

Dynamic Delegation of Experimentation*

Yingni Guo[†]

April 15, 2014

Abstract

I study a dynamic relationship in which a principal delegates experimentation to an agent. Experimentation is modeled as a two-armed bandit whose risky arm yields successes following a Poisson process. Its intensity, unknown to the players, is either high or low. The agent has private information, his type being his prior belief that the intensity is high. The agent values successes more than the principal and therefore prefers to experiment longer. I show how to reduce the analysis to a finite-dimensional problem. In the optimal contract, the principal starts with a calibrated prior belief and updates it as if the agent had no private information. The agent is free to experiment or not if this belief remains above a cutoff. He is required to stop once it reaches the cutoff. The cutoff binds for a positive measure of high enough types. Surprisingly, this delegation rule is time-consistent. I prove that the cutoff rule remains optimal and time-consistent for more general stochastic processes governing payoffs.

Keywords: principal-agent, delegation, experimentation, two-armed bandit, incomplete information, no monetary transfers.

JEL Codes: D82, D83, D86.

*I am grateful to Johannes Hörner, Larry Samuelson and Dirk Bergemann for insightful conversations and constant support. I would like to thank Daniel Barron, Alessandro Bonatti, Florian Ederer, Eduardo Faingold, Mitsuru Igami, Yuichiro Kamada, Nicolas Klein, Daniel Kraehmer, Zvika Neeman, Andrew Postlewaite, Andrzej Skrzypacz, Philipp Strack, Eran Shmaya, Juuso Toikka, Juuso Välimäki, Jörgen Weibull, and various seminar and conference audiences for insightful discussion and comments.

[†]Department of Economics, Yale University, 37 Hillhouse Ave., New Haven, CT, 06520, USA, yingni.guo@yale.edu.

1 Introduction

Motivation. Innovation carries great uncertainty. Firms frequently start R&D projects with little knowledge of eventual success. As experimentation goes on but no success occurs, firms grow pessimistic and taper resource input or even discontinue the project altogether.

This paper studies the optimal mechanism by which a principal (she) delegates experimentation to an agent (he), as is the case of a firm delegating an R&D project to its employee. The literature on experimentation in a principal-agent setting focuses on transferable utilities.¹ Instead, I focus on delegation (Holmström, 1977, 1984 [15] [16]) for three reasons. First, from a practical point of view, it is obvious that an overwhelming number of economic activities (conventional and innovative) are organized by delegation: managers delegate tasks to subordinates by authority, rather than transfer-based trading contracts. Second, it is often cheaper to restrict the agent’s actions than to devise a possibly complex compensation scheme. This is consistent with the transaction-cost economics which discusses the relative efficiency of authority-based organization (“hierarchies”) and contract-based organization (“market”) (Coase, 1937 [10]; Williamson, 1975 [25]). Third, there are cases in which transfers are prohibited outright to prevent corruption, such as constituents delegating reforms to politicians.

Current literature on delegation focuses on static problems which preclude learning. In this paper, I consider the problem of dynamic delegation. As new information arrives over time, the flexibility granted to the agent might be adjusted accordingly.

The R&D project consumes the principal’s resources and the agent’s time. Both wish to discontinue it if they become pessimistic enough. However, the agent’s relative return from the project’s successes typically exceeds the principal’s (high cost of principal’s resources; principal’s moderate benefit from one project out of her many responsibilities; agent’s career advancement as an extra benefit); hence the agent prefers to keep the project alive for a longer time.

Promising projects warrant longer experimentation. Building on his expertise, the agent often has private knowledge on the prospect of the project at the outset. If the principal wishes to take advantage of his information, she has to give the agent some flexibility over resource allocation. But misaligned preferences curtail the flexibility that the principal is willing to grant. Therefore the principal faces a trade-off between using the agent’s information and containing his bias.

The purpose of this paper is to solve for the optimal delegation rule. It addresses the following questions: In the absence of transfers, what instruments does the principal have to extract the agent’s private information? Is there delay in information acquisition? How much of the resource allocation decision should be delegated to the agent? Will some projects be over-experimented and

¹See for instance Bergemann and Hege (1998, 2005) [7] [8], Halac, Kartik, and Liu (2013) [14], Hörner and Samuelson (2013) [17].

others under-experimented? Is the optimal delegation rule time-consistent?

Analysis. I examine a dynamic relationship in which a principal delegates experimentation to an agent. Experimentation is modeled as a continuous-time two-armed bandit problem. See for instance Presman (1990) [23], Keller, Rady, and Cripps (2005) [20], and Keller and Rady (2010) [18]. There is one unit of a perfectly divisible resource per unit of time and the agent continually splits the resource between a safe task and a risky one. In any given time interval, the safe task generates a known flow payoff proportional to the resource allocated to it.² The risky task's payoff depends on an unknown binary *state*. In the *benchmark setting*, if the state is good the risky task yields *successes* at random times. The arrival rate is proportional to the resource allocated to it. If the state is bad, the risky task yields no successes. I assume that the agent values the safe task's flow payoffs and the risky task's successes differently than the principal. Both prefer to allocate the resource to the risky task in the good state and to the safe task in the bad one. However, the agent values successes relatively more than the principal; hence, he prefers to experiment longer if faced with prolonged absence of success.³ At the outset, the agent has private information: his *type* is his prior belief that the state is good. After experimentation begins, the agent's actions and the arrivals of successes are publicly observed.

The principal delegates the decision on how the agent should allocate the resource over time. This decision is made at the outset. Since the agent has private information before experimentation, the principal offers a set of *policies* from which the agent chooses his preferred one. A policy specifies how the agent should allocate the resource in all future contingencies.

Note that the space of all policies is very large. Possible policies include: allocate all resource to the risky task until a fixed time and then switch to the safe task only if no success has realized; gradually reduce the resource input to the risky task if no success occurs and allocate all resource to it after the first success; allocate all resource to the risky task until the first success and then allocate a fixed fraction to the risky task; always allocate a fixed fraction of the unit resource to the risky task; etc.

A key observation is that any policy, in terms of payoffs, can be summarized by a pair of numbers, corresponding to the *total expected discounted resource* allocated to the risky task conditional on the state being good and the *total expected discounted resource* allocated to the risky task conditional on the state being bad. That is, as far as payoffs are concerned, there is a simple, finite-dimensional summary statistic for any given policy. The range of these summary statistics as we vary policies is what I call the *feasible set*—a subset of the plane. Determining the feasible

²The flow payoff generated by the safe task can be regarded as the opportunity cost saved or the payoff from conducting conventional tasks.

³This assumption will be relaxed later and the case in which the bias goes in the other direction will also be studied.

set is a nontrivial problem in general, but it involves no incentive constraints, and so reduces to a standard optimization problem which I solve. This reduces the delegation problem to a static one. Given that the problem is now static, I use Lagrangian optimization methods (similar to those used by Amador, Werning, and Angeletos (2006) [4]) to determine the optimal delegation rule.

Under a mild regularity condition, the optimal delegation rule takes a very simple form. It is a *cutoff rule* with a properly calibrated prior belief that the state is good. This belief is then updated as if the agent had no private information. In other words, this belief drifts down when no success is observed and jumps to one upon the first success. It is updated in the way the principal would if she were carrying out the experiment herself (starting at the calibrated prior belief). The agent freely decides whether to experiment or not as long as the updated belief remains above the cutoff. However, if this belief ever decreases to the cutoff, the agent is required to stop experimenting. This rule turns out not to bind for types with low enough priors, who voluntarily stop experimenting conditional on no success, but does constrain those with high enough priors, who are required to stop when the cutoff is reached.

Given this updating rule, the belief jumps to one upon the first success. Hence, in the benchmark setting the cutoff rule can be implemented by imposing a deadline for experimentation, under which the agent allocates all resource to the risky task after the first success, but is not allowed to experiment past the deadline. Those types with low enough priors stop experimenting before the deadline conditional on no success while a positive measure of types with high enough priors stop at the deadline. In equilibrium, there is no delay in information acquisition as the risky task is operated exclusively until either the first success reveals that the state is good or the agent stops.

Among the positive measure of high enough types who are forced to stop when the cutoff (or the deadline) is reached, the highest subset under-experiment even from the principal's point of view. Every other type over-experiments. This implies that in practice the most promising projects are always terminated too early while less promising ones are stopped too late due to the agency problem.

An important property of the cutoff rule is time consistency. After any history the principal would not adjust the cutoff rule even if she were given a chance to do so. In particular, after the agent experiments for some time yet no success has realized, the principal still finds it optimal to keep the cutoff (or the deadline) at the same level as it was set at the beginning. This property indicates that, surprisingly, implementing the cutoff rule requires minimal commitment on the principal's side.

I then show that both the optimality of the cutoff rule and its time-consistency generalize to situations in which the risky task generates successes in the bad state as well. When successes are inconclusive, the belief is updated differently than in the benchmark setting. It jumps up upon successes and then drifts down. Consequently, the cutoff rule cannot be implemented by imposing

a deadline. Instead, it can be interpreted as a *sliding deadline*. The principal initially extends some time to the agent to operate the risky task. Then, whenever a success realizes, more time is extended. The agent is free to switch to the safe task before he uses up the time granted by the principal. After a long enough period of time elapses without success, the agent is required to switch to the safe task.

I further extend the analysis to the case in which the agent gains less from the experiment than the principal and therefore tends to under-experiment. This happens when an innovative task yields positive externalities, or when it is important to the firm but does not widen the agent's influence. When the agent's bias is small enough, the optimum can be implemented by imposing a lockup period which is extended upon successes. Instead of placing a cap on the length of experimentation in the previous case, the principal enacts a floor. The agent has no flexibility but to experiment before the lockup period ends, yet has full flexibility afterwards. Time-consistency is no longer valid, though, as whenever the agent stops experimenting voluntarily, he reveals that the principal's optimal experimentation length has yet to be reached. The principal is tempted to order the agent to experiment further. Therefore to implement the sliding lockup period, commitment from the principal is required.

My results have two important implications for the practical design of delegation rules (I assume a larger agent's return in this illustration). First, a (sliding) deadline should be in place as a safeguard against abuse of the principal's resources. The continuation of the project is permitted only upon demonstrated successes. Second, the agent should have the flexibility over resource allocation before the (sliding) deadline is reached. In particular, the agent should be free to terminate the project whenever he finds appropriate. Besides in-house innovation, these results apply to various resource allocation problems with experimentation, such as companies budgeting marketing resources for product introduction and funding agencies awarding grants to scholarly research.⁴

Related literature. My paper contributes to the literature on delegation. This literature addresses the incentive problems in organizations which arise due to hidden information and misaligned preferences. Holmström (1977, 1984) [15] [16] provides conditions for the existence of an optimal solution to the delegation problem. He also characterizes optimal delegation sets in a series of examples, under the restriction, for the most part, that only interval delegation sets are allowed. Alonso and Matouschek (2008) [2] and Amador and Bagwell (2012) [3] characterize the optimal delegation set in general environments under some conditions and provide conditions under which simple interval delegation is optimal.

None of these papers consider dynamic delegation. What distinguishes my model from static

⁴Suppose that experimentation has a flow cost. The agent is cash-constrained and his action is contractible. Delegating experimentation equals funding his research project.

delegation problems is that additional information arises over time. The principal ought to use it both to reduce the agent's informational rents and to adjust his behavior. My paper complements the current literature and facilitates the understanding of how to optimally delegate experimentation.

Second, my paper is related to the literature on experimentation in a principal-agent setting.⁵ Since most papers address different issues than I do, here I only mention the most related ones. Gomes, Gottlieb and Maestri (2013) [13] study a multiple-period model in which the agent has private information about both the project quality and his cost of effort. The agent's actions are observable. Unlike my setting, the agent has no benefit from the project and outcome-contingent transfers are allowed for. They identify necessary and sufficient conditions under which the principal only pays rents for the agent's information about his cost, but not for the agent's information about the project quality. Garfagnini (2011) [12] studies a dynamic delegation model without hidden information at the beginning. The principal cannot commit to future actions and transfers are infeasible. Agency conflicts arise because the agent prefers to work on the project regardless of the state. He delays information acquisition to prevent the principal from growing pessimistic. In my model, there is pre-contractual hidden information; transfers are infeasible and the principal is able to commit to long-term contract terms; the agent has direct benefit from experimentation and shares the same preferences as the principal conditional on the state. Agency conflicts arise as the agent is inclined to exaggerate the prospects for success and prolong the experimentation.

The paper is organized as follows. The model is presented in Section 2. Section 3 considers a single player's decision problem. In Section 4, I illustrate how to reduce the delegation problem to a static one. The main results are presented in Section 5. I extend the analysis to more general stochastic processes in Section 6 and discuss other applications of the model. Section 7 concludes.

2 The Model

Players, tasks and states. Time $t \in [0, \infty)$ is continuous. There are two risk-neutral players $i \in \{\alpha, \rho\}$, an agent (he) and a principal (she), and two tasks, a safe task S and a risky one R . The principal is endowed with one unit of perfectly divisible resource per unit of time. She delegates resource allocation to the agent, who continually splits the resource between the two tasks. The safe task yields a known deterministic flow payoff that is proportional to the fraction of the resource

⁵Bergemann and Hege (1998, 2005) [7] [8] study the financing of a new venture in which the principal funds the experiment and the agent makes contract offers. Dynamic agency problem arises as the agent can invest or divert the funds. Hörner and Samuelson (2013) [17] consider a similar model in which the agent's effort requires funding and is unobservable. The principal makes short-term contract offers specifying profit-sharing arrangement. Halac, Kartik, and Liu (2013) [14] study long-term contract for experimentation with adverse selection about the agent's ability and moral hazard about his effort choice.

allocated to it. The risky task's payoff depends on an unknown binary *state*, $\omega \in \{0, 1\}$.

In particular, if the fraction $\pi_t \in [0, 1]$ of the resource is allocated to R over an interval $[t, t+dt)$, and consequently $1 - \pi_t$ to S , player i receives $(1 - \pi_t)s_i dt$ from S , where $s_i > 0$ for both players. The risky task generates a *success* at some point in the interval with probability $\pi_t \lambda^1 dt$ if $\omega = 1$ and $\pi_t \lambda^0 dt$ if $\omega = 0$. Each success is worth h_i to player i . Therefore, the overall expected payoff increment to player i conditional on ω is $[(1 - \pi_t)s_i + \pi_t \lambda^\omega h_i] dt$. All this data is common knowledge.⁶

In the benchmark setting, I assume that $\lambda^1 > \lambda^0 = 0$. Hence, R yields no success in state 0. In Subsection 6.1.1, I extend the analysis to the setting in which $\lambda^1 > \lambda^0 > 0$.

Conflicts of interests. I allow different payoffs to players, *i.e.*, I do not require that $s_\alpha = s_\rho$ or $h_\alpha = h_\rho$. The restriction imposed on payoff parameters is the following:

Assumption 1. Parameters are such that $\lambda^1 h_i > s_i > \lambda^0 h_i$ for $i \in \{\alpha, \rho\}$, and

$$\frac{\lambda^1 h_\alpha - s_\alpha}{s_\alpha - \lambda^0 h_\alpha} > \frac{\lambda^1 h_\rho - s_\rho}{s_\rho - \lambda^0 h_\rho}.$$

Assumption 1 has two implications. First, there is agreement on how to allocate the unit resource if the state is known. Both players prefer to allocate the resource to R in state 1 and the resource to S in state 0. Second, the agent values successes over flow payoffs relatively more than the principal does. Let

$$\eta_i = \frac{\lambda^1 h_i - s_i}{s_i - \lambda^0 h_i}$$

denote player i 's net gain from R 's successes over S 's flow payoffs. The ratio η_α/η_ρ , being strictly greater than one, measures how misaligned players' interests are and is referred to as the agent's *bias*. (The case in which the bias goes in the other direction is discussed in Subsection 6.2.)

Private information. Players do not observe the state. At time 0, the agent has private information about the probability that the state is 1. For ease of exposition, I express the agent's prior belief that the state is 1 in terms of the implied odds ratio of state 1 to state 0, denoted θ and referred to as the agent's type. The agent's type is drawn from a compact interval $\Theta \equiv [\underline{\theta}, \bar{\theta}] \subset \mathbf{R}_+$ according to some continuous density function f . Let F denote the cumulative distribution function.

By the definition of the odds ratio, the agent of type θ assigns probability $p(\theta) = \theta/(1 + \theta)$ to the event that the state is 1 at time 0. The principal knows only the distribution of types. Hence,

⁶It is not necessary that S generates deterministic flow payoffs. What matters to players is that the expected payoff rates of S are known and equal s_i , and that S 's flow payoffs are uncorrelated with the state.

her prior belief that the state is 1 is given by

$$\mathbf{E}[p(\theta)] = \int_{\Theta} \frac{\theta}{1 + \theta} dF(\theta).$$

Actions and successes are publicly observable. The only information asymmetry comes from the agent's private information about the state at time 0. Hence, a resource allocation policy, which I introduce next, conditions on both the agent's past actions and arrivals of successes.

Policies and posterior beliefs. A (pure) resource allocation *policy* is a non-anticipative stochastic process $\pi = \{\pi_t\}_{t \geq 0}$. Here, $\pi_t \in [0, 1]$ is interpreted as the fraction of the unit resource allocated to R at time t , which may depend only on the history of events up to t . A policy π can be described as follows. At time 0, a choice is made of a deterministic function $\pi(t | 0)$, measurable with respect to t , $0 \leq t < \infty$, which takes values in $[0, 1]$ and corresponds to the fraction of the resource allocated to R up to the moment of the first success. If at the random time τ_1 a success occurs, then depending on the value of τ_1 , a new function $\pi(t | \tau_1, 1)$ is chosen, etc. The space of all policies, including randomized ones, is denoted Π . (See Footnote 8.) The agent is said to *experiment* when a positive fraction of the unit resource is allocated to R when the state is still unknown.

Let N_t denote the number of successes observed up to time t . Both players discount payoffs at rate $r > 0$. Player i 's *payoff* given an arbitrary policy $\pi \in \Pi$ and an arbitrary prior belief $p \in [0, 1]$ consists of the expected discounted payoffs from R 's successes and the expected discounted flow payoffs from S

$$U_i(\pi, p) \equiv \mathbf{E} \left[\int_0^{\infty} r e^{-rt} [h_i dN_t + (1 - \pi_t) s_i dt] \mid \pi, p \right].$$

Here, the expectation is taken over the state variable ω and the stochastic processes π and N_t . By the Law of Iterated Expectations, I can rewrite player i 's payoff as the discounted sum of the expected payoff increments

$$U_i(\pi, p) = \mathbf{E} \left[\int_0^{\infty} r e^{-rt} [(1 - \pi_t) s_i + \pi_t \lambda^\omega h_i] dt \mid \pi, p \right].$$

Given prior p , policy π and trajectory N_s on the time interval $0 \leq s \leq t$, I consider the posterior probability p_t that the state is 1. The function p_t may be assumed to be right-continuous with left-hand limits. Because R yields no success in state 0, before the first success of the process N_t , the process p_t satisfies a differential equation

$$\dot{p}_t = -\pi_t \lambda^1 p_t (1 - p_t). \tag{1}$$

At the first success, p_t jumps to one.⁷

Delegation. I consider the situation in which transfers are not allowed and the principal is able to commit to dynamic policies. At time 0, the principal chooses a set of policies from which the agent chooses his preferred one. Since there is hidden information at time 0, by the Revelation Principle, the principal's problem is reduced to solving for a map $\pi : \Theta \rightarrow \Pi$ to maximize her expected payoff subject to the agent's incentive compatibility constraint (IC constraint, hereafter). Formally, I solve

$$\begin{aligned} & \sup \int_{\Theta} U_{\rho}(\pi(\theta), p(\theta)) dF(\theta), \\ & \text{subject to } U_{\alpha}(\pi(\theta), p(\theta)) \geq U_{\alpha}(\pi(\theta'), p(\theta)) \quad \forall \theta, \theta' \in \Theta, \end{aligned}$$

over measurable $\pi : \Theta \rightarrow \Pi$.⁸

3 The Single-Player Benchmark

In this section, I present player i 's preferred policy as a single player. This is a standard problem. The policy preferred by player i is Markov with respect to the posterior belief p_t . It is characterized by a cutoff belief p_i^* such that $\pi_t = 1$ if $p_t \geq p_i^*$ and $\pi_t = 0$ otherwise. By standard results (see Keller, Rady, and Cripps (2005, Proposition 3.1) [20], for instance), the cutoff belief is

$$p_i^* = \frac{s_i}{\lambda^1 h_i + \frac{(\lambda^1 h_i - s_i) \lambda^1}{r}} = \frac{r}{r + (\lambda^1 + r) \eta_i}. \quad (2)$$

Note that this is lower than player i 's myopic cutoff belief $s_i/(\lambda^1 h_i)$, *i.e.*, the probability below which R yields a lower expected flow payoff than S does. The cutoff belief p_i^* decreases in η_i . Therefore, the agent's cutoff belief p_{α}^* is lower than the principal's p_{ρ}^* , as he values R 's successes over S 's flow payoffs more than the principal does.

