-r\tau} q_\tau(d\tau \mid t)\Big]\cdot\Big[-\omega + p_t(W+\omega)\Big], 0\right\}, \forall t.\end{aligned} \tag{23}$$

We fix T, ignore condition (21) and (22) and argue that condition (23) should bind for almost all $t < t^*$. Suppose the strict inequality holds for all $t \in A$, where A has positive measure. Let us consider

$$q_\tau^\epsilon(\cdot \mid t) = \begin{cases}(1-\epsilon)q_\tau(\cdot \mid t) + \epsilon \mathbb{1}_t, & \text{if } t \in A \\ q_\tau(\cdot \mid t), & \text{otherwise.}\end{cases}.$$

For ϵ sufficiently small, the strictly inequality still holds at t. For $s < t$, the incentive is strengthened. For $s > t$, incentive is unaffected. Since q_τ^ϵ first-order stochastic dominates q_τ and $e^{-R\tau}$ is decreasing in τ, the principal receives strictly higher payoff under q_τ^ϵ.

Now let us impose (23) binding for all $t < t^*$ while letting (22) and (23) hold for $t = T$ simultaneously. Moreover, for any $t \geq 0$, let $\tau(t)$ be such that $e^{-r\tau(t)} = \int_t^\infty e^{-r\tau} q_\tau(d\tau \mid t)$ and $e^{-rS} = \int_T^\infty e^{-r\tau} q_T(d\tau)$. We then have

$$\begin{aligned}&\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}} W e^{-r\tau(s)} ds + e^{-rS}\Big[p_t e^{-\lambda^1(T-t)} W - (1-p_t)e^{-\lambda^0(T-t)}\omega\Big] \\ &= e^{-r\tau(t)}\cdot\Big[-\omega + p_t(W+\omega)\Big], \forall t \in [0, t^*).\end{aligned} \tag{24}$$

$$e^{-rS} = e^{-r\tau(T)} \tag{25}$$

(24) and (25) combined give us

$$\int_t^\infty e^{-r\tau} q_\tau^*(d\tau \mid t) = e^{-r\tau(t)} = e^{-r(S-T)} e^{-r\tau^{T*}(t)} \quad \forall t < t^*, \tag{26}$$

where τ^{T*} is the optimal deterministic contract given deadline T. It is then easy to see that q_τ^* satisfies the other conditions as well and therefore is feasible. Therefore it solves the original problem.

Note that any q_τ^* that satisfies (26) must not assign probability 1 to $\{t\}$ for $t < \min\{t^*, T\}$. If this is the case, then

$$\int_0^\infty e^{-r\tau} q_\tau^*(d\tau \mid t) = e^{-rt} > e^{-r\tau^{T^*}(t)} \geq e^{-r(S-T)} e^{-r\tau^{T*}(t)},$$

a contradiction.

We first consider the case where $R \leq r$, that is, the principal is more patient than the agent. The principal's payoff equals

$$\begin{aligned}
&\int_0^T p_0\lambda^1 e^{-\lambda^1 s}\Big[\int_s^\infty e^{-R\tau} q_\tau^*(d\tau \mid s)\Big] V ds + \int_T^\infty e^{-R\tau} q_T(d\tau)\Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big] \\
&\leq \int_0^T p_0\lambda^1 e^{-\lambda^1 s}\Big[\int_s^\infty e^{-r\tau} q_\tau^*(d\tau \mid s)\Big]^{\frac{R}{r}} V ds + \Big[\int_T^\infty e^{-r\tau} q_T(d\tau)\Big]^{\frac{R}{r}}\Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big] \\
&= \int_0^T p_0\lambda^1 e^{-\lambda^1 s}\Big[e^{-r(S-T)} e^{-r\tau^{T*}(s)}\Big]^{\frac{R}{r}} V ds + (e^{-rS})^{\frac{R}{r}}\Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big] \\
&= \int_0^T p_0\lambda^1 e^{-\lambda^1 s} e^{-R\tau^{T*}(s)} e^{-R(S-T)} V ds + e^{-RS}\Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big] \\
&\leq \int_0^T p_0\lambda^1 e^{-\lambda^1 s} e^{-R\tau^{T*}(s)} V ds + e^{-RT}\Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big]
\end{aligned}$$

The first inequality is due to Jensen's inequality. The next equality is from (26). We have thus shown that for any given T, the optimal random contract is at most as good as the optimal deterministic contract with under the same T.