⁷In general, subscripts indicate either time or player. Superscripts refer to state. Parentheses contain type or policy.

⁸Here, I define randomized policies and stochastic mechanisms following Aumann (1964) [6]. Let $\mathcal{B}_{[0,1]}$ (resp. \mathcal{B}^k) denote the σ -algebra of Borel sets of $[0, 1]$ (resp. \mathbf{R}_+^k) and λ the Lebesgue measure on $[0, 1]$, where k is a positive integer. I denote the set of measurable functions from $(\mathbf{R}_+^k, \mathcal{B}^k)$ to $([0, 1], \mathcal{B}_{[0,1]})$ by F^k and endow this set with the σ -algebra generated by sets of the form $\{f : f(s) \in A\}$ with $s \in \mathbf{R}_+^k$ and $A \in \mathcal{B}_{[0,1]}$. The σ -algebra is denoted χ^k . Let Π^* denote the space of pure policies. I impose on Π^* the product σ -algebra generated by (F^k, χ^k) , $\forall k \in N_+$. Following Aumann (1964) [6], I define randomized policies as measurable functions $\hat{\pi} : [0, 1] \rightarrow \Pi^*$. According to $\hat{\pi}$, a value $\epsilon \in [0, 1]$ is drawn uniformly from $[0, 1]$ and then the pure policy $\hat{\pi}(\epsilon)$ is implemented. Analogously, I define stochastic mechanisms as measurable functions $\hat{\pi} : [0, 1] \times \Theta \rightarrow \Pi^*$. A value $\epsilon \in [0, 1]$ is drawn uniformly from $[0, 1]$, along with the agent's report θ , determines which element of Π is chosen. For ease of exposition, my descriptions assume pure policies and deterministic mechanisms. My results do not.

Given the law of motion of beliefs (1) and the cutoff belief (2), player i 's preferred policy given prior $p(\theta)$ can be identified with a *stopping time* $\tau_i(\theta)$: if the first success occurs before the stopping time, use R forever after the first success; otherwise, use R until the stopping time and then switch to S . Player i 's preferred stopping time for a given θ is stated as follows:

Claim 1. *Player i 's stopping time given odds ratio $\theta \in \Theta$ is*

$$\tau_i(\theta) = \begin{cases} \frac{1}{\lambda^1} \log \frac{(r+\lambda^1)\theta\eta_i}{r} & \text{if } \frac{(r+\lambda^1)\theta\eta_i}{r} \geq 1, \\ 0 & \text{if } \frac{(r+\lambda^1)\theta\eta_i}{r} < 1. \end{cases}$$

Figure 1 illustrates the two players' cutoff beliefs and their preferred stopping times associated with two possible odds ratios θ', θ'' (with $\theta' < \theta''$).⁹ The prior beliefs are thus $p(\theta'), p(\theta'')$ (with $p(\theta') < p(\theta'')$). The x -axis variable is time t and the y -axis variable is the posterior belief. On the y -axis is labeled the two players' cutoff beliefs p_ρ^* and p_α^* . The solid and dashed lines depict how posterior beliefs evolve when R is used exclusively and no success realizes.

The figure on the left-hand side shows that for a given odds ratio the agent prefers to experiment longer than the principal does because his cutoff is lower than the principal's. The figure on the right-hand side shows that for a given player i , the stopping time increases in the odds ratio, *i.e.*, $\tau_i(\theta') < \tau_i(\theta'')$. Therefore, both players prefer to experiment longer given a higher odds ratio. Figure 1 makes clear what agency problem the principal faces. The principal's stopping time $\tau_\rho(\theta)$ is an increasing function of θ . The agent prefers to stop later than the principal for a given θ and thus has incentives to misreport his type. More specifically, lower types (those types with a lower θ) have incentives to mimic high types to prolong the experimentation.

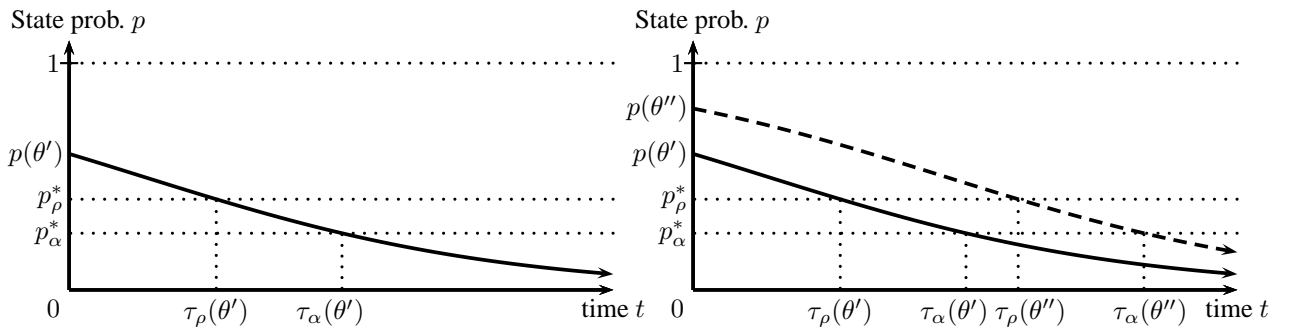


Figure 1: Thresholds and stopping times

Given that a single player's preferred policy is always characterized by a stopping time, one might expect that the solution to the delegation problem is a set of stopping times. This is the

⁹Parameters in Figure 1 are $\eta_\alpha = 3/2, \eta_\rho = 3/4, r/\lambda^1 = 1, \theta = 3/2, \theta' = 4$.

case if there is no private information or no bias. For example, if the distribution F is degenerate, information is symmetric. The optimal delegation set is the principal's preferred stopping time given her prior. If η_α/η_ρ equals one, the two players' preferences are perfectly aligned. The principal, knowing that for any prior the agent's preferred stopping time coincides with hers, offers the set of her preferred stopping times $\{\tau_\rho(\theta) : \theta \in \Theta\}$ for the agent to choose from.

However, if the agent has private information and is also biased, it is unclear how the principal should restrict his actions. Particularly, it is unclear whether the principal would still offer a set of stopping times. For this reason, I am led to consider the space of all policies.

4 A Finite-Dimensional Characterization of the Policy Space

The space of all policies is large. In the first half of this section, I associate to each policy—a (possibly complicated) stochastic process—a pair of numbers, called *total expected discounted resource* pair, and show that this pair is a sufficient statistic for this policy in terms of both players' payoffs. Then, I solve for the set of feasible *total expected discounted resource* pairs, which is a subset of \mathbf{R}^2 and can be treated as the space of all policies.

This transformation allows me to reduce the dynamic delegation problem to a static one. In the second half of this section, I characterize players' preferences over the feasible pairs and reformulate the delegation problem.

4.1 A Policy as a Pair of Numbers

For a fixed policy π , I define $\mathbf{w}^1(\pi)$ and $\mathbf{w}^0(\pi)$ as follows:

$$\mathbf{w}^1(\pi) \equiv \mathbf{E} \left[\int_0^\infty r e^{-rt} \pi_t dt \mid \pi, 1 \right] \text{ and } \mathbf{w}^0(\pi) \equiv \mathbf{E} \left[\int_0^\infty r e^{-rt} \pi_t dt \mid \pi, 0 \right]. \quad (3)$$

The first term $\mathbf{w}^1(\pi)$ is the expected discounted sum of the resource allocated to R if policy π is implemented conditional on the state being 1. I refer to $\mathbf{w}^1(\pi)$ as the *total expected discounted resource* (*expected resource*, hereafter) allocated to R under π in state 1.¹⁰ Similarly, the second term $\mathbf{w}^0(\pi)$ is the *expected resource* allocated to R under π in state 0. Both $\mathbf{w}^1(\pi)$ and $\mathbf{w}^0(\pi)$ are in $[0, 1]$ because π takes values in $[0, 1]$. Therefore, $(\mathbf{w}^1, \mathbf{w}^0)$ defines a mapping from the policy space Π to $[0, 1]^2$.

To calculate the payoff of implementing a policy for a given prior p , I first calculate the payoff of implementing this policy if the state is 1 (or equivalently, $p = 1$) and the payoff if the state is 0

¹⁰For a fixed policy π , the *expected resource* spent on R in state 1 is proportional to the expected discounted number of successes, i.e., $\mathbf{w}^1(\pi) = \mathbf{E} \left[\int_0^\infty r e^{-rt} dN_t \mid \pi, 1 \right] / \lambda^1$.

(or equivalently, $p = 0$). Multiplying these payoffs by the initial distribution of the state gives the payoff of this policy. Then I show that, conditional on the state, the payoff of implementing one policy is linear in the *expected resource* allocated to R .¹¹ As a result, what is relevant for evaluating a policy π is the *expected resource* pair $(\mathbf{w}^1(\pi), \mathbf{w}^0(\pi))$. I summarize this in the following lemma.

Lemma 1 (A policy as a pair of numbers).

For a given policy $\pi \in \Pi$ and a given prior $p \in [0, 1]$, player i 's payoff can be written as

$$U_i(\pi, p) = p (\lambda^1 h_i - s_i) \mathbf{w}^1(\pi) + (1 - p) (\lambda^0 h_i - s_i) \mathbf{w}^0(\pi) + s_i. \quad (4)$$

Proof. Player i 's payoff given policy $\pi \in \Pi$ and prior $p \in [0, 1]$ is

$$\begin{aligned} U_i(\pi, p) &= \mathbf{E} \left[\int_0^\infty r e^{-rt} [(1 - \pi_t) s_i + \pi_t \lambda^\omega h_i] dt \mid \pi, p \right] \\ &= p \mathbf{E} \left[\int_0^\infty r e^{-rt} [s_i + \pi_t (\lambda^1 h_i - s_i)] dt \mid \pi, 1 \right] + (1 - p) \mathbf{E} \left[\int_0^\infty r e^{-rt} [s_i + \pi_t (\lambda^0 h_i - s_i)] dt \mid \pi, 0 \right] \\ &= p (\lambda^1 h_i - s_i) \mathbf{E} \left[\int_0^\infty r e^{-rt} \pi_t dt \mid \pi, 1 \right] + (1 - p) (\lambda^0 h_i - s_i) \mathbf{E} \left[\int_0^\infty r e^{-rt} \pi_t dt \mid \pi, 0 \right] + s_i \\ &= p (\lambda^1 h_i - s_i) \mathbf{w}^1(\pi) + (1 - p) (\lambda^0 h_i - s_i) \mathbf{w}^0(\pi) + s_i. \end{aligned}$$

■

According to Lemma 1, $(\mathbf{w}^1(\pi), \mathbf{w}^0(\pi))$ is a sufficient statistic in terms of payoffs for the policy π . Moreover, $U_i(\pi, p)$ is a linear function of $(\mathbf{w}^1(\pi), \mathbf{w}^0(\pi))$ as specified in (4). Lemma 1 claims that instead of working with a generic policy π , it is without loss of generality to focus on $(\mathbf{w}^1(\pi), \mathbf{w}^0(\pi))$.

4.2 Feasible Pairs

Let Γ denote the image of the mapping $(\mathbf{w}^1, \mathbf{w}^0) : \Pi \rightarrow [0, 1]^2$. I call Γ the *feasible set* since it contains all possible (w^1, w^0) pairs that can be achieved by some policy π . To illustrate the concept of *expected resource*, I calculate the image of $(\mathbf{w}^1, \mathbf{w}^0)$ for two classes of policies, which turn out to be important for characterizing Γ .

1. *Stopping-time policies*: allocate all resource to R until the stopping time; if at least one success occurs by then, allocate all resource to R forever; otherwise, switch to S forever. Let π_τ^{st} denote such a policy with τ being the stopping time. If the state is 0, no success will occur. All resource is allocated to R from time 0 to τ and then S is used exclusively until infinity. If the state is

¹¹Recall that, conditional on the state and over any interval, S generates a flow payoff proportional to the resource allocated to it and R yields a success with probability proportional to the resource allocated to it.

1, switch to S will be triggered if and only if no success occurs before τ , which happens with probability $e^{-\lambda^1 \tau}$. Therefore, I have

$$(\mathbf{w}^1(\pi_\tau^{\text{st}}), \mathbf{w}^0(\pi_\tau^{\text{st}})) = (1 - e^{-(r+\lambda^1)\tau}, 1 - e^{-r\tau}).$$

As the stopping time τ ranges from 0 to ∞ , $\mathbf{w}^1(\pi_\tau^{\text{st}})$ and $\mathbf{w}^0(\pi_\tau^{\text{st}})$ increase from 0 to 1. The image of all stopping-time policies under $(\mathbf{w}^1, \mathbf{w}^0)$ is denoted Γ^{st}

$$\Gamma^{\text{st}} \equiv \left\{ (w^1, w^0) \mid w^0 = 1 - (1 - w^1)^{\frac{r}{r+\lambda^1}}, w^1 \in [0, 1] \right\}.$$

2. *Slack-after-success policies*: allocate all resource to R until the first success occurs; then allocate a fixed fraction to R . Let π_ϕ^{sl} denote such a policy with $\phi \in [0, 1]$ being the fixed fraction allocated to R after the first success. If the state is 0, all resource is directed to R because no success will occur. Hence, $\mathbf{w}^0(\pi_\phi^{\text{sl}})$ always equals one. If the state is 1, the arrival time of the first success is exponentially distributed with rate parameter λ^1 . After the first success, the fraction ϕ is allocated to R until infinity. Therefore, I have

$$(\mathbf{w}^1(\pi_\phi^{\text{sl}}), \mathbf{w}^0(\pi_\phi^{\text{sl}})) = \left(\frac{r + \lambda^1 \phi}{r + \lambda^1}, 1 \right).$$

As the fixed fraction ϕ varies from 0 to 1, $\mathbf{w}^1(\pi_\phi^{\text{sl}})$ increases from $r/(r + \lambda^1)$ to 1. The image of all slack-after-success policies under $(\mathbf{w}^1, \mathbf{w}^0)$ is denoted Γ^{sl}

$$\Gamma^{\text{sl}} \equiv \left\{ (w^1, w^0) \mid w^0 = 1, w^1 \in \left[\frac{r}{r + \lambda^1}, 1 \right] \right\}.$$

Figure 2 depicts the image of all stopping-time policies and slack-after-success policies when r and λ^1 both equal $1/5$. The shaded area is $\text{co}(\Gamma^{\text{st}} \cup \Gamma^{\text{sl}})$. Here, $\text{co}(X)$ denotes the convex hull of X . The (w^1, w^0) pairs on the southeast boundary correspond to stopping-time policies. Those pairs on the north boundary correspond to slack-after-success policies.

The following lemma characterizes the feasible set, which is $\text{co}(\Gamma^{\text{st}} \cup \Gamma^{\text{sl}})$.

Lemma 2 (Feasible set).

The image of the mapping $(\mathbf{w}^1, \mathbf{w}^0) : \Pi \rightarrow [0, 1]^2$ is $\text{co}(\Gamma^{\text{st}} \cup \Gamma^{\text{sl}})$.

Proof. Given $\gamma = (\gamma^1, \gamma^0) \in \mathbf{R}^2$, define the supremum score in direction γ and the associated half

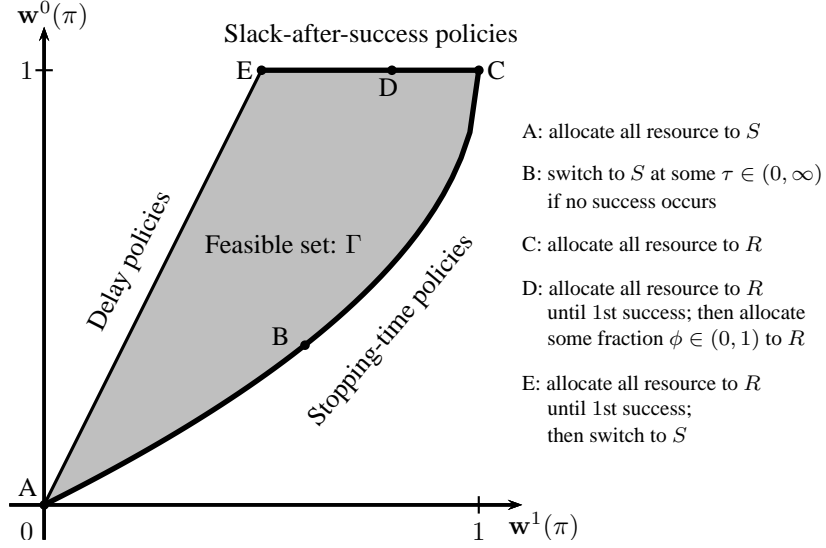


Figure 2: Feasible set and example policies ($r/\lambda^1 = 1$)

space as

$$K(\gamma) \equiv \sup_{\pi \in \Pi} [\gamma^1 \mathbf{w}^1(\pi) + \gamma^0 \mathbf{w}^0(\pi)],$$

$$\mathcal{H}(\gamma) \equiv \{v \in \mathbf{R}^2 : \gamma \cdot v \leq K(\gamma)\}.$$

Define the intersection of all half spaces as $\mathcal{H} \equiv \bigcap_{\gamma \in \mathbf{R}^2} \mathcal{H}(\gamma)$. Since $\Gamma \subset \mathcal{H}(\gamma)$ for any γ , it follows that $\Gamma \subset \mathcal{H}$. On the other hand, the feasible set Γ is convex given that the policy space Π and hence Γ are convexified. (Recall that Π includes all randomized policies (see Footnote 8).) It follows that $\Gamma = \mathcal{H}$.

If $\gamma^1 \geq 0, \gamma^0 \geq 0$, $K(\gamma)$ equals $\gamma^1 + \gamma^0$, achieved by the policy which directs all resources to R . If $\gamma^1 \leq 0, \gamma^0 \leq 0$, $K(\gamma)$ equals 0, achieved by the policy which directs all resources to S . If $\gamma^1 > 0, \gamma^0 < 0$, finding $K(\gamma)$ is equivalent to a (Bayesian) decision problem of player i who chooses π to maximize $U_i(\pi, p)$. His prior p is given by $p/(1-p) = -\gamma^1/(\gamma^0 \eta_i)$. Therefore, $K(\gamma)$ is achieved by a cutoff Markov policy under which R is used exclusively if the posterior belief is above p_i^* and S is used otherwise. This Markov policy is effectively a stopping-time policy, the stopping time increasing from 0 to ∞ as the ratio $-\gamma^1/\gamma^0$ increases from 0 to ∞ . Hence, the southeast boundary of \mathcal{H} is Γ^{st} (more detailed proof is given in Appendix 8.1).

If $\gamma^1 < 0, \gamma^0 > 0$, finding $K(\gamma)$ is equivalent to a decision problem of player i who chooses π to minimize $U_i(\pi, p)$. His prior p is given by $p/(1-p) = -\gamma^1/(\gamma^0 \eta_i)$. According to Keller and Rady (2013) [19], $K(\gamma)$ is achieved by a cutoff Markov policy such that R is used exclusively if the posterior belief is below the cutoff and S is used if above. Depending on his prior (or the ratio

$-\gamma^1/\gamma^0$), player i either uses R until the first success and then switches to S , or uses S forever. The former policy corresponds to the pair $(w^1, w^0) = (r/(r + \lambda^1), 1)$ and the latter to $(w^1, w^0) = (0, 0)$. Hence, the set \mathcal{H} is bounded from above by

$$\Gamma^{sl} \cup \left\{ (w^1, w^0) \mid \epsilon \left(\frac{r}{r + \lambda^1}, 1 \right) + (1 - \epsilon)(0, 0), \epsilon \in [0, 1] \right\}.$$

Therefore, $\Gamma = \mathcal{H} = co(\Gamma^{st} \cup \Gamma^{sl})$.

Note that the image of all pure policies under $(\mathbf{w}^1, \mathbf{w}^0)$ is also $co(\Gamma^{st} \cup \Gamma^{sl})$. Pick any pair (w^1, w^0) in the interior of $co(\Gamma^{st} \cup \Gamma^{sl})$. The line connecting the origin and (w^1, w^0) intersects $\Gamma^{st} \cup \Gamma^{sl}$ at some point $(\tilde{w}^1, \tilde{w}^0)$ which is the image of a stopping-time or slack-after-success policy $\tilde{\pi}$ under $(\mathbf{w}^1, \mathbf{w}^0)$. There exists a $t \in (0, \infty)$ such that (w^1, w^0) corresponds to a policy which directs all resource to S until time t and then the policy $\tilde{\pi}$ is implemented from time t on. ■

According to Lemma 2, the image of any policy π under $(\mathbf{w}^1, \mathbf{w}^0)$ is in $co(\Gamma^{st} \cup \Gamma^{sl})$. Also, for any $(w^1, w^0) \in co(\Gamma^{st} \cup \Gamma^{sl})$, I can identify a policy π such that $(w^1, w^0) = (\mathbf{w}^1(\pi), \mathbf{w}^0(\pi))$. From now on, when I refer to a pair $(w^1, w^0) \in \Gamma$, I have in mind a policy such that w^1 is the *expected resource* allocated to R under this policy in state 1 and w^0 is that in state 0. A (w^1, w^0) pair which lies in Γ is called a *bundle*.

The feasible set Γ is bounded from above by the union of Γ^{sl} and $\{(w^1, w^0) \mid \epsilon(r/(r + \lambda^1), 1) + (1 - \epsilon)(0, 0), \epsilon \in [0, 1]\}$. The latter set can be achieved by delaying the policy corresponding to point E (see Figure 2) for some fixed amount of time. I call this class of policies *delay policies*. From now on, I also refer to the union of Γ^{sl} and $\{(w^1, w^0) \mid \epsilon(r/(r + \lambda^1), 1) + (1 - \epsilon)(0, 0), \epsilon \in [0, 1]\}$ as the *northwest* boundary of Γ . The fact that the northwest boundary is piecewise linear is peculiar to the benchmark setting due to its degenerate feature that R yields no success in state 0. In Subsection 6.1.1, I will characterize the feasible sets of more general stochastic processes.

The shape of the feasible set only depends on the ratio r/λ^1 . Figure 3 shows that as r/λ^1 decreases, the feasible set expands in both the southeast and the northwest directions. Intuitively, if future payoffs are discounted to a less extent, a player has more time to learn about the state. As a result, he is more capable of directing resources to R in one state while avoiding wasting resources on R in the other state.

For future reference, I also write the feasibility set Γ as $\Gamma = \{(w^1, w^0) \mid \beta^{se}(w^1) \leq w^0 \leq \beta^{nw}(w^1), w^1 \in [0, 1]\}$ where β^{se}, β^{nw} are functions from $[0, 1]$ to $[0, 1]$, characterizing the southeast

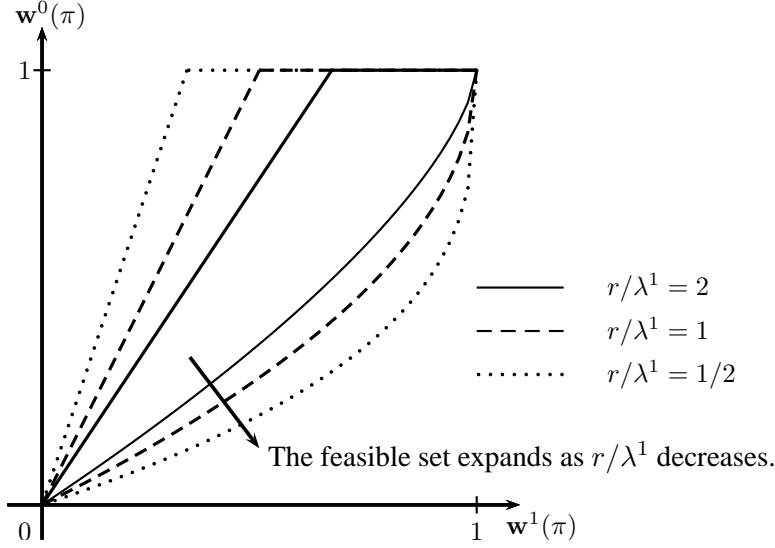


Figure 3: Feasible sets as r/λ^1 varies

and northwest boundaries of the feasible set,

$$\beta^{\text{se}}(w^1) \equiv 1 - (1 - w^1)^{\frac{r}{r+\lambda^1}},$$

$$\beta^{\text{nw}}(w^1) \equiv \begin{cases} (r + \lambda^1) w^1 & \text{if } w^1 \in [0, \frac{1}{r+\lambda^1}], \\ 1 & \text{if } w^1 \in (\frac{1}{r+\lambda^1}, \frac{1}{r}]. \end{cases}$$

4.3 Preferences over Feasible Pairs

If a player knew the state, he would allocate all resources to R in state 1 and all resources to S in state 0. However, this allocation cannot be achieved if the state is unknown. A policy, being state-independent, necessarily entails the cost of learning. If a player wants to direct more resources to R in state 1, he has to allocate more resources to R before the arrival of the first success. Inevitably, more resources will be wasted on R if the state is actually 0.

A player's attitude toward this trade-off between spending more resources on R in state 1 and wasting less resources on R in state 0 depends on how likely the state is 1 and how much he gains from R 's successes over S 's flow payoffs. According to Lemma 1, player i 's payoff given policy π and odds ratio θ is

$$U_i(\pi, p(\theta)) = \left(\frac{\theta}{1+\theta} \eta_i \mathbf{w}^1(\pi) - \frac{1}{1+\theta} \mathbf{w}^0(\pi) \right) (s_i - \lambda^0 h_i) + s_i.$$

Recall that $p(\theta) = \theta/(1 + \theta)$ is the prior that the state is 1 given θ . Player i 's preferences over (w^1, w^0) are characterized by upward-sloping indifference curves with the slope being $\theta \eta_i$. For a

fixed player, the indifference curves are steeper for higher odds ratios. For a fixed odds ratio, the agent's indifference curves are steeper than the Principal's (see Figure 4).¹²

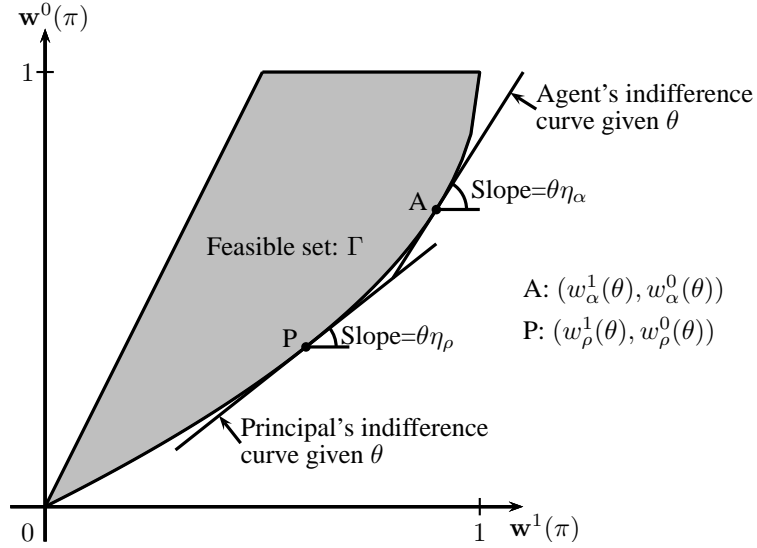


Figure 4: Indifference curves and preferred bundles

Player i 's preferred bundle given θ , denoted $(w_i^1(\theta), w_i^0(\theta))$, is the point at which his indifference curve is tangent to the southeast boundary of Γ . It is easy to verify that $(w_i^1(\theta), w_i^0(\theta))$ corresponds to a stopping-time policy with $\tau_i(\theta)$ being the stopping time. Also, for a fixed θ the agent's preferred bundle lies to the northeast of the principal's.

4.4 Delegation Problem Reformulated

Based on Lemma 1 and 2, I reformulate the delegation problem. Given that each policy can be represented by a bundle in Γ , the principal simply offers a direct mechanism $(w^1, w^0) : \Theta \rightarrow \Gamma$, called a *contract*, such that

$$\max_{w^1, w^0} \int_{\Theta} \left(\frac{\theta}{1+\theta} \eta_{\rho} w^1(\theta) - \frac{1}{1+\theta} w^0(\theta) \right) dF(\theta), \quad (5)$$

$$\text{subject to } \theta \eta_{\alpha} w^1(\theta) - w^0(\theta) \geq \theta \eta_{\alpha} w^1(\theta') - w^0(\theta'), \forall \theta, \theta' \in \Theta. \quad (6)$$

The IC constraint (6) ensures that the agent reports his type truthfully. The data relevant to this problem include: (i) two payoff parameters $\eta_{\alpha}, \eta_{\rho}$; (ii) the feasible set parametrized by r and λ^1 ; and (iii) the agent's private information θ drawn from Θ with distribution function F . The solution

¹²Parameters in Figure 4 are $\eta_{\alpha} = 3/2, \eta_{\rho} = 3/4, r/\lambda^1 = 1, \theta = \sqrt{10}/3$.

to this problem, called the optimal contract, is denoted $(w^{1*}(\theta), w^{0*}(\theta))$.¹³

5 Main Results

5.1 A Special Case: Two Types

I begin by studying the delegation problem with binary types, high type θ_h or low type θ_l , and then return to the case with a continuum. Let $q(\theta)$ denote the probability that the agent's type is θ . Formally, I solve for $(w^1, w^0) : \{\theta_l, \theta_h\} \rightarrow \Gamma$ such that

$$\begin{aligned} & \max_{w^1, w^0} \sum_{\theta \in \{\theta_l, \theta_h\}} q(\theta) \left(\frac{\theta}{1+\theta} \eta_\rho w^1(\theta) - \frac{1}{1+\theta} w^0(\theta) \right), \\ & \text{subject to } \theta_l \eta_\alpha w^1(\theta_l) - w^0(\theta_l) \geq \theta_l \eta_\alpha w^1(\theta_h) - w^0(\theta_h), \\ & \theta_h \eta_\alpha w^1(\theta_h) - w^0(\theta_h) \geq \theta_h \eta_\alpha w^1(\theta_l) - w^0(\theta_l). \end{aligned}$$

For ease of exposition, I refer to the contract for the low (high) type agent as the low (high) type contract and the principal who believes to face the low (high) type agent as the low (high) type principal. The optimum is characterized as follows.

Proposition 1 (Two types).

Suppose that $(r + \lambda^1)\theta_l \eta_\rho / r > 1$. There exists a $b' \in (1, \theta_h / \theta_l)$ such that

- 1.1 If $\eta_\alpha / \eta_\rho \in [1, b']$, the principal's preferred bundles $\{(w_\rho^1(\theta_l), w_\rho^0(\theta_l)), (w_\rho^1(\theta_h), w_\rho^0(\theta_h))\}$ are implementable.
- 1.2 If $\eta_\alpha / \eta_\rho \in (b', \theta_h / \theta_l)$, separating is optimal, i.e., $(w^{1*}(\theta_l), w^{0*}(\theta_l)) < (w^{1*}(\theta_h), w^{0*}(\theta_h))$. The low type contract is a stopping-time policy, the stopping time between $\tau_\rho(\theta_l)$ and $\tau_\alpha(\theta_l)$. The low type's IC constraint binds and the high type's does not.
- 1.3 If $\eta_\alpha / \eta_\rho \geq \theta_h / \theta_l$, pooling is optimal, i.e., $(w^{1*}(\theta_l), w^{0*}(\theta_l)) = (w^{1*}(\theta_h), w^{0*}(\theta_h))$.

In all cases, the optimum can be attained using bundles on the boundary of Γ .

Proof. See Appendix 8.2. ■

Without loss of generality, the presumption $(r + \lambda^1)\theta_l \eta_\rho / r > 1$ ensures that both the low type principal's preferred stopping time $\tau_\rho(\theta_l)$ and the high type principal's preferred stopping time

¹³Since both players' payoffs are linear in (w^1, w^0) , the optimal mechanism is deterministic.

$\tau_\rho(\theta_h)$ are strictly positive. The degenerate cases of $\tau_\rho(\theta_h) > \tau_\rho(\theta_l) = 0$ and $\tau_\rho(\theta_h) = \tau_\rho(\theta_l) = 0$ yield similar results to Proposition 1 and thus are relegated to Appendix 8.2.

Proposition 1 describes the optimal contract as the bias level varies. According to result (I.1), if the bias is low enough, the principal simply offers her preferred policies given θ_l and θ_h . This is incentive compatible because even though the low type agent prefers longer experimentation than the low type principal, at a low bias level he still prefers the low type principal's preferred bundle instead of the high type principal's. Consequently the principal pays no informational rents. This result does not hold with a continuum of types. The principal's preferred bundles are two points on the southeast boundary of Γ with binary types, but they become an interval on the southeast boundary with a continuum of types in which case lower types are strictly better off mimicking higher types.

The result (I.2) corresponds to medium bias level. As the bias has increased, offering the principal's preferred policies is no longer incentive compatible. Instead, both the low type contract and the high type one deviate from the principal's preferred policies. The low type contract is always a stopping-time policy while the high type contract takes one of three possible forms: stopping-time, slack-after-success or delay policies.¹⁴ One of the latter two forms is assigned as the high type contract if the agent's type is likely to be low and his bias is relatively large. All three forms of high type contract are meant to impose a significant cost—excessive experimentation, constrained exploitation of success, or delay in experimentation—on the high type contract so as to deter the low type agent from misreporting. However the principal can more than offset the cost by effectively shortening the low type agent's experimentation. In the end, the low type agent over-experiments slightly and the high type contract deviates from the principal's preferred policy $(w_\rho^1(\theta_h), w_\rho^0(\theta_h))$ as well. One interesting observation is that the optimal contract can take a form other than a stopping-time policy.

If the bias is even higher, as shown by result (I.3), pooling is preferable. The condition $\eta_\alpha/\eta_\rho \geq \theta_h/\theta_l$ has an intuitive interpretation that the low type agent prefers to experiment longer than even the high type principal. The screening instruments utilized in result (I.2) impair the high type principal's payoff more than the low type agent's. As a result, the principal is better off offering her uninformed preferred bundle. Notably, for fixed types the prior probabilities of the types do not affect the pooling decision. Only the bias level does.

Before moving to the continuous type case, I make two observations. First, the principal chooses to take advantage of the agent's private information unless the agent's bias is too large. This result carries over to the continuous type case. Second, the optimal contract can be tailored to

¹⁴Here, I give an example in which the high type contract is a slack-after-success policy. Parameters are $\eta_\alpha = 6, \eta_\rho = 1, \theta_l = 3/2, \theta_h = 19, r = \lambda^1 = 1$. The agent's type is low with probability 2/3. The optimum is $(w^{1*}(\theta_h), w^{0*}(\theta_h)) \approx (0.98, 1)$ and $(w^{1*}(\theta_l), w^{0*}(\theta_l)) \approx (0.96, 0.79)$.

the likelihood of the two types. For example, if the type is likely to be low, the principal designs the low type contract close to her low type bundle and purposefully makes the high type contract less attractive to the low type agent. Similarly, if the type is likely to be high, the principal starts with a high type contract close to her high type bundle without concerning about the low type's over-experimentation. This "type targeting", however, becomes irrelevant when the principal faces a continuum of types and has no incentives to target certain types.

5.2 The General Case

I return to the case that the distribution of types is represented by a continuous density f over the bounded interval $\Theta \subset \mathbf{R}_+$. The first step involves simplifying further the problem.

Given a direct mechanism $(w^1(\theta), w^0(\theta))$, let $U_\alpha(\theta)$ denote the payoff that the agent of type θ gets by maximizing over his report, *i.e.*, $U_\alpha(\theta) = \max_{\theta' \in \Theta} [\theta \eta_\alpha w^1(\theta') - w^0(\theta')]$. As the optimal mechanism is truthful, $U_\alpha(\theta)$ equals $\theta \eta_\alpha w^1(\theta) - w^0(\theta)$ and the envelope condition implies that $U'_\alpha(\theta) = \eta_\alpha w^1(\theta)$. The principal's payoff for a fixed θ is

$$\frac{\theta}{1+\theta} \eta_\rho w^1(\theta) - \frac{1}{1+\theta} w^0(\theta) = \frac{U_\alpha(\theta)}{1+\theta} + \frac{(\eta_\rho - \eta_\alpha) \theta w^1(\theta)}{1+\theta}.$$

The first term on the right-hand side corresponds to the "shared preference" between the two players because they both prefer higher w^1 value for a higher θ . The second term captures the "preference divergence" as the principal is less willing to spend resources on R in state 0 for a given increase in w^1 than the agent.

By integrating the envelope condition, one obtains the standard integral condition

$$\theta \eta_\alpha w^1(\theta) - w^0(\theta) = \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1(\tilde{\theta}) d\tilde{\theta} + \underline{\theta} \eta_\alpha w^1(\underline{\theta}) - w^0(\underline{\theta}). \quad (7)$$

Incentive compatibility of (w^1, w^0) also requires w^1 to be a nondecreasing function of θ : higher types (those types with a higher θ) are more willing to spend resources on R in state 0 for a given increase in w^1 than low types. Thus, condition (7) and the monotonicity of w^1 are necessary for incentive compatibility. As is standard, these two conditions are also sufficient.

The principal's problem is thus to maximize the expected payoff (5) subject to the feasibility constraint $w^0(\theta) \in [\beta^{\text{se}}(w^1(\theta)), \beta^{\text{nw}}(w^1(\theta))]$, the IC constraint (7), and monotonicity $w^1(\theta') \geq w^1(\theta)$ for $\theta' > \theta$. Note that this problem is convex because the expected payoff (5) is linear in $(w^1(\theta), w^0(\theta))$ and the constraint set is convex.

Substituting the IC constraint (7) into (5) and the feasibility constraint, and integrating by parts allows me to eliminate $w^0(\theta)$ from the problem except its value at $\underline{\theta}$. I denote $w^0(\underline{\theta})$ by \underline{w}^0 .

Consequently, the principal's problem reduces to finding a function $w^1 : \Theta \rightarrow [0, 1]$ and a scalar \underline{w}^0 that solves

$$\max_{w^1, \underline{w}^0 \in \Phi} \left(\eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} w^1(\theta) G(\theta) d\theta + \underline{\theta} \eta_\alpha w^1(\underline{\theta}) - \underline{w}^0 \right), \quad (\text{OBJ})$$

subject to

$$\theta \eta_\alpha w^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1(\tilde{\theta}) d\tilde{\theta} - \underline{\theta} \eta_\alpha w^1(\underline{\theta}) + \underline{w}^0 - \beta^{\text{se}}(w^1(\theta)) \geq 0, \quad \forall \theta \in \Theta, \quad (8)$$

$$\beta^{\text{nw}}(w^1(\theta)) - \left(\theta \eta_\alpha w^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1(\tilde{\theta}) d\tilde{\theta} - \underline{\theta} \eta_\alpha w^1(\underline{\theta}) + \underline{w}^0 \right) \geq 0, \quad \forall \theta \in \Theta, \quad (9)$$

where

$$\Phi \equiv \{w^1, \underline{w}^0 \mid w^1 : \Theta \rightarrow [0, 1], w^1 \text{ nondecreasing}; \underline{w}^0 \in [0, 1]\},$$

$$G(\theta) = \frac{H(\bar{\theta}) - H(\theta)}{H(\bar{\theta})} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1 \right) \theta \frac{h(\theta)}{H(\bar{\theta})}, \quad \text{where } h(\theta) = \frac{f(\theta)}{1 + \theta} \text{ and } H(\theta) = \int_{\underline{\theta}}^{\theta} h(\tilde{\theta}) d\tilde{\theta}.$$

Here, $G(\theta)$ consists of two terms. The first term is positive as it corresponds to the ‘‘shared preference’’ between the two players toward higher w^1 value for a higher θ . The second term is negative as it captures the impact of the incentive problem on the principal's expected payoff due to the agent's bias toward longer experimentation.