Now let us consider the case where $R > r$. We show that for any given T, the optimal random contract given with the deterministic deadline T is strictly better than the optimal deterministic contract. To see this, note that the principal's payoff under this random mechanism equals

$$\begin{aligned}
&\int_0^T p_0\lambda^1 e^{-\lambda^1 s}\Big[\int_s^\infty e^{-R\tau} q_\tau^*(d\tau \mid s)\Big] V ds + \Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big] e^{-RT} \\
&> \int_0^T p_0\lambda^1 e^{-\lambda^1 s}\Big[\int_s^\infty e^{-r\tau} q_\tau^*(d\tau \mid s)\Big]^{\frac{R}{r}} V ds + \Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big] e^{-RT} \\
&= \int_0^T p_0\lambda^1 e^{-\lambda^1 s}\Big[e^{-r\tau^{T*}(s)}\Big]^{\frac{R}{r}} V ds + \Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big] e^{-RT} \\
&= \int_0^T p_0\lambda^1 e^{-\lambda^1 s} e^{-R\tau^{T*}(s)} V ds + \Big[p_0 e^{-\lambda^1 T} V - (1-p_0)e^{-\lambda^0 T}\nu\Big] e^{-RT}.
\end{aligned}$$

To see that the optimal random contract featuring q_τ^* and a deterministic deadline is indeed random, notice that if not, then by (26), $q_\tau^*(t)$ must assign probability 1 to $\tau^{T*}(t)$ for all t. Then the random contract is identical to τ^{T*}, a contradiction. ∎

Proof of Proposition 4. We first consider a relaxed problem. The *relaxed problem* maximizes (16) subject to (13)-(15) given T. Suppose that τ and q solve the relaxed problem. We argue that $q(t) = 0$ for almost every $t \in [0, T]$. Suppose that $q(t) > 0$ for all $t \in A$, where $A \subseteq [0, T]$ has positive measure. For $t \notin A$, define $q'(t) = q(t)$ and $\tau(t) = \tau'(t)$. For $t \in A$, define $q'(t) < q(t)$ (close enough to $q'(t)$) and $\tau'(t)$ such that[22]

$$e^{-r\tau'(t)}\Big(p_tW+(1-p_t)(-\omega)\Big)+e^{-rt}q'(t)=e^{-r\tau(t)}\Big(p_tW+(1-p_t)(-\omega)\Big)+e^{-rt}q(t).$$

This means that when the uninformed agent deviates he gets the same payoff under both the original and the new contract. In particular, the informed agent at t finds contract τ' and q' more attractive than τ and q:

$$e^{-r\tau'(t)}W+e^{-rt}q'(t)\geq e^{-r\tau(t)}W+e^{-rt}q(t) \tag{27}$$

with strict inequality for $t\in A$. We now argue that τ' and q' satisfy all the constraints in the relaxed problem. To verify (13), note that due to (27), the expected payoff that the uninformed agent gets from truthful revelation under τ' and q' is strictly larger than under τ and q due to (27)). Thus, since the uninformed agent at t finds τ' and q' as attractive as τ and q, (13) holds. Analogously, to verify (15), we note again that

$$e^{-r\tau'(t)}W+e^{-rt}q'(t)\geq We^{-r\tau(t)}+e^{-rt}q(t)\geq e^{-rT}W.$$

It follows that τ' and q' are feasible for the relaxed problem. Since $\tau'\leq\tau$ and $q'\leq q$ with strict inequality in some set of positive measure, it follows that τ and q cannot be optimal for the principal.