I denote this problem by \mathcal{P} . The set Φ is convex and includes the monotonicity constraint. Any contract $(w^1, \underline{w}^0) \in \Phi$ uniquely determines an incentive compatible direct mechanism based on (7). A contract is *admissible* if $(w^1, \underline{w}^0) \in \Phi$ and the feasibility constraint, (8) and (9), is satisfied.

5.3 A Robust Result: Pooling on Top

With a continuum of types, I first show that types above some threshold are offered the same (w^1, w^0) bundle. Intuitively, types at the very top prefer to experiment more than what the principal prefers to do for any prior. Therefore, the cost of separating those types exceeds the benefit. This can be seen from the fact that the first term—the ‘‘shared preference’’ term—of $G(\theta)$ reduces to 0 as θ approaches $\bar{\theta}$. As a result, the principal finds it optimal to pool those types at the very top.

Let θ_p be the lowest value in Θ such that

$$\int_{\hat{\theta}}^{\bar{\theta}} G(\theta) d\theta \leq 0, \quad \text{for any } \hat{\theta} \geq \theta_p. \quad (10)$$

My next result shows that types with $\theta \geq \theta_p$ are pooled.

Proposition 2 (Pooling on top).

An optimal contract $(w^{1*}, \underline{w}^{0*})$ satisfies $w^{1*}(\theta) = w^{1*}(\theta_p)$ for $\theta \geq \theta_p$. It is optimal for (8) or (9) to hold with equality at θ_p .

Proof. The contribution to (OBJ) from types with $\theta > \theta_p$ is $\eta_\alpha \int_{\theta_p}^{\bar{\theta}} w^1(\theta)G(\theta)d\theta$. Substituting $w^1(\theta) = \int_{\theta_p}^{\theta} dw^1 + w^1(\theta_p)$ and integrating by parts, I obtain

$$\eta_\alpha w^1(\theta_p) \int_{\theta_p}^{\bar{\theta}} G(\theta)d\theta + \eta_\alpha \int_{\theta_p}^{\bar{\theta}} \int_{\theta}^{\bar{\theta}} G(\tilde{\theta})d\tilde{\theta}dw^1(\theta). \quad (11)$$

The first term only depends on $w^1(\theta_p)$. The second term depends on $dw^1(\theta)$ for all $\theta \in [\theta_p, \bar{\theta}]$. According to the definition of θ_p , the integrand of the second term, $\int_{\theta}^{\bar{\theta}} G(\tilde{\theta})d\tilde{\theta}$, is weakly negative for all $\theta \in [\theta_p, \bar{\theta}]$. Therefore, it is optimal to set $dw^1(\theta) = 0$ for all $\theta \in [\theta_p, \bar{\theta}]$. If $\theta_p = \underline{\theta}$, all types are pooled. The principal offers her preferred uninformed bundle, which is on the southeast boundary of Γ . If $\theta_p > \underline{\theta}$, the first term of (11) is zero as well because $\int_{\theta_p}^{\bar{\theta}} G(\theta)d\theta = 0$. Adjusting $w^1(\theta_p)$ does not affect the objective function, so $w^1(\theta_p)$ can be increased until either (8) or (9) binds. ■

The slope of the principal's indifference curves is bounded from above by $\bar{\theta}\eta_\rho$, which is the slope if she believes that the agent's type is $\bar{\theta}$. An agent whose type is above $\bar{\theta}\eta_\rho/\eta_\alpha$ has indifference curves with slope steeper than $\bar{\theta}\eta_\rho$. The following corollary states that types above $\bar{\theta}\eta_\rho/\eta_\alpha$ are offered the same bundle.

Corollary 1. *The threshold of the top pooling segment, θ_p , is below $\bar{\theta}\eta_\rho/\eta_\alpha$.*

Proof. See Appendix 8.3. ■

Note that for a fixed type distribution, the value of θ_p depends only on the ratio η_ρ/η_α but not on the magnitudes of η_ρ, η_α . If $\eta_\rho/\eta_\alpha = 1$, both parties' preferences are perfectly aligned. The function G is positive for any θ , and thus the principal optimally sets θ_p to be $\bar{\theta}$. As η_ρ/η_α decreases, the agent's bias grows. The principal enlarges the top pooling segment by lowering θ_p . When η_ρ/η_α is sufficiently close to zero, the principal optimally sets θ_p to be $\underline{\theta}$ in which case all types are pooled.

Corollary 2. *For a fixed type distribution, θ_p increases in η_ρ/η_α . Moreover, $\theta_p = \bar{\theta}$ if $\eta_\rho/\eta_\alpha = 1$, and there exists $z^* \in [0, 1)$ such that $\theta_p = \underline{\theta}$ if $\eta_\rho/\eta_\alpha \leq z^*$.*

Proof. See Appendix 8.4. ■

If $\theta_p = \underline{\theta}$, all types are pooled. The optimal contract consists of the principal's preferred uninformed bundle. For the rest of this section, I focus on the more interesting case in which $\theta_p > \underline{\theta}$.

5.4 Imposing a Cutoff

To make progress, I assume that the distribution of types satisfies the following condition. In Subsection 5.4.4, I will examine how results change when this condition fails.

Assumption 2. For all $\theta \leq \theta_p$, $1 - G(\theta)$ is nondecreasing.

When the density function f is differentiable, Assumption 2 is equivalent to the following condition:

$$\theta \frac{f'(\theta)}{f(\theta)} \geq - \left(\frac{\eta_\alpha}{\eta_\alpha - \eta_\rho} + \frac{1}{1 + \theta} \right), \quad \forall \theta \leq \theta_p.$$

This condition is satisfied for all density functions that are nondecreasing and holds for the exponential distribution, the log-normal, the Pareto and the Gamma distribution for a subset of their parameters. Also, it is satisfied for any density f with $\theta f'/f$ bounded from below when η_α/η_ρ is sufficiently close to 1.

My next result (Proposition 3) shows that under Assumption 2 the optimal contract takes a very simple form. To describe it formally, I introduce the following:

Definition 1. The *cutoff rule* is the contract (w^1, w^0) such that

$$(w^1(\theta), w^0(\theta)) = \begin{cases} (w_\alpha^1(\theta), w_\alpha^0(\theta)) & \text{if } \theta \leq \theta_p, \\ (w_\alpha^1(\theta_p), w_\alpha^0(\theta_p)) & \text{if } \theta > \theta_p. \end{cases}$$

Under the cutoff rule, types with $\theta \leq \theta_p$ are offered their preferred bundles $(w_\alpha^1(\theta), w_\alpha^0(\theta))$ whereas types with $\theta > \theta_p$ are pooled at $(w_\alpha^1(\theta_p), w_\alpha^0(\theta_p))$. I denote the cutoff rule by $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$.

Figure 5 shows the delegation set corresponding to the cutoff rule.¹⁵ With a slight abuse of notation, I identify a bundle on the southeast boundary of Γ with the slope of the tangent line at that point. As θ varies, the principal's preferred bundle ranges from $\underline{\theta}\eta_\rho$ to $\bar{\theta}\eta_\rho$ and the agent's ranges from $\underline{\theta}\eta_\alpha$ to $\bar{\theta}\eta_\alpha$. The delegation set is the interval between $\underline{\theta}\eta_\alpha$ and $\theta_p\eta_\alpha$. According to Corollary 1, $\theta_p\eta_\alpha$ is smaller than $\bar{\theta}\eta_\rho$. Therefore, the upper bound of the delegation set is lower than $\bar{\theta}\eta_\rho$, the principal's preferred bundle given the highest type $\bar{\theta}$.

The next proposition shows that $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$ is the optimum under Assumption 2.

Proposition 3 (Sufficiency).

The cutoff rule $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$ is optimal if Assumption 2 holds.

In what follows, I first illustrate how to implement the cutoff rule and prove that it is time-consistent. Next, I present the proof of Proposition 3. Then, I discuss the condition required by Assumption 2 and how results change if this assumption fails.

¹⁵Parameters in Figure 5 are $\eta_\alpha = 1, \eta_\rho = 3/5, r/\lambda^1 = 1, \underline{\theta} = 1, \bar{\theta} = 5$. The type variable θ is uniformly distributed. The pooling threshold is $\theta_p \approx 1.99$.

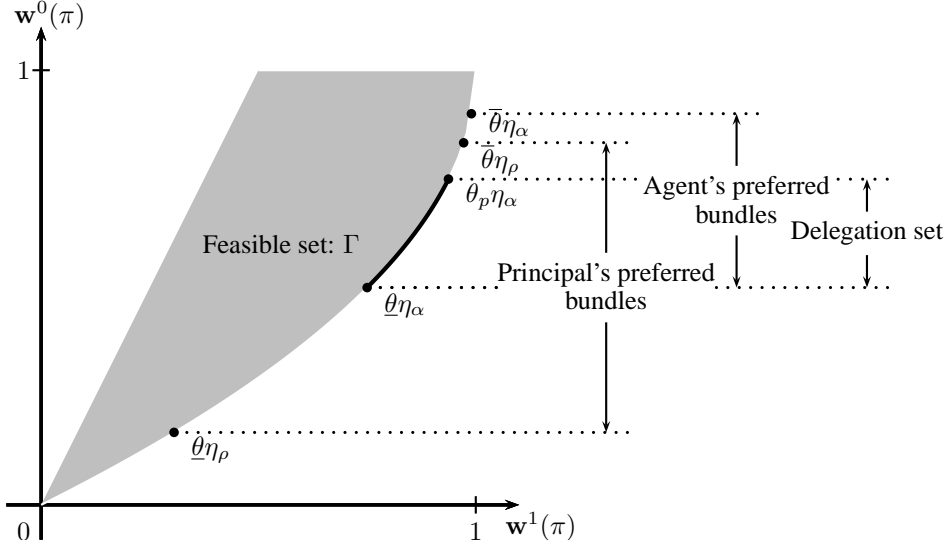


Figure 5: Delegation set under cutoff rule

5.4.1 Properties of the Cutoff Rule

Implementation. Under the cutoff rule, the agent need not report his type at time 0. Instead, the optimal outcome for the principal can be implemented indirectly by calibrating a *constructed belief* that the state is 1. It starts with the prior belief $p(\theta_p \eta_\alpha / \eta_\rho)$ and then is updated as if the agent had no private information about the state. More specifically, if no success occurs this belief is downgraded according to the differential equation $\dot{p}_t = -\lambda^1 p_t (1 - p_t)$. Upon the first success this belief jumps to one.

The principal imposes a cutoff at p_ρ^* . As long as the constructed belief stays above the cutoff, the agent can decide whether to continue experimenting or not. As soon as it drops to the cutoff, the agent is not allowed to operate R any more. This rule does not bind for those types below θ_p , who switch to S voluntarily conditional on no success, but does constrain those types above θ_p , who are forced to stop by the principal.

Figure 6 illustrates how the constructed belief evolves over time. The solid arrow shows that the belief is downgraded when R is used and no success has realized. The dashed arrow shows that the belief jumps to one at the first success. The gray area shows that those types below θ_p stop voluntarily as their posterior beliefs drop to the agent's cutoff p_α^* . As illustrated by the black dot, a mass of higher types are required to stop when the cutoff is reached.

There are many other ways to implement the cutoff rule. For example, the constructed belief may start with the prior belief $p(\theta_p)$ and the principal imposes a cutoff at p_α^* . What matters is that the prior belief and the cutoff are chosen collectively to ensure that exactly those types below θ_p are given the freedom to decide whether to experiment or not.

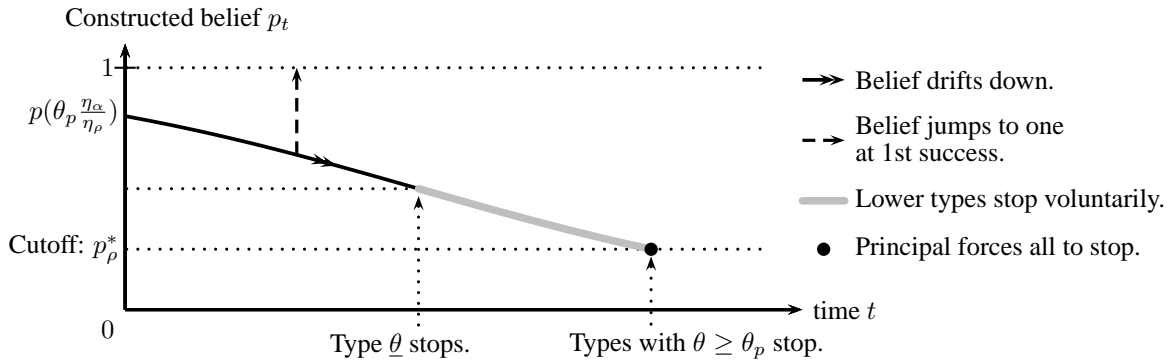


Figure 6: Implementing the cutoff rule

Time consistency. Given that the agent need not report his type at time 0 and that the principal commits to the cutoff rule before experimentation, a natural question to follow is whether the cutoff rule is *time-consistent*. Does the principal find it optimal to fulfill the contract that she commits to at time 0 after any history? Put differently, were the principal given a chance to adjust the contract, would she choose not to do so? As my next result shows, the cutoff rule is indeed time-consistent.

Proposition 4 (Time consistency).

If Assumption 2 holds, the cutoff rule is time-consistent.

To show time-consistency, I need to consider three classes of histories on the equilibrium path: (i) the first success occurs before the cutoff is reached; (ii) the agent stops experimenting before the cutoff is reached; (iii) no success has occurred and the agent has not stopped. Clearly, the principal has no incentives to alter the contract after the first two classes of histories. Upon the arrival of the first success, it is optimal to let the agent use R exclusively thereafter. Also, if the agent stops experimenting before the cutoff is reached, his type is revealed. From the principal's point of view, the agent already over-experiments. Hence, she has no incentives to ask the agent to use R any more.

If the agent has not stopped and there is no success, it is unclear whether the principal still finds the cutoff rule set at time 0 to be optimal. On the one hand, the principal learns from the lack of success that the agent's type is more likely to be low. One might expect that the principal wishes at this moment that she had set up a higher cutoff at time 0. On the other, as time elapses without success, the constructed belief is downgraded and hence there is less time left before the cutoff is reached. This might eliminate the need to adjust the cutoff. It turns out that these two factors exactly cancel out. After updating the constructed belief, the principal finds it optimal to keep the cutoff intact.

Here, I sketch the proof. First, I calculate the principal's updated belief about the type dis-

tribution given no success and that the agent has not stopped. By continuing experimenting, the agent signals that his type is above some level. Hence, the updated type distribution is a truncated one. Since the agent has also updated his belief about the state, I then rewrite the type distribution in terms of the agent's updated odds ratio. Next I show that given the new type distribution the optimal contract is to continue the cutoff rule set at time 0.

Let $\theta_\alpha^* = p_\alpha^*/(1 - p_\alpha^*)$ be the odds ratio at which the agent is indifferent between continuing and stopping. After operating R for $\delta > 0$ without success, the agent of type θ updates his belief about the state. He assigns odds ratio $\theta e^{-\lambda^1 \delta}$ to the state being 1, referred to as his type at time δ . Let $\underline{\theta}_\delta = \max\{\underline{\theta}, \theta_\alpha^* e^{\lambda^1 \delta}\}$. After a period of δ with no success, only those types above $\underline{\theta}_\delta$ remain. The principal's updated belief about the agent's type distribution, *in terms of his type at time 0*, is given by the density function

$$f(\theta | \delta) = \begin{cases} \frac{[1-p(\theta)(1-e^{-\lambda^1 \delta})]f(\theta)}{\int_{\underline{\theta}_\delta}^{\bar{\theta}} [1-p(\theta)(1-e^{-\lambda^1 \delta})]f(\theta)d\theta} & \text{if } \theta \in [\underline{\theta}_\delta, \bar{\theta}], \\ 0 & \text{otherwise.} \end{cases}$$

Here, $1 - p(\theta)(1 - e^{-\lambda^1 \delta})$ is the probability that no success occurs from time 0 to δ conditional on the agent's type being θ at time 0. The principal's belief about the agent's type distribution, *in terms of his type at time δ* , is given by the density function

$$f_\delta(\theta) = \begin{cases} f(\theta e^{\lambda^1 \delta} | \delta) e^{\lambda^1 \delta} & \text{if } \theta \in [\underline{\theta}_\delta e^{-\lambda^1 \delta}, \bar{\theta} e^{-\lambda^1 \delta}], \\ 0 & \text{otherwise.} \end{cases}$$

I prove that continuing the cutoff rule is optimal by showing two things. First, given the distribution f_δ at time δ , the threshold of the top pooling segment is $\theta_p e^{-\lambda^1 \delta}$. Second, if Assumption 2 holds for $\theta \leq \theta_p$ under distribution f , then it holds for $\theta \leq \theta_p e^{-\lambda^1 \delta}$ under f_δ . The detailed proof is relegated to Appendix 8.5.

Over- and under-experimentation. The result of Proposition 3 can also be represented by a delegation rule mapping types into stopping times as only stopping-time policies are assigned in equilibrium. Figure 7 depicts such a rule. The x -axis variable is θ , ranging from $\underline{\theta}$ to $\bar{\theta}$. The dotted line represents the agent's preferred stopping time and the dashed line represents the principal's.¹⁶ The delegation rule consists of (i) segment $[\underline{\theta}, \theta_p]$ where the stopping time equals the agent's preferred stopping time and (ii) segment $[\theta_p, \bar{\theta}]$ where the stopping time is independent of the agent's report (*i.e.*, pooling segment). To implement, the principal simply imposes a deadline at $\tau_\alpha(\theta_p)$.

¹⁶Parameters in Figure 7 are $\eta_\alpha = 6/5, \eta_p = 1, r/\lambda^1 = 1, \underline{\theta} = 1, \bar{\theta} = 5$. The type variable θ is uniformly distributed.

Those types with $\theta \geq \theta_p$ all stop at $\tau_\alpha(\theta_p)$, which is the principal's preferred stopping time if she believes that the agent's type is *above* θ_p . Since the principal's stopping time given type $\theta_p \eta_\alpha / \eta_\rho$ equals the agent's stopping time given θ_p , the delegation rule intersects the principal's stopping time at type $\theta_p \eta_\alpha / \eta_\rho$. From the principal's point of view, those types with $\theta < \theta_p \eta_\alpha / \eta_\rho$ experiment too long while those types with $\theta > \theta_p \eta_\alpha / \eta_\rho$ stop too early.

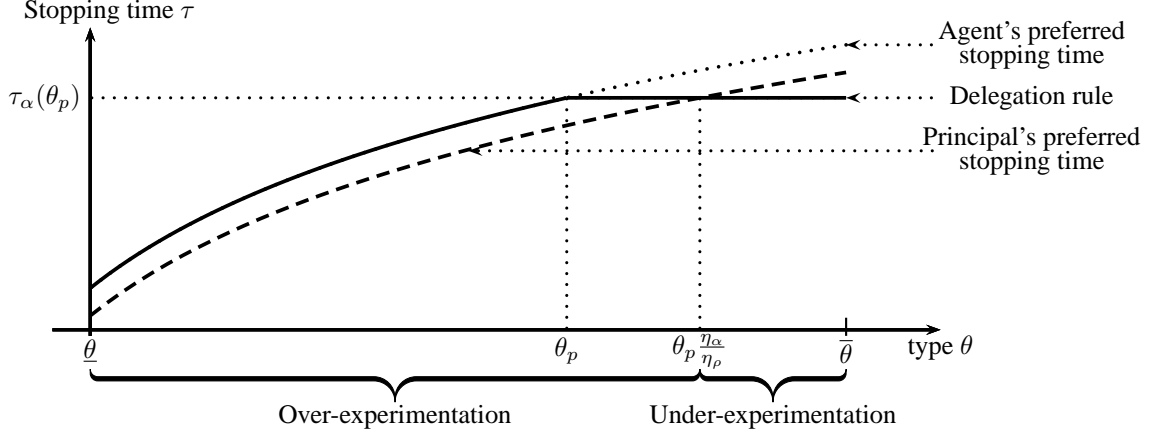


Figure 7: Equilibrium stopping times

5.4.2 Proof of Proposition 3: the Optimality of the Cutoff Rule

To prove Proposition 3, I utilize Lagrangian optimization methods (similar to those used by Amador, Werning, and Angeletos (2006) [4]). It suffices to show that $(w_{\theta_p}^1, w_{\theta_p}^0)$ maximizes some Lagrangian functional. Then I establish the sufficient first-order conditions and prove that they are satisfied at the conjectured contract and Lagrange multipliers.

I first extend β^{se} to the real line in the following way:

$$\hat{\beta}(w^1) = \begin{cases} (\beta^{\text{se}})'(0)w^1 & \text{if } w^1 \in (-\infty, 0), \\ \beta^{\text{se}}(w^1) & \text{if } w^1 \in [0, \hat{w}^1], \\ \beta^{\text{se}}(\hat{w}^1) + (\beta^{\text{se}})'(\hat{w}^1)(w^1 - \hat{w}^1) & \text{if } w^1 \in (\hat{w}^1, \infty). \end{cases}$$

for some value \hat{w}^1 such that $\hat{w}^1 \in (w_\alpha^1(\bar{\theta}), 1)$.¹⁷ The newly defined function $\hat{\beta}$ is continuously differentiable, convex and lower than β^{se} on $[0, 1]$.

I then define a new problem $\hat{\mathcal{P}}$ which differs from \mathcal{P} in two aspects: (i) the upper bound

¹⁷Such a \hat{w}^1 exists because $\bar{\theta}$ is finite and hence $w_\alpha^1(\bar{\theta})$ is bounded away from 1.

constraint (9) is dropped; and (ii) the lower bound constraint (8) is replaced with the following:

$$\theta\eta_\alpha w^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1(\tilde{\theta})d\tilde{\theta} - \theta\eta_\alpha w^1(\underline{\theta}) + \underline{w}^0 - \hat{\beta}(w^1(\theta)) \geq 0, \quad \forall \theta \in \Theta. \quad (12)$$

If (w^1, \underline{w}^0) satisfies the feasibility constraint (8) and (9), it also satisfies (12). Therefore, the newly defined problem $\hat{\mathcal{P}}$ is a relaxation of \mathcal{P} . If the solution to $\hat{\mathcal{P}}$ is admissible, I claim that it is also the solution to \mathcal{P} .

Define the Lagrangian functional associated with $\hat{\mathcal{P}}$ as

$$\begin{aligned} \hat{L}(w^1, \underline{w}^0 \mid \Lambda) &= \theta\eta_\alpha w^1(\underline{\theta}) - \underline{w}^0 + \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} w^1(\theta)G(\theta)d\theta \\ &+ \int_{\underline{\theta}}^{\bar{\theta}} \left(\theta\eta_\alpha w^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1(\tilde{\theta})d\tilde{\theta} - \theta\eta_\alpha w^1(\underline{\theta}) + \underline{w}^0 - \hat{\beta}(w^1(\theta)) \right) d\Lambda, \end{aligned}$$

where the function Λ is the Lagrange multiplier associated with (12). Fixing a nondecreasing multiplier Λ , the Lagrangian is a concave functional on Φ because all terms in $\hat{L}(w^1, \underline{w}^0 \mid \Lambda)$ are linear in (w^1, \underline{w}^0) except $\int_{\underline{\theta}}^{\bar{\theta}} -\hat{\beta}(w^1(\theta))d\Lambda$ which is concave in w^1 . Without loss of generality I set $\Lambda(\bar{\theta}) = 1$. Integrating the Lagrangian by parts yields

$$\begin{aligned} \hat{L}(w^1, \underline{w}^0 \mid \Lambda) &= (\theta\eta_\alpha w^1(\underline{\theta}) - \underline{w}^0) \Lambda(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \left(\theta\eta_\alpha w^1(\theta) - \hat{\beta}(w^1(\theta)) \right) d\Lambda \quad (13) \\ &+ \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} w^1(\theta) [\Lambda(\theta) - (1 - G(\theta))] d\theta. \end{aligned}$$

The following lemma provides a sufficient condition for a contract $(\tilde{w}^1, \tilde{\underline{w}}^0) \in \Phi$ to solve $\hat{\mathcal{P}}$.

Lemma 3 (Lagrangian—sufficiency).

A contract $(\tilde{w}^1, \tilde{\underline{w}}^0) \in \Phi$ solves $\hat{\mathcal{P}}$ if (12) holds with equality and there exists a nondecreasing $\tilde{\Lambda}$ such that

$$\hat{L}(\tilde{w}^1, \tilde{\underline{w}}^0 \mid \tilde{\Lambda}) \geq \hat{L}(w^1, \underline{w}^0 \mid \tilde{\Lambda}), \quad \forall (w^1, \underline{w}^0) \in \Phi.$$

Proof. I first introduce the problem studied in section 8.4 of Luenberger (1969, p. 220) [21]: $\max_{x \in X} Q(x)$ subject to $x \in \Omega$ and $J(x) \in P$, where Ω is a subset of the vector space X , $Q : \Omega \rightarrow \mathbf{R}$ and $J : \Omega \rightarrow Z$; where Z is a normed vector space, and P is a nonempty positive

cone in Z . To apply Theorem 1 in Luenberger (1969, p. 220) [21], set

$$X = \{w^1, \underline{w}^0 \mid w^1 : \Theta \rightarrow \mathbf{R} \text{ and } \underline{w}^0 \in \mathbf{R}\}, \quad (14)$$

$$\Omega = \Phi, \quad (15)$$

$$Z = \{z \mid z : \Theta \rightarrow \mathbf{R} \text{ with } \sup_{\theta \in \Theta} |z(\theta)| < \infty\},$$

$$\text{with the norm } \|z\| = \sup_{\theta \in \Theta} |z(\theta)|,$$

$$P = \{z \mid z \in Z \text{ and } z(\theta) \geq 0, \forall \theta \in \Theta\}.$$

I let the objective function in (OBJ) be Q and let the left-hand side of (12) be defined as J . This result holds because the hypotheses of Theorem 1 in Luenberger (1969, p. 220) [21] are met. ■

To apply Lemma 3 and show that a proposed contract $(\tilde{w}^1, \tilde{w}^0)$ maximizes $\hat{L}(w^1, \underline{w}^0 | \tilde{\Lambda})$ for some candidate Lagrangian multiplier $\tilde{\Lambda}$, I modify Lemma 1 in Luenberger (1969, p. 227) [21] which concerns the maximization of a concave functional in a convex cone. Note that set Φ is not a convex cone, so Lemma 1 in Luenberger (1969, p. 227) [21] does not apply directly in the current setting.

Lemma 4 (First-order conditions).

Let L be a concave functional on Ω , a convex subset of a vector space X . Take $\tilde{x} \in \Omega$. Suppose that the Gâteaux differentials $\partial L(\tilde{x}; x)$ and $\partial L(\tilde{x}; x - \tilde{x})$ exist for any $x \in \Omega$ and that $\partial L(\tilde{x}; x - \tilde{x}) = L(\tilde{x}; x) - L(\tilde{x}; \tilde{x})$.¹⁸ A sufficient condition that $\tilde{x} \in \Omega$ maximizes L over Ω is that

$$\partial L(\tilde{x}; x) \leq 0, \quad \forall x \in \Omega,$$

$$\partial L(\tilde{x}; \tilde{x}) = 0.$$

Proof. See Appendix 8.6. ■

Next, I prove Proposition 3 based on Lemma 3 and Lemma 4.

Proof. To apply Lemma 4, let X and Ω be the same as in (14) and (15). Fixing a nondecreasing multiplier Λ , the Lagrangian (13) is a concave functional on Φ . By applying Lemma A.1 in Amador, Werning, and Angeletos (2006) [4], it is easy to verify that $\partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; w^1, \underline{w}^0 \mid \Lambda)$ and

¹⁸Let X be a vector space, Y a normed space and $D \subset X$. Given a transformation $T : D \rightarrow Y$, if for $\tilde{x} \in D$ and $x \in X$ the limit

$$\lim_{\epsilon \rightarrow 0} \frac{T(\tilde{x} + \epsilon x) - T(\tilde{x})}{\epsilon}$$

exists, then it is called the Gâteaux differential at \tilde{x} with direction x and is denoted $\partial T(\tilde{x}; x)$. If the limit exists for each $x \in X$, T is said to be Gâteaux differentiable at \tilde{x} .

$\partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; (w^1, \underline{w}^0) - (w_{\theta_p}^1, \underline{w}_{\theta_p}^0) \mid \Lambda)$ exist for any $(w^1, \underline{w}^0) \in \Phi$.¹⁹ The linearity condition is also satisfied

$$\partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; (w^1, \underline{w}^0) - (w_{\theta_p}^1, \underline{w}_{\theta_p}^0) \mid \Lambda) = \partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; w^1, \underline{w}^0 \mid \Lambda) - \partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; w_{\theta_p}^1, \underline{w}_{\theta_p}^0 \mid \Lambda).$$

So the hypotheses of Lemma 4 are met. Also, the Gâteaux differential at $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$ is given by

$$\begin{aligned} \partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; w^1, \underline{w}^0 \mid \Lambda) &= (\theta \eta_\alpha w^1(\underline{\theta}) - \underline{w}^0) \Lambda(\underline{\theta}) + \eta_\alpha \int_{\theta_p}^{\bar{\theta}} (\theta - \theta_p) w^1(\theta) d\Lambda \\ &+ \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} w^1(\theta) [\Lambda(\theta) - (1 - G(\theta))] d\theta, \quad \forall (w^1, \underline{w}^0) \in \Phi. \end{aligned} \quad (16)$$

Next, I construct a nondecreasing multiplier $\tilde{\Lambda}$, in a similar manner as in Proposition 3 in Amador, Werning, and Angeletos (2006) [4], such that the first-order conditions $\partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; w^1, \underline{w}^0 \mid \tilde{\Lambda}) \leq 0$ and $\partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; w_{\theta_p}^1, \underline{w}_{\theta_p}^0 \mid \tilde{\Lambda}) = 0$ are satisfied for any $(w^1, \underline{w}^0) \in \Phi$.

Let $\tilde{\Lambda}(\underline{\theta}) = 0$, $\tilde{\Lambda}(\theta) = 1 - G(\theta)$ for $(\underline{\theta}, \theta_p]$, and $\tilde{\Lambda}(\theta) = 1$ for $\theta \in (\theta_p, \bar{\theta}]$. Given that $\theta_p > \underline{\theta}$, I need to show that jumps at $\underline{\theta}$ and θ_p are upward. The jump at $\underline{\theta}$ is upward since $1 - G(\underline{\theta})$ is nonnegative. The jump at θ_p is $G(\theta_p)$, which is nonnegative based on the definition of θ_p . Therefore, $\tilde{\Lambda}$ is nondecreasing.

Substituting the multiplier $\tilde{\Lambda}$ into the Gâteaux differential (16) yields

$$\begin{aligned} \partial \hat{L}(w_{\theta_p}^1, \underline{w}_{\theta_p}^0; w^1, \underline{w}^0 \mid \tilde{\Lambda}) &= \eta_\alpha \int_{\theta_p}^{\bar{\theta}} w^1(\theta) G(\theta) d\theta \\ &= \eta_\alpha \int_{\theta_p}^{\bar{\theta}} \left(\int_{\underline{\theta}}^{\bar{\theta}} G(\tilde{\theta}) d\tilde{\theta} \right) dw^1(\theta), \end{aligned}$$

where the last equality follows by integrating by parts, which can be done given the monotonicity of w^1 and by the definition of θ_p . This Gâteaux differential is zero at $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$ and, by the definition of θ_p , it is nonpositive for all w^1 nondecreasing. It follows that the first-order conditions are satisfied for all $(w^1, \underline{w}^0) \in \Phi$. By Lemma 4, $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$ maximizes $\hat{L}(w^1, \underline{w}^0 \mid \tilde{\Lambda})$ over Φ . By Lemma 3, $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$ solves $\hat{\mathcal{P}}$. Because $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$ is admissible, it solves \mathcal{P} . ■

5.4.3 Discussion of the Results

Here, I explain the intuitions for the cutoff rule $(w_{\theta_p}^1, \underline{w}_{\theta_p}^0)$ to be optimal, particularly how the threshold of the pooling segment θ_p is determined and what is the economic content of Assumption

¹⁹The observations made by Ambrus and Egorov (2013) [5] do not apply to the results of Amador, Werning and Angeletos (2006) [4] that my proof relies on.

2. To do so, I consider a delegation problem in which only bundles on the southeast boundary of Γ are considered. This restriction reduces the action space to a one-dimensional line segment and allows me to compare the condition in Assumption 2 with those identified in Alonso and Matouschek (2008) [2].

Recall that Γ^{st} denotes the southeast boundary of Γ , characterized by $\beta^{\text{se}}(w^1) = 1 - (1 - w^1)^{r/(r+\lambda^1)}$ such that

$$\Gamma^{\text{st}} = \{(w^1, w^0) \mid w^0 = \beta^{\text{se}}(w^1), w^1 \in [0, 1]\}.$$

The function β^{se} is twice continuously differentiable with $\beta^{\text{se}}(0)$ and $\beta^{\text{se}}(1)$ being 0 and 1 respectively. The derivative $(\beta^{\text{se}})'(w^1)$ strictly increases in w^1 and approaches infinity as w^1 approaches 1. I identify an element $(w^1, \beta^{\text{se}}(w^1)) \in \Gamma^{\text{st}}$ with the derivative $(\beta^{\text{se}})'(w^1)$ at that point. The set of possible derivatives is denoted $Y = [(\beta^{\text{se}})'(0), \infty]$. Since there is a one-to-one mapping between Γ^{st} and Y , I let Y be the action space and refer to $y \in Y$ as an action. The principal simply assigns a non-empty subset of Y as the permissible set. Let $n(y) = ((\beta^{\text{se}})')^{-1}(y)$ be the inverse of the mapping from w^1 to the derivative $(\beta^{\text{se}})'(w^1)$.

Player i 's preferred action given θ is $y_i(\theta) = \eta_i \theta$. Player i 's payoff given type θ and action y is denoted

$$V_i(\theta, y) = \eta_i \frac{\theta}{1 + \theta} n(y) - \frac{1}{1 + \theta} \beta^{\text{se}}(n(y)).$$

I first solve the principal's preferred action if she believes that the agent's type is below θ . The principal chooses $y \in Y$ to maximize $\int_{\underline{\theta}}^{\theta} V_{\rho}(\tilde{\theta}, y) f(\tilde{\theta}) d\tilde{\theta}$. The maximum is achieved by choosing action

$$\eta_{\rho} \frac{\int_{\underline{\theta}}^{\theta} \tilde{\theta} h(\tilde{\theta}) d\tilde{\theta}}{H(\theta)}.$$

Following Alonso and Matouschek (2008) [2], I define the *backward bias* for a given type θ as

$$T(\theta) \equiv \frac{H(\theta)}{H(\bar{\theta})} \left(\eta_{\alpha} \theta - \eta_{\rho} \frac{\int_{\underline{\theta}}^{\theta} \tilde{\theta} h(\tilde{\theta}) d\tilde{\theta}}{H(\theta)} \right).$$

Here, $T(\theta)$ measures the difference between the agent's preferred action given θ and the principal's preferred action if she believes that the type is below θ . Similarly, the *forward bias*

$$R(\theta) \equiv \frac{H(\bar{\theta}) - H(\theta)}{H(\bar{\theta})} \left(\eta_{\alpha} \theta - \eta_{\rho} \frac{\int_{\underline{\theta}}^{\bar{\theta}} \tilde{\theta} h(\tilde{\theta}) d\tilde{\theta}}{H(\bar{\theta}) - H(\theta)} \right)$$

measures the difference between the agent's preferred action given θ and the principal's preferred action if she believes that the type is above θ . Since players' preferred actions increase in θ and the

agent is biased toward larger actions, the backward bias is strictly positive for any θ . The sign of the forward bias depends on the parameters. It is easy to verify that

$$T(\theta) = \eta_\alpha \int_{\underline{\theta}}^{\theta} (1 - G(\tilde{\theta})) d\tilde{\theta},$$

$$R(\theta) = -\eta_\alpha \int_{\theta}^{\bar{\theta}} G(\tilde{\theta}) d\tilde{\theta}.$$

Hence, according to the definition of θ_p , θ_p is the lowest type $\hat{\theta}$ such that $R(\theta) \geq 0$ for any $\theta \in [\hat{\theta}, \bar{\theta}]$. If $\theta_p > \underline{\theta}$, $R(\theta_p) = 0$. The definition of θ_p ensures that for any $\theta > \theta_p$, the agent's preferred action given θ is larger than the principal's preferred action if she believes that the type is above θ . Therefore, the principal finds it optimal to push the threshold of the pooling segment until it reaches θ_p .

Assumption 2 is equivalent to requiring that the backward bias is convex when $\theta \leq \theta_p$. This is the condition that Alonso and Matouschek (2008) [2] find for the interval delegation set to be optimal in their setting. Intuitively, when this condition holds, the principal finds it optimal to fill in the “holes” in the delegate set. I shall emphasize that this is *not* a proof of the optimality of the cutoff rule, because considering only bundles on the southeast boundary might be restrictive. For example, I have shown that with two types the optimal contract involves bundles which are not on the southeast boundary for certain parameter values. With a continuum of types, there exist examples such that the principal is strictly better off by offering policies which are not necessarily stopping-time policies. By using Lagrangian methods, I prove that the cutoff rule is indeed optimal under Assumption 2. In my setting, the principal's preferred bundle is not more sensitive than the agent's to the agent's private information and Assumption 2 ensures that the type distribution is sufficiently smooth so the principal has no particular interest to screen some types. Hence, the interval delegation set is optimal.

5.4.4 Results if Assumption 2 Fails

I first define the x_p -cutoff contract, given by $(w^1(\theta), w^0(\theta)) = (w_\alpha^1(\theta), w_\alpha^0(\theta))$ for $\theta < x_p$ and $(w^1(\theta), w^0(\theta)) = (w_\alpha^1(x_p), w_\alpha^0(x_p))$ for $\theta \geq x_p$. The x_p -cutoff contract is denoted $(w_{x_p}^1, w_{x_p}^0)$. The purpose of this subsection is to show that no x_p -cutoff contract is optimal for any $x_p \in \Theta$ if Assumption 2 does not hold.