We can therefore solve the relaxed problem by setting $q\equiv 0$. Such relaxed problem has solution τ^T as we described in Section 4 that satisfy all the constraints in our original problem (16). ■

Proof of Proposition 5. We first argue part (a). Similar to Lemma 1, equation (18) implies that $\tau(t)>t$. Moreover, similar to Lemma 2, (18) binds for $t<T$ for any optimal solution. Taking derivatives, the fact that (18) binds is equivalent to

$$\tau(T)=T,\quad \dot{\tau}(t)=\frac{(1-p_t)\lambda^0\omega-ce^{r(\tau(t)-t)}}{r(p_tW+(1-p_t)(-\omega))}\quad\forall t\leq T. \tag{28}$$

We first argue that there exists $\hat{c}(T)>0$ such that for all $c<\hat{c}(T)$, any solution $\tau(t)$ to (28) is bounded. To see this, note first that $\dot{\tau}(t)\leq\frac{\lambda^0\omega(1-p_t)}{r(p_tW+(1-p_t)(-\omega))}<1$ and, since $\tau(T)=T$, $\tau(t)\geq 0$ for all $t\in[0,T]$. Now, from (28),

$$\dot{\tau}(t)e^{-r\tau(t)}\geq-\frac{c}{r}e^{-rt}\frac{1}{(p_tW+(1-p_t(-\omega))}$$

[22] Note that we can always pick $q'(t)<q(t)$ close enough to $q(t)$ so that

$$e^{-r\tau(t)}+\frac{e^{-rt}-e^{-rt}}{p_tW+(1-p_t)(-\omega)}(q(t)-q'(t))<1$$

and thus $\tau'(t)$ is well defined.

and therefore, $\frac{d}{dt}\left(e^{-r\tau(t)}\right) \leq c\frac{e^{-rt}}{p_tW+(1-p_t)(-\omega)} \leq c \cdot \max_{t\in[0,t^*]} \frac{e^{-rt}}{p_tW+(1-p_t)(-\omega)}$. We thus deduce that

$$\tau(t) \leq \frac{1}{r}\ln\left(e^{-rT} - c\kappa T\right)$$

with $\kappa := \max_{t\in[0,t^*]} \frac{e^{-rt}}{p_tW+(1-p_t)(-\omega)}$. The claim follows by taking $\hat{c} = \frac{e^{-rT}}{\kappa T} > 0$.

Define

$$\Phi(\tau, t, c) = \frac{(1-p_t)\lambda^0\omega - ce^{r(\tau(t)-t)}}{r(p_tW + (1-p_t)(-\omega))}.$$

Being continuously differentiable, Φ is a Lispchitz function over any bounded domain. Theorem 7.4 in (Coddington and Levinson, 1955) implies that the ordinary differential equation (28) has a unique solution τ^T that therefore solves the relaxed problem. Moreover, from Theorem 7.4 in (Coddington and Levinson, 1955), any solution $\tau^{T,c}(t)$ to (28) (where we emphasize the dependence of the solution on c) is continuous in (c, t). In particular, since (28) holds, the derivative $\frac{\partial\tau}{\partial t}$ is continuous in (c, t). When $c = 0$, $0 < \frac{\partial\tau}{\partial t} < 1$ for all $t \in [0, T]$. As a result, we take $\bar{c} \in]0, \hat{c}[$ such that for all $t \in [0, T]$, $\tau^{T,c}(t)$ is increasing and its slope is less than 1.

To prove part (b), we note that for $c < \bar{c}$, τ^T satisfies (2)-(3)-(4). It remains to show that τ^T satisfies (17). To see this, note that for all $t \in \text{dom}(S)$ and all $S \in \text{dom}(\tau^T)$