Define the Lagrangian functional associated with \mathcal{P} as

$$\begin{aligned}
L(w^1, \underline{w}^0 \mid \Lambda^{se}, \Lambda^{nw}) &= \underline{\theta} \eta_\alpha w^1(\underline{\theta}) - w^0(\underline{\theta}) + \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} w^1(\theta) G(\theta) d\theta \\
&+ \int_{\underline{\theta}}^{\bar{\theta}} \left(\theta \eta_\alpha w^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1(\tilde{\theta}) d\tilde{\theta} - \underline{\theta} \eta_\alpha w^1(\underline{\theta}) + \underline{w}^0 - \beta^{se}(w^1(\theta)) \right) d\Lambda^{se} \\
&+ \int_{\underline{\theta}}^{\bar{\theta}} \left[\beta^{nw}(w^1(\theta)) - \left(\theta \eta_\alpha w^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1(\tilde{\theta}) d\tilde{\theta} - \underline{\theta} \eta_\alpha w^1(\underline{\theta}) + \underline{w}^0 \right) \right] d\Lambda^{nw},
\end{aligned} \tag{17}$$

where the function $\Lambda^{se}, \Lambda^{nw}$ are the Lagrange multiplier associated with constraints (8) and (9). I first show that if $(w_{x_p}^1, \underline{w}_{x_p}^0)$ is optimal for some x_p , there must exist some Lagrange multipliers $\tilde{\Lambda}^{se}, \tilde{\Lambda}^{nw}$ such that $L(w^1, \underline{w}^0 \mid \tilde{\Lambda}^{se}, \tilde{\Lambda}^{nw})$ is maximized at $(w_{x_p}^1, \underline{w}_{x_p}^0)$. Since any x_p -cutoff contract is continuous, I can restrict attention to the set of continuous contracts

$$\hat{\Phi} \equiv \{w^1, \underline{w}^0 \mid w^1 : \Theta \rightarrow [0, 1], w^1 \text{ nondecreasing and continuous; } \underline{w}^0 \in [0, 1]\}.$$

Lemma 5 (Lagrangian—necessity).

If $(w_{x_p}^1, \underline{w}_{x_p}^0)$ solves \mathcal{P} , then there exist two nondecreasing functions $\tilde{\Lambda}^{se}, \tilde{\Lambda}^{nw} : \Theta \rightarrow \mathbf{R}$ such that

$$L(w_{x_p}^1, \underline{w}_{x_p}^0 \mid \tilde{\Lambda}^{se}, \tilde{\Lambda}^{nw}) \geq L(w^1, \underline{w}^0 \mid \tilde{\Lambda}^{se}, \tilde{\Lambda}^{nw}), \quad \forall (w^1, \underline{w}^0) \in \hat{\Phi}.$$

Furthermore, it is the case that

$$\begin{aligned}
0 &= \int_{\underline{\theta}}^{\bar{\theta}} \left(\theta \eta_\alpha w_{x_p}^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w_{x_p}^1(\tilde{\theta}) d\tilde{\theta} - \underline{\theta} \eta_\alpha w_{x_p}^1(\underline{\theta}) + \underline{w}_{x_p}^0 - \beta^{se}(w_{x_p}^1(\theta)) \right) d\tilde{\Lambda}^{se} \\
&+ \int_{\underline{\theta}}^{\bar{\theta}} \left[\beta^{nw}(w_{x_p}^1(\theta)) - \left(\theta \eta_\alpha w_{x_p}^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w_{x_p}^1(\tilde{\theta}) d\tilde{\theta} - \underline{\theta} \eta_\alpha w_{x_p}^1(\underline{\theta}) + \underline{w}_{x_p}^0 \right) \right] d\tilde{\Lambda}^{nw}.
\end{aligned} \tag{18}$$

Proof. This is a direct application of Theorem 1 in Luenberger (1969, p. 217) [21]. The proof is relegated to Appendix 8.7. ■

My next result shows that no x_p -cutoff contract is optimal if Assumption 2 fails.

Proposition 5.

If Assumption 2 does not hold, then no x_p -cutoff contract is optimal for any $x_p \in \Theta$.

Proof. The proof proceeds by contradiction. Suppose that $(w_{x_p}^1, \underline{w}_{x_p}^0)$ is optimal for some $x_p \in \Theta$. According to Lemma 5, there exist nondecreasing $\tilde{\Lambda}^{se}, \tilde{\Lambda}^{nw}$ such that the Lagrangian (17) is maximized at $(w_{x_p}^1, \underline{w}_{x_p}^0)$ and (18) holds. This implies that $\tilde{\Lambda}^{nw}$ is constant so the integral related to $\tilde{\Lambda}^{nw}$ can be dropped. Without loss of generality I set $\tilde{\Lambda}^{se}(\bar{\theta}) = 1$. Integrating the Lagrangian by

parts yields

$$L(w^1, \underline{w}^0 \mid \tilde{\Lambda}^{\text{se}}) = (\underline{\theta} \eta_\alpha w^1(\underline{\theta}) - \underline{w}^0) \tilde{\Lambda}^{\text{se}}(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} (\theta \eta_\alpha w^1(\theta) - \beta^{\text{se}}(w^1(\theta))) d\tilde{\Lambda}^{\text{se}} \\ + \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} w^1(\theta) \left[\tilde{\Lambda}^{\text{se}}(\theta) - (1 - G(\theta)) \right] d\theta.$$

Then, I establish the necessary first-order conditions for $L(w^1, \underline{w}^0 \mid \tilde{\Lambda}^{\text{se}})$ to be maximized at x_p -cutoff rule and show that they cannot be satisfied if Assumption 2 fails. The rest of the proof is relegated to Appendix 8.8. ■

6 Extensions

6.1 More General Stochastic Processes

6.1.1 Inconclusive Successes

Now suppose that $\lambda^0 > 0$, so R generates successes in state 0 as well and players never fully learn the state. In this subsection, unless otherwise specified, I use the same notations as in the benchmark setting. Recall that $(\mathbf{w}^1(\pi), \mathbf{w}^0(\pi))$ denotes the *expected resource* allocated to R under policy π conditional on the state being 1 and 0, and Γ the image of the mapping $(\mathbf{w}^1, \mathbf{w}^0) : \Pi \rightarrow [0, 1]^2$. The following lemma characterizes the feasible set Γ .

Lemma 6 (Inconclusive news—feasible set).

There exist two functions $\beta^{\text{se}}, \beta^{\text{nw}} : [0, 1] \rightarrow [0, 1]$ such that $\Gamma = \{(w^1, w^0) \mid \beta^{\text{se}}(w^1) \leq w^0 \leq \beta^{\text{nw}}(w^1), w^1 \in [0, 1]\}$. The southeast boundary is given by $\beta^{\text{se}}(w^1) = 1 - (1 - w^1)^{\mu/(1+\mu)}$, for some constant $\mu > 0$. The northwest boundary β^{nw} is concave, nondecreasing, once continuously differentiable, having end points $(0, 0)$ and $(1, 1)$.

Proof. The proof is similar to that of Lemma 2 and relegated to Appendix 8.9 ■

Figure 8 depicts the feasible set when $r = 1/5, \lambda_1 = 2/5, \lambda_0 = (2 - \sqrt{2})/5$. Unlike the benchmark setting, the northwest boundary is characterized by a once continuously differentiable function. My next result shows that, if Assumption 2 holds, the cutoff rule as defined in Definition 1 is optimal. This is the case because the proof of Proposition 3, which only relies on the properties of the southeast boundary of the feasible set, applies directly to the current setting.

Proposition 6 (Inconclusive news—sufficiency).

If Assumption 2 holds, the cutoff rule is optimal.

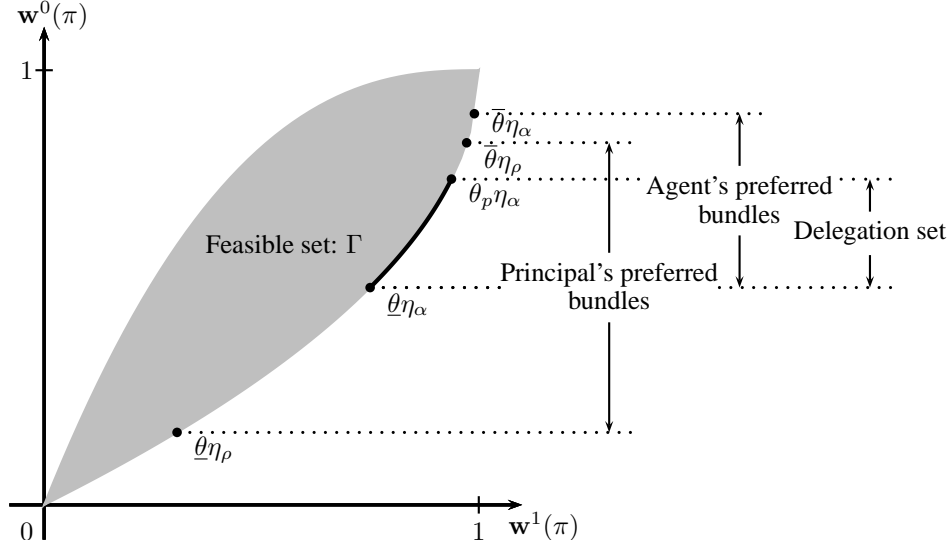


Figure 8: Delegation set under cutoff rule: inconclusive news case

Proof. According to Lemma 6, the southeast boundary of the feasible set is given by $\beta^{\text{se}}(w^1) = 1 - (1 - w^1)^{\mu/(1+\mu)}$, for some constant $\mu > 0$. The proof of Proposition 3 applies directly here. ■

The implementation of the cutoff rule, similar to that of the benchmark setting, is achieved by calibrating a constructed belief which starts with the prior belief $p(\theta_p \eta_\rho / \eta_\alpha)$. This belief is updated as follows: (i) if no success occurs, the belief drifts down according to the differential equation $\dot{p}_t = -(\lambda^1 - \lambda^0)p_t(1 - p_t)$; and (ii) if a success occurs at time t , the belief jumps from p_{t-} to

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

The principal imposes a cutoff at p_ρ^* . The agent can choose to experiment or not if this belief stays above the cutoff and are required to stop when it drops to the cutoff. The gray areas in Figure 9 show that lower types stop voluntarily as their posterior beliefs reach p_α^* . The black dot shows that those types with $\theta > \theta_p$ are required to stop.

Figure 9 also highlights the difference between the benchmark setting, where the belief jumps to one after the first success, and the inconclusive news setting, where the belief jumps up upon successes and then drifts down. Consequently, when successes are inconclusive, the optimum can no longer be implemented by imposing a fixed deadline. Instead, it takes the form of a sliding deadline. The principal initially extends some time to the agent. Then, whenever a success realizes, more time is extended. The agent is allowed to give up his time voluntarily. That is, the agent can choose to switch to S before he uses up the time granted by the principal. After a long enough period of time elapses without success, the principal requires the agent to switch to S .

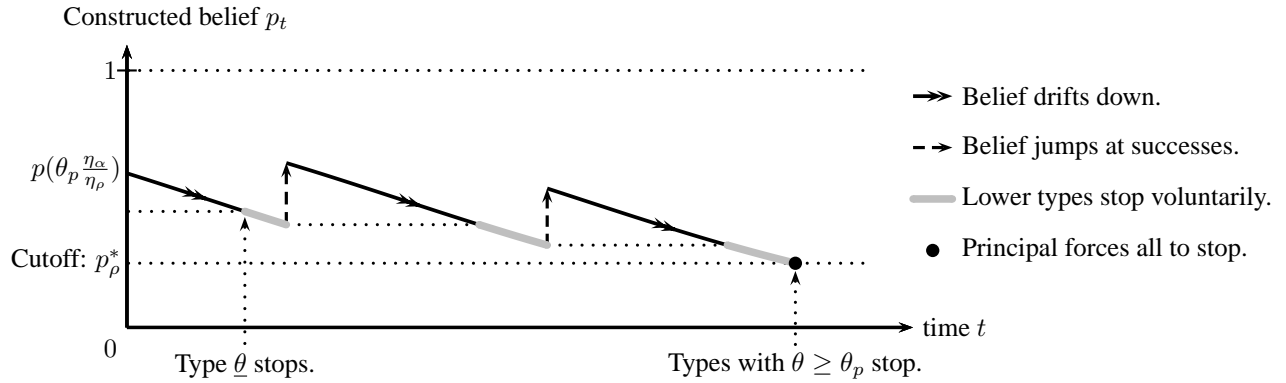


Figure 9: Implementing the cutoff rule: inconclusive news

My next result shows that the cutoff rule is time consistent when successes are inconclusive. The only difference from the benchmark setting is that the constructed belief does not jump to one upon the first success. I need to show that the principal finds it optimal not to adjust the cutoff each time a success occurs. The intuition, similar to that in the benchmark setting, is that there is no need to adjust the cutoff after the constructed belief is upgraded upon successes. The proof is also similar and hence relegated to Appendix 8.10.

Proposition 7 (Inconclusive news—time consistency).

If Assumption 2 holds, the cutoff rule is time-consistent.

Proof. See Appendix 8.10. ■

6.1.2 Lévy Processes and Lévy Bandits

Here, I extend the analysis to the more general Lévy bandits (Cohen and Solan, 2013 [11]). The risky task’s payoff is driven by a Lévy process whose Lévy triplet depends on an unknown binary state. In what follows, I start with a reminder about Lévy processes and Lévy bandits. Then, I show that the optimality of the cutoff rule and its time consistency property generalize to Lévy bandits.

Lévy processes. A Lévy process $L = (L(t))_{t \geq 0}$ is a continuous-time stochastic process that (i) starts at the origin: $L(0) = 0$; (ii) admits càdlàg modification;²⁰ (iii) has stationary independent increments. Examples of Lévy processes include a Brownian motion, a Poisson process, and a compound Poisson process.

²⁰It is continuous from the right and has limits from the left.

Let (Ω, P) be the underlying probability space. For every Borel measurable set $A \in \mathcal{B}(\mathbf{R} \setminus \{0\})$, and every $t \geq 0$, let the Poisson random measure $N(t, A)$ be the number of jumps of L in the time interval $[0, t]$ with jump size in A : $N(t, A) = \#\{0 \leq s \leq t \mid \Delta L(s) \equiv L(s) - L(s-) \in A\}$. The measure ν defined by

$$\nu(A) \equiv \mathbf{E}[N(1, A)] = \int N(1, A)(\omega) dP(\omega).$$

is called the *Lévy measure* of the process L .

I focus on Lévy processes that have finite expectation for each t . For a fixed Lévy process L , there exists a constant $\mu \in \mathbf{R}$, a Brownian motion $\sigma Z(t)$ with standard deviation $\sigma \geq 0$, and an independent Poisson random measure $N_\nu(t, dh)$ with the associated Lévy measure ν such that, for each $t \geq 0$, the Lévy-Itô decomposition of $L(t)$ is

$$L(t) = \mu t + \sigma Z(t) + \int_{\mathbf{R} \setminus \{0\}} h \tilde{N}_\nu(t, dh),$$

where $\tilde{N}_\nu(t, A) \equiv N_\nu(t, A) - t\nu(A)$ is the compensated Poisson random measure.²¹ Hence, a Lévy process L is characterized by a triplet $\langle \mu, \sigma, \nu \rangle$.

Lévy bandits. The agent operates a two-armed bandit in continuous time, with a safe arm S that yields a known flow payoff s_i to player i , and a risky arm R whose payoff, depending on an unknown state $x \in \{0, 1\}$, is given by the process L^x . For ease of exposition, I assume that both players derive the same payoff from R but different payoffs from S . For a fixed state x , L^x is a Lévy process characterized by the triplet $\langle \mu^x, \sigma^x, \nu^x \rangle$. For an arbitrary prior p that the state is 1, I denote by P_p the probability measure over space of realized paths.

I keep the same assumptions (A1–A6) on the Lévy processes L^x as in Cohen and Solan (2013) [11] and modify A5 to ensure that both players prefer to use R in state 1 and S in state 0. That is, $\mu^1 > s_i > \mu^0$, for $i \in \{\alpha, \rho\}$.²² Let $\eta_i = (\mu^1 - s_i)/(s_i - \mu^0)$ denote player i 's net gain from the experiment. I assume that the agent gains more from the experiment, *i.e.*, $\eta_\alpha > \eta_\rho$.²³

²¹Consider a set $A \in \mathcal{B}(\mathbf{R} \setminus \{0\})$ and a function $f : \mathbf{R} \rightarrow \mathbf{R}$. The integral with respect to a Poisson random measure $N(t, A)$ is defined as $\int_A f(h)N(t, dh) = \sum_{s \leq t} f(\Delta L(s)) \mathbf{1}_A(\Delta L(s))$.

²²The assumptions are (A1) $\mathbf{E}[(L^x)^2(1)] = (\mu^x)^2 + (\sigma^x)^2 + \int h^2 \nu^x(dh) < \infty$; (A2) $\sigma^1 = \sigma^0$; (A3) $|\nu^1(\mathbf{R} \setminus \{0\}) - \nu^0(\mathbf{R} \setminus \{0\})| < \infty$; (A4) $|\int h(\nu^1(dh) - \nu^0(dh))| < \infty$; (A5) $\mu^0 < s_\alpha < s_\rho < \mu^1$; (A6) For every $A \in \mathcal{B}(\mathbf{R} \setminus \{0\})$, $\nu^0(A) < \nu^1(A)$. Assumption (A1) states that both L^1 and L^0 have finite quadratic variation. It follows that both have finite expectation. Assumptions (A2) to (A4) ensure that players cannot distinguish between the two states in any infinitesimal time. Assumption (A5) states that the expected payoff rate of R is higher than that of S in state 1 and lower in state 0. The last assumption (A6) requires that jumps of any size h , both positive or negative, occur more often in state 1 than in state 0. Consequently, jumps always provide good news, and increase the posterior belief of state 1.

²³The results generalize to the case in which, for a fixed state x , the drift term of the Lévy process L^x differs for the

Policies and feasible set. A (pure) allocation policy is a non-anticipative stochastic process $\pi = \{\pi_t\}_{t \geq 0}$. Here, $\pi_t \in [0, 1]$ (resp. $1 - \pi_t$) may be interpreted as the fraction of time in the interval $[t, t + dt)$ that is devoted to R (resp. S), which may depend only on the history of events up to t .²⁴ The space of all policies, including randomized ones, is denoted Π . (See Footnote 8.)

Player i 's payoff given a policy $\pi \in \Pi$ and a prior belief $p \in [0, 1]$ that the state is 1 is

$$U_i(\pi, p) \equiv \mathbf{E} \left[\int_0^\infty r e^{-rt} \left[dL^x \left(\int_0^t \pi_s ds \right) + (1 - \pi_t) s_i dt \right] \mid \pi, p \right].$$

Over an interval $[t, t + dt)$, if the fraction π_t of time is allocated to R , the expected payoff increment to player i conditional on x is $[(1 - \pi_t)s_i + \pi_t \mu^x] dt$. By the Law of Iterated Expectations, I can write player i 's payoff as the discounted sum of the expected payoff increments

$$U_i(\pi, p) = \mathbf{E} \left[\int_0^\infty r e^{-rt} [\pi_t \mu^x + (1 - \pi_t) s_i] dt \mid \pi, p \right].$$

For a fixed policy π , I define $\mathbf{w}^1(\pi)$ and $\mathbf{w}^0(\pi)$ as follows:

$$\mathbf{w}^1(\pi) \equiv \mathbf{E} \left[\int_0^\infty r e^{-rt} \pi_t dt \mid \pi, 1 \right] \text{ and } \mathbf{w}^0(\pi) \equiv \mathbf{E} \left[\int_0^\infty r e^{-rt} \pi_t dt \mid \pi, 0 \right].$$

Then, player i 's payoff can be written as

$$U_i(\pi, p) = p (\mu^1 - s_i) \mathbf{w}^1(\pi) + (1 - p) (\mu^0 - s_i) \mathbf{w}^0(\pi) + s_i.$$

Let Γ denote the image of the mapping $(\mathbf{w}^1, \mathbf{w}^0) : \Pi \rightarrow [0, 1]^2$, referred to as the feasible set. The following lemma characterizes the southeast boundary of Γ .

Lemma 7. *There exists $a^* > 0$ such that the southeast boundary of Γ is given by*

$$\{(w^1, w^0) \mid w^0 = 1 - (1 - w^1)^{a^*/(1+a^*)}, w^1 \in [0, 1]\}$$

Proof. The proof is similar to that of Lemma 2 and relegated to Appendix 8.11 ■

Given Lemma 7, the proof of Proposition 3, which only relies on the properties of the southeast boundary of the feasible set, applies directly to the current setting. Therefore, the cutoff rule as defined in Definition 1 is optimal under Assumption 2.

principal and the agent, as long as the relation $\eta_\alpha > \eta_\rho$ holds.

²⁴Suppose the process L is a Lévy process L^1 with probability $p \in (0, 1)$ and L^0 with probability $1 - p$. Let \mathcal{F}_s^L be the sigma-algebra generated by the process $(L(t))_{t \leq s}$. Then it is required that the process π satisfies that $\{\int_0^t \pi_s ds \leq t'\} \in \mathcal{F}_{t'}^L$, for any $t, t' \in [0, \infty)$.

Proposition 8 (Lévy bandits—sufficiency).

The cutoff rule is optimal if Assumption 2 holds.

For every prior $p \in [0, 1]$ that the state is 1, the probability measure P_p satisfies $P_p = pP_1 + (1 - p)P_0$. An important auxiliary process is the Radon-Nikodym density, given by

$$\psi_t \equiv \frac{d(P_0 | \mathcal{F}_{K(t)})}{d(P_1 | \mathcal{F}_{K(t)})}, \text{ where } K(t) = \int_0^t \pi_s ds \text{ and } t \in [0, \infty).$$

According to Lemma 1 in Cohen and Solan (2013) [11], if the prior belief is p , the posterior belief at time t is given by

$$p_t = \frac{p}{p + (1 - p)\psi_t}.$$

The agent of type θ updates his belief about the state. He assigns odds ratio θ/ψ_t to the state being 1, referred to as his type at time t . Let $\underline{\theta}_t = \max\{\underline{\theta}, \theta_\alpha^* \psi_t\}$. Recall that θ_α^* denotes the odds ratio at which the agent is indifferent between continuing and stopping. At time t , only those types above $\underline{\theta}_t$ remain. The principal's updated belief about the agent's type distribution, *in terms of his type at time 0*, is given by the density function

$$f(\theta | t) = \begin{cases} \frac{[p(\theta) + (1 - p(\theta))\psi_t]f(\theta)}{\int_{\underline{\theta}_t}^{\bar{\theta}} [p(\theta) + (1 - p(\theta))\psi_t]f(\theta)d\theta} & \text{if } \theta \in [\underline{\theta}_t, \bar{\theta}], \\ 0 & \text{otherwise.} \end{cases}$$

The principal's belief about the agent's type distribution, *in terms of his type at time t* , is given by the density function

$$f_t(\theta) = \begin{cases} f(\theta\psi_t | t)\psi_t & \text{if } \theta \in [\underline{\theta}_t/\psi_t, \bar{\theta}/\psi_t], \\ 0 & \text{otherwise.} \end{cases}$$

I prove that continuing the cutoff rule is optimal by showing two things. First, given the distribution f_t at time t , the threshold of the top pooling segment is θ_p/ψ_t . Second, if Assumption 2 holds for $\theta \leq \theta_p$ under distribution f , then it holds for $\theta \leq \theta_p/\psi_t$ under f_t . The detailed proof is similar to that of Proposition 7 (see Appendix 8.5) and hence omitted.

Proposition 9 (Lévy bandits—time consistency).

If Assumption 2 holds, the cutoff rule is time-consistent.

6.2 Biased Toward the Safe Task: a Lockup Period

In this subsection, I consider the situations in which the agent is biased toward the safe task, *i.e.*, $\eta_\alpha < \eta_\rho$. This happens, for example, when a division (agent) conducts an experiment which yields

positive externalities to other divisions. Hence, the agent does not internalize the total benefit that the experiment brings to the organization (principal). Another possibility is that the agent does not perceive the risky task to generate significant career opportunities compared with alternative activities. In both cases, the agent has a preference to stop experimenting earlier.

To illustrate the main intuition, I assume that the lowest type agent prefers a positive length of experimentation, *i.e.*, $\tau_\alpha(\underline{\theta}) > 0$. Using the same methods as in the main model, I first show that types below some threshold are pooled. Intuitively, types at the bottom prefer to experiment less than what the principal prefers to do for any prior. The cost of separating those types exceeds the benefit. Then, I show that under certain condition the optimal outcome for the principal can be implemented by starting with a properly calibrated prior belief that the state is 1. This belief is then updated as if the agent had no private information. As long as this belief remains above a cutoff belief, the agent is required to operate R . As soon as it drops to the cutoff, the principal keeps her hands off the project and lets the agent decide whether to experiment or not.

Notably, in contrast to the main model, the agent has no flexibility until the cutoff belief is reached. I call this mechanism the *reversed cutoff rule*. Those types with low enough priors stop experimenting as soon as the cutoff is reached. Those with higher priors are not constrained and thus implement their preferred policies.

If successes are conclusive, the principal simply sets up a *lockup period* during which the agent uses R regardless of the outcome. After the lockup period ends, the agent is free to experiment or not. If successes are inconclusive, the principal initially sets up a lockup length. Each time a success occurs, the lockup length is extended. The agent has no freedom until the lockup period ends.

Given a direct mechanism $(w^1(\theta), w^0(\theta))$, let $U_\alpha(\theta)$ denote the payoff that the agent of type θ gets by maximizing over his report. As the optimal mechanism is truthful, $U_\alpha(\theta)$ equals $\theta\eta_\alpha w^1(\theta) - w^0(\theta)$ and the envelope condition implies that $U'_\alpha(\theta) = \eta_\alpha w^1(\theta)$. By integrating the envelope condition, one obtains the standard integral condition

$$\theta\eta_\alpha w^1(\theta) - w^0(\theta) = \bar{\theta}\eta_\alpha w^1(\bar{\theta}) - \bar{w}^0 - \int_{\theta}^{\bar{\theta}} \eta_\alpha w^1(\tilde{\theta}) d\tilde{\theta},$$

where \bar{w}^0 stands for $w^0(\bar{\theta})$. Substituting $w^0(\theta)$ and simplifying, I reduce the problem to finding a function $w^1 : \Theta \rightarrow [0, 1]$ and a scalar \bar{w}^0 that solves

$$\max_{w^1, \bar{w}^0 \in \Phi^s} \left(\bar{\theta}\eta_\alpha w^1(\bar{\theta}) - \bar{w}^0 - \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} w^1(\theta) G^s(\theta) d\theta \right), \quad (\text{OBJ-S})$$

subject to

$$\theta\eta_\alpha w^1(\theta) + \int_\theta^{\bar{\theta}} \eta_\alpha w^1(\tilde{\theta})d\tilde{\theta} - \bar{\theta}\eta_\alpha w^1(\bar{\theta}) + \bar{w}^0 - \beta^{\text{se}}(w^1(\theta)) \geq 0, \forall \theta \in \Theta, \quad (19)$$

$$\beta^{\text{nw}}(w^1(\theta)) - \left(\theta\eta_\alpha w^1(\theta) + \int_\theta^{\bar{\theta}} \eta_\alpha w^1(\tilde{\theta})d\tilde{\theta} - \bar{\theta}\eta_\alpha w^1(\bar{\theta}) + \bar{w}^0 \right) \geq 0, \forall \theta \in \Theta, \quad (20)$$

where

$$\Phi^s \equiv \{w^1, \bar{w}^0 \mid w^1 : \Theta \rightarrow [0, 1], w^1 \text{ nondecreasing}; \bar{w}^0 \in [0, 1]\},$$

$$G^s(\theta) = \frac{H(\theta)}{H(\bar{\theta})} + \left(1 - \frac{\eta_\rho}{\eta_\alpha}\right) \theta \frac{h(\theta)}{H(\bar{\theta})}, \quad \text{where } h(\theta) = \frac{f(\theta)}{1+\theta} \text{ and } H(\theta) = \int_{\underline{\theta}}^\theta h(\tilde{\theta})d\tilde{\theta}.$$

I denote this problem by \mathcal{P}^s . Let θ_p^s be the highest value in Θ such that

$$\int_{\underline{\theta}}^{\hat{\theta}} G^s(\theta)d\theta \leq 0, \text{ for any } \hat{\theta} \leq \theta_p^s. \quad (21)$$

My next result shows that types with $\theta \leq \theta_p^s$ are pooled.

Proposition 10 (Pooling on bottom).

An optimal contract (w^{1}, \bar{w}^{0*}) satisfies $w^{1*}(\theta) = w^{1*}(\theta_p^s)$ for $\theta \leq \theta_p^s$. It is optimal for (19) or (20) to hold with equality at θ_p^s .*

Proof. The contribution to (OBJ-S) from types with $\theta < \theta_p^s$ is $-\eta_\alpha \int_{\underline{\theta}}^{\theta_p^s} w^1(\theta)G^s(\theta)d\theta$. Substituting $w^1(\theta) = w^1(\theta_p^s) - \int_{\theta}^{\theta_p^s} dw^1$ and integrating by parts, I obtain

$$-\eta_\alpha \int_{\underline{\theta}}^{\theta_p^s} w^1(\theta)G^s(\theta)d\theta = -\eta_\alpha w^1(\theta_p^s) \int_{\underline{\theta}}^{\theta_p^s} G^s(\theta)d\theta + \eta_\alpha \int_{\underline{\theta}}^{\theta_p^s} \int_{\underline{\theta}}^\theta G^s(\tilde{\theta})d\tilde{\theta}dw^1(\theta). \quad (22)$$

The first term only depends on $w^1(\theta_p^s)$. The second term depends on $dw^1(\theta)$ for all $\theta \in [\underline{\theta}, \theta_p^s]$. According to the definition of θ_p^s , the integrand of the second term, $\int_{\underline{\theta}}^\theta G^s(\tilde{\theta})d\tilde{\theta}$, is weakly negative for all $\theta \in [\underline{\theta}, \theta_p^s]$. Therefore, it is optimal to set $dw^1(\theta) = 0$ for all $\theta \in [\underline{\theta}, \theta_p^s]$. If $\theta_p^s = \bar{\theta}$, all types are pooled. The principal offers her preferred uninformed bundle, which is on the southeast boundary of Γ . If $\theta_p^s < \bar{\theta}$, the first term of (22) is zero as well because $\int_{\underline{\theta}}^{\theta_p^s} G^s(\theta)d\theta = 0$. Adjusting $w^1(\theta_p^s)$ does not affect the objective function, so $w^1(\theta_p^s)$ can be decreased until either (19) or (20) binds. ■

For the rest of this subsection, I focus on the more interesting case in which $\theta_p^s < \bar{\theta}$. I first define the reversed cutoff rule.

Definition 2. The *reversed cutoff rule* is the contract (w^1, w^0) such that

$$(w^1(\theta), w^0(\theta)) = \begin{cases} (w_\alpha^1(\theta_p^s), w_\alpha^0(\theta_p^s)) & \text{if } \theta \leq \theta_p^s, \\ (w_\alpha^1(\theta), w_\alpha^0(\theta)) & \text{if } \theta > \theta_p^s. \end{cases}$$

I denote this rule by $(w_{\theta_p^s}^1, \bar{w}_{\theta_p^s}^0)$. My next result gives a sufficient condition under which the reserved cutoff rule is optimal.

Proposition 11 (Sufficiency-reversed cutoff rule).

The reversed cutoff rule $(w_{\theta_p^s}^1, \bar{w}_{\theta_p^s}^0)$ is optimal if $G^s(\theta)$ is nondecreasing when $\theta \geq \theta_p^s$.

If $f(\theta)$ is differentiable, $G^s(\theta)$ being nondecreasing for $\theta \geq \theta_p^s$ is equivalent to requiring that

$$\frac{\theta f'(\theta)}{f(\theta)} \leq \frac{\eta_\alpha}{\eta_\rho - \eta_\alpha} - \frac{1}{1 + \theta}, \quad \forall \theta \in [\theta_p^s, \bar{\theta}].$$

It is satisfied for any density f with $\theta f'/f$ bounded from above when η_ρ/η_α is sufficiently close to 1, or equivalently when two players' preferences are sufficiently aligned.

Proof. I define a new problem $\hat{\mathcal{P}}^s$ which differs from \mathcal{P}^s in two aspects: (i) the upper bound constraint (20) is dropped; and (ii) the lower bound constraint (19) is replaced with the following:

$$\theta \eta_\alpha w^1(\theta) + \int_\theta^{\bar{\theta}} \eta_\alpha w^1(\tilde{\theta}) d\tilde{\theta} - \bar{\theta} \eta_\alpha w^1(\bar{\theta}) + \bar{w}^0 - \hat{\beta}(w^1(\theta)) \geq 0, \quad \forall \theta \in \Theta.$$

Define the Lagrangian functional associated with $\hat{\mathcal{P}}^s$ as

$$\begin{aligned} \hat{L}^s(w^1, \bar{w}^0 \mid \Lambda) &= \bar{\theta} \eta_\alpha w^1(\bar{\theta}) - \bar{w}^0 - \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} w^1(\theta) G^s(\theta) d\theta \\ &+ \int_{\underline{\theta}}^{\bar{\theta}} \left(\theta \eta_\alpha w^1(\theta) + \int_\theta^{\bar{\theta}} \eta_\alpha w^1(\tilde{\theta}) d\tilde{\theta} - \bar{\theta} \eta_\alpha w^1(\bar{\theta}) + \bar{w}^0 - \hat{\beta}(w^1(\theta)) \right) d\Lambda. \end{aligned}$$

Integrating by parts and simplifying, I obtain

$$\begin{aligned} \hat{L}^s(w^1, \bar{w}^0 \mid \Lambda) &= (\bar{\theta} \eta_\alpha w^1(\bar{\theta}) - \bar{w}^0) (1 - \Lambda(\bar{\theta}) + \Lambda(\underline{\theta})) + \int_{\underline{\theta}}^{\bar{\theta}} \left(\theta \eta_\alpha w^1(\theta) - \hat{\beta}(w^1(\theta)) \right) d\Lambda \\ &+ \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} (\Lambda(\theta) - \Lambda(\underline{\theta}) - G^s(\theta)) w^1(\theta) d\theta \end{aligned}$$

Based on Lemma 3 and Lemma 4, it suffices to show that the following first-order conditions hold

for some candidate multiplier Λ ,

$$\begin{aligned}\hat{L}^s(w_{\theta_p^s}^1, \bar{w}_{\theta_p^s}^0; w^1, \bar{w}^0 \mid \Lambda) &\leq 0, \quad \forall (w^1, \bar{w}^0) \in \Phi^s \\ \hat{L}^s(w_{\theta_p^s}^1, \bar{w}_{\theta_p^s}^0; w_{\theta_p^s}^1, \bar{w}_{\theta_p^s}^0 \mid \Lambda) &= 0.\end{aligned}$$

If $G^s(\theta)$ is nondecreasing when $\theta \in [\theta_p, \bar{\theta}]$, the first-order conditions are satisfied given the following candidate multiplier

$$\Lambda(\theta) = \begin{cases} 0 & \text{if } \theta \in [\theta, \theta_p^s), \\ G^s(\theta) & \text{if } \theta \in [\theta_p^s, \bar{\theta}), \\ 1 & \text{if } \theta = \bar{\theta}.\end{cases}$$

The jump at θ_p^s is nonnegative according to the definition of θ_p^s . The jump at $\bar{\theta}$ is nonnegative because $G^s(\theta) \leq 1$ for all θ . This completes the proof. \blacksquare

6.3 Heterogeneous Beliefs

The analysis can be extended to the situation in which beliefs are heterogeneous. Suppose that the two players differ in their prior beliefs about the state. At time 0, the agent obtains a private informative signal about the state. Due to difference in their prior beliefs, the principal's belief about the state distribution would differ from that of the agent even if she observed the agent's signal. To illustrate, suppose that there is no fundamental preference conflict. The results in the benchmark setting are applicable in the situation where the agent is more optimistic and assigns a higher odds ratio to state 1 than the principal for any fixed signal. The agent has an incentive to misrepresent his information to counteract the principal's pessimism (from the agent's point of view). There are many other possible forms of different opinions. For example, the agent might be more pessimistic or disagree with the principal about the informativeness of his private signal. The analysis can be easily modified to incorporate those situations.

6.4 Social Planner's Problem

Here, I consider a social planner who seeks to maximize the weighted sum of the two players' payoffs. The social planner determines a delegation set at time 0, knowing that the agent chooses a bundle to maximize his own payoff. The social planner's gain from R 's successes relative to S 's flow payoffs can be summarized by a constant, denoted η_s , which is a weighted average of η_ρ and η_α . As a result, the social planner's problem is similar to the principal's problem, the only difference being that the bias term is smaller since the agent's welfare is taken into account. In general, the social planner prefers to give the agent more flexibility than the principal does.

If assumption 2 holds for $\theta \in \Theta$, the optimal solution to the social planner's problem is also a cutoff rule. Moreover, based on Corollary 2, the threshold of the pooling segment θ_p is an increasing function of the weight put on the agent's welfare. Therefore, the more weight is put on the agent's welfare, the more flexibility shall be granted to the agent.

7 Concluding Remarks

This paper discusses how organizations can optimally manage innovative activities, particularly how much control right over resource allocation shall be left to the agent over time with the presence of misaligned preferences and hidden information. From this aspect, this paper contributes to the discussion of how to optimally allocate formal authority and real authority within organizations (Aghion and Tirole, 1997 [1]).

The optimal delegation rule requires the agent to achieve a success before the next deadline to keep the project alive. It is simple, time-consistent and already implemented in organizations such as Google. Google encourages its employees to come up with new ideas and build a prototype. If the initial result is satisfactory, Google makes it an official project, funds it and sets the next pair of goal and deadline. The goal must be met before the deadline to secure future funding. Successful products including Gmail, Google Earth and Google Maps survived all the deadlines. Needless to say, many more did not. For example, the once highly publicized and well-funded Google Wave was canceled in August 2010 as it failed to achieve the goal set by Google executives before then.

Besides in-house innovation, my results also apply to the government sector, which often experiments reforms in public policy. The constituents delegate reforms to politicians. Legislatures delegate policy-making to their standing committees. Throughout the process transfers are prohibited to prevent corruption. It has been concluded that every reform has consequences that cannot be fully known until it has been implemented (Strulovici, 2010 [24]). The constituents as well as the government learn the effects of a reform that gradually unfold. If a reform is thought to be a failure, it can be legislatively repealed, executively overturned or allowed to automatically expire with a sunset provision. Politicians hope to prolong the policy experimentation as they gain the most popularity from successful reforms that they initiated. My sliding deadline rule suggests that if politicians are better informed on policies, every reform should carry a sunset provision. They should be renewed only upon demonstrated successes.

8 Appendix

8.1 Proof of Lemma 2

I have shown that the supremum $K(\gamma)$ is achieved by stopping-time policies if $\gamma^1 > 0, \gamma^0 < 0$. If $p \leq p_i^*$ (or equivalently, $-\gamma^1/\gamma^0 \leq r/(r + \lambda^1)$), the stopping time is 0. Hence, $(w^1, w^0) = (0, 0)$ and $K(\gamma) = 0$. If $p > p_i^*$, the stopping time τ satisfies

$$e^{-\lambda^1 \tau} = \frac{p_i^*}{1 - p_i^*} \frac{1 - p}{p} = \frac{r}{\lambda^1 + r} \frac{-\gamma^0}{\gamma^1}.$$

The corresponding (w^1, w^0) pair is

$$(w^1, w^0) = \left(1 - \left(\frac{r}{\lambda^1 + r} \frac{-\gamma^0}{\gamma^1} \right)^{\frac{r + \lambda^1}{\lambda^1}}, 1 - \left(\frac{r}{\lambda^1 + r} \frac{-\gamma^0}{\gamma^1} \right)^{\frac{r}{\lambda^1}} \right).$$

The derivative of the boundary Γ^{st} at this point equals $-\gamma^1/\gamma^0$. As the ratio $-\gamma^1/\gamma^0$ increases from $r/(r + \lambda^1)$ to ∞ the stopping time ranges from 0 to ∞ . This proves that the southeast boundary of \mathcal{H} is Γ^{st} .