$$\begin{aligned}
&\frac{\partial u^{\langle T,\tau^T\rangle}(t,S)}{\partial S} \\
&= (1-p_t)\frac{\lambda^0}{e^{-\lambda^0 t}}e^{-r\tau^T(S)}\omega + p_t\frac{e^{-\lambda^1 S}}{e^{-\lambda^1 t}}(-rW)\dot{\tau}^T(S)e^{-r\tau^T(S)} \\
&+(1-p_t)\frac{e^{-\lambda^0 S}}{e^{-\lambda^0 t}}(r\omega)\dot{\tau}^T(S)e^{-r\tau^T(S)} - ce^{-rS}\left(p_t\frac{e^{-\lambda^1 S}}{e^{-\lambda^1 t}} + (1-p_t)\frac{e^{-\lambda^0 S}}{e^{-\lambda^0 t}}\right)
\end{aligned}$$

which equals 0 iff for all $t \in [0, T]$

$$\dot{\tau}^T(t) = \frac{\lambda^0\omega(1-p_t) - ce^{r(\tau(t)-t)}}{r(p_tW + (1-p_t)(-\omega))}.$$

Since τ^T satisfies (19), $\frac{\partial u^{\langle T,\tau^T\rangle}(t,S)}{\partial S} = 0$. Thus (17) holds and τ^T solves the dynamic delegation problem.

To prove part (c), note that the slope of $\tau^{\hat{T}}$ is less than 1. Since $\tau^T(T) = T$, and $\tau^{\hat{T}}(\hat{T}) = \hat{T}$, it follows that $\tau^T(T) < \tau^{\hat{T}}(T)$.

We now prove that $\tau^{\hat{T}}(t) > \tau^T(t)$ for all $t \in [0, T]$. Otherwise, there exists $t \in [0, T]$ such that $\tau^{\hat{T}}(t) \leq \tau^T(t)$. Take $\bar{t} = \max\{t \in [0, T] \mid \tau^{\hat{T}}(t) \leq \tau^T(t)\}$, which exists since $\tau^{\hat{T}}$ and τ^T are continuous and $[0, T]$ is compact. Since $\tau^{\hat{T}}(T) > \tau^T(T)$, $\bar{t} < T$ and for all $t \in (\bar{t}, T]$, $\tau^{\hat{T}}(t) > \tau^T(t)$. Moreover, $\tau^{\hat{T}}(\bar{t}) = \tau^T(\bar{t})$ since otherwise $\bar{t}$ would not be the greatest time at which $\tau^{\hat{T}}$ is below τ^T. There must exists $t' \in [\bar{t}, T]$ such that

$\dot{\tau}^{\hat{T}}(t') > \dot{\tau}^T(t')$. Otherwise, for all $t' \in [\bar{t}, T]$,

$$\dot{\tau}^{\hat{T}}(t') \leq \dot{\tau}^T(t') \tag{29}$$

and thus

$$\tau^{\hat{T}}(T) = \tau^{\hat{T}}(\bar{t}) + \int_{\bar{t}}^{T} \dot{\tau}^{\hat{T}}(t')dt' \leq \tau^T(\bar{t}) + \int_{\bar{t}}^{T} \dot{\tau}^T(t')dt' = \tau^T(T)$$

where the inequality follows from (29). This contradicts the fact that $\tau^{\hat{T}}(T) > \tau^T(T)$. Take $t' \in [\bar{t}, T]$ such that $\dot{\tau}^{\hat{T}}(t') > \dot{\tau}^T(t')$. Since both $\tau^{\hat{T}}$ and τ^T satisfy (28), it follows that $\tau^{\hat{T}}(t') < \tau^T(t')$. This contradicts the fact for all $t \in [\bar{t}, T]$, $\tau^{\hat{T}}(t) \geq \tau^T(t)$. ■

Proof of Lemma 4. To prove part (a), we will show that $v(p,t) = \max\{e^{-r(\tau(t)-t)}(pW + (1-p)(-\omega)), 0\}$ satisfies the HJB equation. The result will follow after we verify that each of the terms inside the max is less than or equal to 0 (since by definition of v the first term is actually 0). Since $\dot{\tau}(t) \geq 0$ and $v(p,t) \geq 0$

$$v_t(p,t) - rv(p,t) = -r(\dot{\tau}(t) - 1)v(p,t) - rv(p,t) = -r\dot{\tau}(t)v(p,t) \leq 0.$$

We also note that the condition

$$-c+p\lambda^1 e^{-r(\tau(t)-t)}W-(r+p\lambda^1+(1-p)\lambda^0)v(p,t)+v_p(p,t)(\lambda^0-\lambda^1)p(1-p)+v_t(p,t) \leq 0$$

can be written as

$$\begin{aligned}&-c + p\lambda^1 e^{-r(\tau(t)-t)}W - (r + p\lambda^1 + (1-p)\lambda^0)e^{-r(\tau(t)-t)}(pW + (1-p)(-\omega)) \\ &\quad + e^{-r(\tau(t)-t)}(W+\omega)(\lambda^0 - \lambda^1)p(1-p) + (pW + (1-p)(-\omega))(-r)(\dot{\tau}(t) - 1)e^{-r(\tau(t)-t)} \\ &\quad \leq 0.\end{aligned}$$

This is equivalent to

$$\dot{\tau}(t) \geq \frac{\omega\lambda^0(1-p) - ce^{r(\tau(t)-t)}}{r(pW + (1-p)(-\omega))}. \tag{30}$$

Since τ satisfies (19) and $p \geq p_t$, (30) holds. It follows that $v(p,t) = \max\{pW + (1-p)(-\omega), 0\}$ satisfies the HJB equation. The fact that $v(p,t)$ is actually the value function for the agent's problem follows from a verification theorem.

We now prove part (b). Since τ is defined so that (18) binds,

$$\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}}\Big(e^{-r\tau(s)}W - c\frac{1-e^{-rs}}{r}\Big)ds + \int_t^T (1-p_t)\lambda^0 \frac{e^{-\lambda^0 s}}{e^{-\lambda^0 t}}(-c)\frac{1-e^{-rs}}{r}ds$$
$$+ p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}}e^{-r\tau(T)}W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}e^{-r\tau(T)}(-\omega) - c\frac{1-e^{-rT}}{r}\Big(p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}\Big)$$
$$= e^{-r\tau(t)}\Big(p_t W + (1-p_t)(-\omega)\Big) - c\frac{1-e^{-rt}}{r}.$$

This equality can be rewritten as

$$\int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}}\Big(e^{-r(\tau(s)-t)}W - c\frac{1-e^{-r(s-t)}}{r}\Big)ds + \int_t^T (1-p_t)\lambda^0 \frac{e^{-\lambda^0 s}}{e^{-\lambda^0 t}}(-c)\frac{1-e^{-r(s-t)}}{r}ds$$
$$+ p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}}e^{-r(\tau(T)-t)}W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}e^{-r(\tau(T)-t)}(-\omega) - c\frac{1-e^{-r(T-t)}}{r}\Big(p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}\Big)$$
$$= e^{-r(\tau(t)-t)}\Big(p_t W + (1-p_t)(-\omega)\Big).$$

This means that at belief $p = p_t$,

$$v(p_t, t) = \int_t^T p_t\lambda^1 \frac{e^{-\lambda^1 s}}{e^{-\lambda^1 t}}\Big(e^{-r(\tau(s)-t)}W - c\frac{1-e^{-r(s-t)}}{r}\Big)ds + \int_t^T (1-p_t)\lambda^0 \frac{e^{-\lambda^0 s}}{e^{-\lambda^0 t}}(-c)\frac{1-e^{-r(s-t)}}{r}ds$$
$$+p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}}e^{-r(\tau(T)-t)}W + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}e^{-r(\tau(T)-t)}(-\omega) - c\frac{1-e^{-r(T-t)}}{r}\Big(p_t\frac{e^{-\lambda^1 T}}{e^{-\lambda^1 t}} + (1-p_t)\frac{e^{-\lambda^0 T}}{e^{-\lambda^0 t}}\Big).$$

In other words, $v(p_t, t)$ is the discounted expected payoff the agent gets by incurring effort cost and being truthful over $[t, T]$. This proves that at (p_t, t), it is optimal to make the learning effort and declare truthfully. When $p > p_t$, $v(p, t) = e^{-r(\tau(t)-t)}(pW + (1-p)(-\omega))$ and therefore it is optimal to declare the good signal. ∎

References

Aghion, P., and J. Tirole (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1), 1–29.

Alonso, R., and N. Matouschek (2008): "Optimal Delegation," *The Review of Economic Studies*, 75(1), 259–293.

Amador, M., and K. Bagwell (2013): "The Theory of Optimal Delegation with an Application to Tariff Caps," *Econometrica*, 81(4), 1541–1599.

Ambrus, A., and G. Egorov (2017): "Delegation and Nonmonetary Incentives," *Journal of Economic Theory*, 171, 101–135.

Armstrong, M., and J. Vickers (2010): "A Model of Delegated Project Choice," *Econometrica*, 78(1), 213–244.

Bergemann, D., and J. Valimaki (forthcoming): "Dynamic Mechanism Design: An Introduction," *Journal of Economic Literature*.

Bolton, P., and M. Dewatripont (2005): *Contract Theory*. MIT Press.

Chen, Y. (2018): "Dynamic Communication with Commitment," Discussion paper.

Coddington, E., and N. Levinson (1955): *Theory of Ordinary Differential Equations*. Tata McGraw-Hill Education.

Deimen, I., and D. Szalay (forthcoming): "Delegated Expertise, Authority, and Communication," *American Economic Review*.

Frankel, A. (2016): "Discounted Quotas," *Journal of Economic Theory*, 166, 396–444.

George, S., and M. Buyse (2015): "Data Fraud in Clinical Trials," *Clinical Investigation*, 5(2), 161.

Grenadier, S., A. Malenko, and N. Malenko (2016): "Timing Decisions in Organizations: Communication and Authority in a Dynamic Environment," *The American Economic Review*, 106(9), 2552–2581.

Guo, Y. (2016): "Dynamic Delegation of Experimentation," *The American Economic Review*, 106(8), 1969–2008.

Guo, Y., and J. Hörner (2018): "Dynamic Allocation without Money," Discussion paper.

Henry, E., and M. Ottaviani (forthcoming): "Research and the Approval Process," *American Economic Review*.

Holmstrom, B. (1984): "On the Theory of Delegation," in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom. Elsevier.

Keller, G., S. Rady, and M. Cripps (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73(1), 39–68.

Lewis, T. R., and M. Ottaviani (2008): "Search agency," *Unpublished paper, Northwestern University.[428]*.

Li, J., N. Matouschek, and M. Powell (2017): "Power Dynamics in Organizations," *American Economic Journal: Microeconomics*, 9(1), 217–41.

Lipnowski, E., and J. Ramos (2017): "Repeated Delegation," Discussion paper.

Liptser, R., and A. Shiryaev (2013): *Statistics of Random Processes: General theory*, vol. 5. Springer Science & Business Media.

Madsen, E. (2018): "Expert Advice and Optimal Project Termination," Discussion paper.

McClellan, A. (2017): "Experimentation and Approval Mechanisms," Discussion paper.

Melumad, N., and T. Shibano (1991): "Communication in Settings with No Transfers," *The RAND Journal of Economics*, pp. 173–198.

Orlov, D., A. Skrzypacz, and P. Zryumov (forthcoming): "Persuading the Principal To Wait," *Journal of Political Economy*.

Pavan, A., I. Segal, and J. Toikka (2014): "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica*, 82(2), 601–653.

Seife, C. (2015): "Research Misconduct Identified by the US Food and Drug Administration: Out of Sight, Out of Mind, Out of the Peer-Reviewed Literature," *JAMA Internal Medicine*, 175(4), 567–577.

Sugaya, T., and A. Wolitzky (2017): "Revelation principles in multistage games," Discussion paper.

Szalay, D. (2005): "The Economics of Clear Advice and Extreme Options," *The Review of Economic Studies*, 72(4), 1173–1198.