I have shown that if $\gamma^1 < 0, \gamma^0 > 0$ the supremum $K(\gamma)$ is achieved by a cutoff Markov policy. According to Keller and Rady (2013) [19], the cutoff, denoted p_i^{**} , is given by $p_i^{**}/(1 - p_i^{**}) = (\lambda^1 + r)/(r\eta_i)$. If $p \geq p_i^{**}$ (or equivalently, $-\gamma^1/\gamma^0 \geq (\lambda^1 + r)/r$), the supremum is achieved by playing S forever. Hence, $(w^1, w^0) = (0, 0)$. If $-\gamma^1/\gamma^0 < (\lambda^1 + r)/r$, the supremum is achieved by using R until the first success and then switching to S . Hence, $(w^1, w^0) = (r/(r + \lambda^1), 1)$. Therefore, the northwest boundary of \mathcal{H} is $\Gamma^{\text{sl}} \cup \{(w^1, w^0) \mid \epsilon(r/(r + \lambda^1), 1) + (1 - \epsilon)(0, 0), \epsilon \in [0, 1]\}$.

8.2 Proof of Proposition 1

Let α_l (resp. α_h) denote the low (resp. high) type agent and ρ_l (resp. ρ_h) the low (resp. high) type principal.

1. Suppose that $\theta_l \eta_\rho > r/(\lambda^1 + r)$. Both ρ_l 's and ρ_h 's preferred bundles lie in the interior of Γ^{st} . Given that $\theta_h > \theta_l$ and $\eta_\alpha > \eta_\rho$, the slopes of players' indifference curves are ranked as follows

$$\theta_h \eta_\alpha > \max\{\theta_h \eta_\rho, \theta_l \eta_\alpha\} \geq \min\{\theta_h \eta_\rho, \theta_l \eta_\alpha\} > \theta_l \eta_\rho.$$

Let ICL and ICH denote α_l 's and α_h 's IC constraints. Let I_{α_l} denote α_l 's indifference curves.

If α_l prefers $(w_\rho^1(\theta_l), w_\rho^0(\theta_l))$ to $(w_\rho^1(\theta_h), w_\rho^0(\theta_h))$, the optimum is

$$\{(w_\rho^1(\theta_l), w_\rho^0(\theta_l)), (w_\rho^1(\theta_h), w_\rho^0(\theta_h))\}.$$

This is the case when the slope of the line connecting $(w_\rho^1(\theta_l), w_\rho^0(\theta_l))$ and $(w_\rho^1(\theta_h), w_\rho^0(\theta_h))$ is greater than $\theta_l \eta_\alpha$. This condition is satisfied when η_α / η_ρ is bounded from above by

$$b' \equiv \frac{\theta_h(\lambda^1 + r) \left(\theta_h^{\frac{r}{\lambda^1}} - \theta_l^{\frac{r}{\lambda^1}} \right)}{r \left(\theta_h^{\frac{r+\lambda^1}{\lambda^1}} - \theta_l^{\frac{r+\lambda^1}{\lambda^1}} \right)}.$$

If this condition does not hold, at least one IC constraint binds. I explain how to find the optimal bundles.

Step 1. ICL binds. Suppose not. It must be the case that ICH binds. Given that ICL does not bind and ICH binds, the principal offers two distinct bundles $(w^1(\theta_l), w^0(\theta_l)) < (w^1(\theta_h), w^0(\theta_h))$ which lie on the same indifference curve of α_h . Given that $\theta_h \eta_\alpha > \max\{\theta_h \eta_\rho, \theta_l \eta_\rho\}$, both ρ_h and ρ_l strictly prefer $(w^1(\theta_l), w^0(\theta_l))$ to $(w^1(\theta_h), w^0(\theta_h))$. The principal is strictly better off by offering a pooling bundle $(w^1(\theta_l), w^0(\theta_l))$. Contradiction. Hence, ICL holds with equality.

Step 2. If $\theta_h \eta_\rho < \theta_l \eta_\alpha$, the optimum is pooling. Suppose not. Suppose that the principal offers two distinct bundles $(w^1(\theta_l), w^0(\theta_l)) < (w^1(\theta_h), w^0(\theta_h))$ which are on the same indifference curve of α_l . Given that $\theta_l \eta_\rho < \theta_h \eta_\rho < \theta_l \eta_\alpha$, α_l 's indifference curves are steeper than ρ_h 's and ρ_l 's. Both ρ_h and ρ_l strictly prefer $(w^1(\theta_l), w^0(\theta_l))$ to $(w^1(\theta_h), w^0(\theta_h))$. The principal is strictly better off by offering a pooling bundle $(w^1(\theta_l), w^0(\theta_l))$. Contradiction.

Step 3. If $\theta_h \eta_\rho > \theta_l \eta_\alpha$, the optimal bundles are on the boundary of Γ . Suppose not. Suppose that $(w^1(\theta_l), w^0(\theta_l))$ or $(w^1(\theta_h), w^0(\theta_h))$ is in the interior. The indifference curve of α_l going through $(w^1(\theta_l), w^0(\theta_l))$ intersects the boundary at $(\tilde{w}^1(\theta_l), \tilde{w}^0(\theta_l))$ and $(\tilde{w}^1(\theta_h), \tilde{w}^0(\theta_h))$. Given that $\theta_h \eta_\rho > \theta_l \eta_\alpha > \theta_l \eta_\rho$, ρ_h prefers $(\tilde{w}^1(\theta_h), \tilde{w}^0(\theta_h))$ to $(w^1(\theta_h), w^0(\theta_h))$ and ρ_l prefers $(\tilde{w}^1(\theta_l), \tilde{w}^0(\theta_l))$ to $(w^1(\theta_l), w^0(\theta_l))$. The principal is strictly better off by offering $(\tilde{w}^1(\theta_l), \tilde{w}^0(\theta_l))$ and $(\tilde{w}^1(\theta_h), \tilde{w}^0(\theta_h))$ (see Figure 10). Therefore, the optimal bundles are on the boundary. The problem is reduced to locate the low type agent's indifference curve on which $(w^{1*}(\theta_l), w^{0*}(\theta_l))$ and $(w^{1*}(\theta_h), w^{0*}(\theta_h))$ lie. I want to show that this indifference curve must be between the indifference curves of α_l which go through $(w_\rho^1(\theta_l), w_\rho^0(\theta_l))$ and $(w_\rho^1(\theta_h), w_\rho^0(\theta_h))$ (see Figure 11). Suppose not. Suppose the optimal bundles are pinned down by the dashed indifference curve of

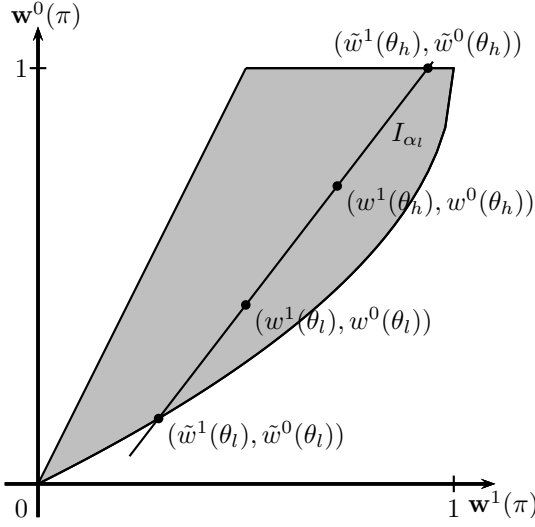


Figure 10: Optimum bundles are on the boundary

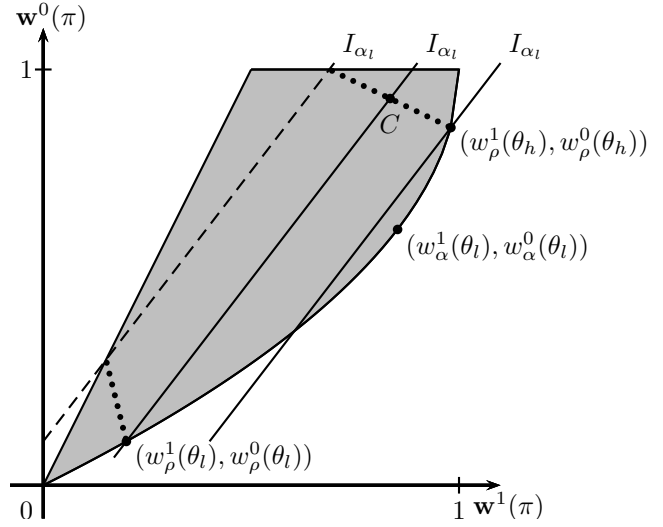


Figure 11: Location of optimal bundles

α_l as shown in Figure 11. The principal is strictly better off by offering $(w_\rho^1(\theta_l), w_\rho^0(\theta_l))$ and C . Analogously, the optimal bundles cannot lie on the indifference curve which lies to the southeast of the indifference curve that goes through $(w_\rho^1(\theta_h), w_\rho^0(\theta_h))$.

Step 4. *If $\theta_h \eta_\rho = \theta_l \eta_\alpha$, there exists an optimal pooling bundle.* If the principal finds it optimal to offer two distinct contracts $\{(w^1(\theta_l), w^0(\theta_l)), (w^1(\theta_h), w^0(\theta_h))\}$, it must be the case that $(w^1(\theta_l), w^0(\theta_l))$ and $(w^1(\theta_h), w^0(\theta_h))$ are on the same indifference curve of α_l . Since α_l 's indifference curve is steeper than ρ_l 's, $(w^1(\theta_l), w^0(\theta_l))$ lies on the boundary of Γ and $(w^1(\theta_h), w^0(\theta_h))$ is located to the northeast of $(w^1(\theta_l), w^0(\theta_l))$. Since ρ_h has the same indifference curves as α_l , it is optimal for the principal to offer a pooling contract $(w^1(\theta_l), w^0(\theta_l))$ if $\{(w^1(\theta_l), w^0(\theta_l)), (w^1(\theta_h), w^0(\theta_h))\}$ is optimal.

Combining Step 2 and 4, I obtain that pooling is optimal when $\eta_\alpha/\eta_\rho \geq \theta_h/\theta_l$. This completes the proof of Proposition 1.

2. Suppose that $\theta_h \eta_\rho \leq r/(r + \lambda^1)$. Both ρ_h and ρ_l prefer to stop at time 0. The optimum is $(w^{1*}(\theta_l), w^{0*}(\theta_l)) = (w^{1*}(\theta_h), w^{0*}(\theta_h)) = (0, 0)$.
3. Suppose that $\theta_l \eta_\rho \leq r/(r + \lambda^1) < \theta_h \eta_\rho$, the principal optimally offers her preferred bundles if

$$\theta_l \eta_\alpha \leq \frac{1 - \left(\frac{\eta_\rho \theta_h (\lambda^1 + r)}{r} \right)^{-\frac{r}{\lambda^1}}}{1 - \left(\frac{\eta_\rho \theta_h (\lambda^1 + r)}{r} \right)^{-\frac{\lambda^1 + r}{\lambda^1}}}.$$

If this does not hold and $\theta_l \eta_\alpha < \theta_h \eta_\rho$, the principal offers two bundles on the same indifference

curve of the low type agent. Both bundles are on the boundary of Γ . If $\theta_l \eta_\alpha \geq \theta_h \eta_\rho$, pooling is optimal. The proof is similar as in case 1.

8.3 Proof of Corollary 1

Substituting $G(\theta) = \frac{1}{H(\bar{\theta})} \left(\int_{\hat{\theta}}^{\bar{\theta}} h(\tilde{\theta}) d\tilde{\theta} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1 \right) \theta h(\theta) \right)$ and integrating by parts, I obtain

$$\begin{aligned} \int_{\hat{\theta}}^{\bar{\theta}} G(\theta) d\theta &= \frac{1}{H(\bar{\theta})} \int_{\hat{\theta}}^{\bar{\theta}} \left[\int_{\theta}^{\bar{\theta}} h(\tilde{\theta}) d\tilde{\theta} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1 \right) \theta h(\theta) \right] d\theta \\ &= \frac{1}{H(\bar{\theta})} \int_{\hat{\theta}}^{\bar{\theta}} \left[(\theta - \hat{\theta}) h(\theta) + \left(\frac{\eta_\rho}{\eta_\alpha} - 1 \right) \theta h(\theta) \right] d\theta \\ &= \frac{1}{H(\bar{\theta})} \int_{\hat{\theta}}^{\bar{\theta}} \left(\frac{\eta_\rho}{\eta_\alpha} \theta - \hat{\theta} \right) h(\theta) d\theta. \end{aligned}$$

For any $\hat{\theta} \in [\bar{\theta} \eta_\rho / \eta_\alpha, \bar{\theta}]$, the integrand $(\theta \eta_\rho / \eta_\alpha - \hat{\theta}) h(\theta)$ is weakly negative for any $\theta \in \Theta$. Therefore, $\theta_p \leq \bar{\theta} \eta_\rho / \eta_\alpha$. If $\theta_p > \underline{\theta}$, θ_p is strictly below $\bar{\theta} \eta_\rho / \eta_\alpha$. Otherwise, $h(\theta)$ must equal 0 for $\theta \in [\bar{\theta} \eta_\rho / \eta_\alpha, \bar{\theta}]$. Contradiction.²⁵

8.4 Proof of Corollary 2

Let $G(\theta, z)$ be the value of $G(\theta)$ if η_ρ / η_α equals $z \in [0, 1]$, i.e.,

$$G(\theta, z) = \frac{H(\bar{\theta}) - H(\theta)}{H(\bar{\theta})} + (z - 1) \theta \frac{h(\theta)}{H(\bar{\theta})}.$$

Let $\theta_p(z)$ be the lowest value in Θ such that $\int_{\hat{\theta}}^{\bar{\theta}} G(\theta, z) d\theta \leq 0$ for any $\hat{\theta} \geq \theta_p(z)$. Because $G(\theta, z)$ is an increasing function of z for a fixed $\theta \in \Theta$, $\theta_p(z)$ also increases in z .

If $z = 1$, $G(\theta, 1) = (H(\bar{\theta}) - H(\theta)) / H(\bar{\theta})$. Suppose $\theta_p(1) < \bar{\theta}$. Given the definition of $\theta_p(1)$, I have

$$\begin{aligned} \int_{\theta_p(1)}^{\bar{\theta}} G(\theta, 1) d\theta &= \frac{1}{H(\bar{\theta})} \int_{\theta_p(1)}^{\bar{\theta}} (H(\bar{\theta}) - H(\theta)) d\theta \\ &= \frac{1}{H(\bar{\theta})} \int_{\theta_p(1)}^{\bar{\theta}} (\theta - \theta_p(1)) h(\theta) d\theta \leq 0. \end{aligned}$$

This implies that $h(\theta) = 0$ for all $\theta \in [\theta_p(1), \bar{\theta}]$. Contradiction. Therefore, $\theta_p(1) = \bar{\theta}$.

²⁵If $\theta_p > \underline{\theta}$, I obtain that $\int_{\theta_p}^{\bar{\theta}} (\eta_\rho \theta - \eta_\alpha \theta_p) h(\theta) d\theta = 0$.

For any $\hat{\theta} \in \Theta$ and $z \in [0, 1]$, I have

$$\begin{aligned} \int_{\hat{\theta}}^{\bar{\theta}} G(\theta, z) d\theta &= \frac{1}{H(\bar{\theta})} \left(\int_{\hat{\theta}}^{\bar{\theta}} (z-1)\theta h(\theta) d\theta + \int_{\hat{\theta}}^{\bar{\theta}} (\theta - \hat{\theta}) h(\theta) d\theta \right) \\ &= \frac{1}{H(\bar{\theta})} \left(z \int_{\hat{\theta}}^{\bar{\theta}} \theta h(\theta) d\theta - \hat{\theta} \int_{\hat{\theta}}^{\bar{\theta}} h(\theta) d\theta \right). \end{aligned}$$

Let z^* be $\min_{\hat{\theta} \in \Theta} \left(\hat{\theta} \int_{\hat{\theta}}^{\bar{\theta}} h(\theta) d\theta \right) / \left(\int_{\hat{\theta}}^{\bar{\theta}} \theta h(\theta) d\theta \right)$. If $z \leq z^*$, $\int_{\hat{\theta}}^{\bar{\theta}} G(\theta, z) d\theta \leq 0$ for any $\hat{\theta} \in \Theta$, and thus $\theta_p(z) = \underline{\theta}$.

8.5 Proof of Proposition 4

Given the distribution f_δ over $\theta \in [\underline{\theta}_\delta e^{-\lambda^1 \delta}, \bar{\theta} e^{-\lambda^1 \delta}]$, I define functions $h_\delta, H_\delta, G_\delta$ as follows:

$$\begin{aligned} h_\delta(\theta) &= \frac{f_\delta(\theta)}{1+\theta}, \quad \text{and} \quad H_\delta(\theta) = \int_{\underline{\theta}_\delta e^{-\lambda^1 \delta}}^{\theta} h_\delta(\tilde{\theta}) d\tilde{\theta}, \\ G_\delta(\theta) &= \frac{H_\delta(\bar{\theta} e^{-\lambda^1 \delta}) - H_\delta(\theta)}{H_\delta(\bar{\theta} e^{-\lambda^1 \delta})} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1 \right) \theta \frac{h_\delta(\theta)}{H_\delta(\bar{\theta} e^{-\lambda^1 \delta})}. \end{aligned}$$

Substituting $f_\delta(\theta) = f(\theta e^{\lambda^1 \delta} | \delta) e^{\lambda^1 \delta}$ into $h_\delta(\theta)$ and simplifying, I obtain

$$h_\delta(\theta) = \frac{f(\theta e^{\lambda^1 \delta}) e^{\lambda^1 \delta}}{C(1 + \theta e^{\lambda^1 \delta})}, \quad \text{where } C = \int_{\underline{\theta}_\delta}^{\bar{\theta}} \left[1 - p(\theta) (1 - e^{-\lambda^1 \delta}) \right] f(\theta) d\theta.$$

First, I show that the threshold of the pooling segment is $\theta_p e^{-\lambda^1 \delta}$ given f_δ . By simplifying and making a change of variables $z = e^{\lambda^1 \delta} \theta$, I obtain the following

$$\begin{aligned} \int_{\theta_p e^{-\lambda^1 \delta}}^{\bar{\theta} e^{-\lambda^1 \delta}} \left(\eta_\rho \theta - \eta_\alpha \theta_p e^{-\lambda^1 \delta} \right) \frac{f_\delta(\theta)}{1+\theta} d\theta &= \int_{\theta_p e^{-\lambda^1 \delta}}^{\bar{\theta} e^{-\lambda^1 \delta}} \left(\eta_\rho \theta - \eta_\alpha \theta_p e^{-\lambda^1 \delta} \right) \frac{f(\theta e^{\lambda^1 \delta}) e^{\lambda^1 \delta}}{C(1 + \theta e^{\lambda^1 \delta})} d\theta \\ &= \frac{1}{C e^{\lambda^1 \delta}} \int_{\theta_p}^{\bar{\theta}} (\eta_\rho z - \eta_\alpha \theta_p) \frac{f(z)}{1+z} dz. \end{aligned}$$

Therefore, if the threshold of the pooling segment given f is θ_p , then the threshold of the pooling segment at time δ is $\theta_p e^{-\lambda^1 \delta}$.

The condition required by Assumption 2 is that $1 - G_\delta(\theta)$ is nondecreasing for $\theta \leq \theta_p e^{-\lambda^1 \delta}$.

The term $1 - G_\delta(\theta)$ can be written in terms of $f(\theta)$,

$$\begin{aligned} 1 - G_\delta(\theta) &= \frac{1}{H_\delta(\bar{\theta}e^{-\lambda^1\delta})} \left[H_\delta(\theta) + \left(1 - \frac{\eta_\rho}{\eta_\alpha} \right) \theta h_\delta(\theta) \right] \\ &= \frac{1}{C} \left[\int_{\underline{\theta}_\delta}^{\theta} \frac{f(\tilde{\theta}e^{\lambda^1\delta})e^{\lambda^1\delta}}{1 + \tilde{\theta}e^{\lambda^1\delta}} d\tilde{\theta} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1 \right) \theta \frac{f(\theta e^{\lambda^1\delta})e^{\lambda^1\delta}}{1 + \theta e^{\lambda^1\delta}} \right] \\ &= \frac{1}{C} \left[\int_{\underline{\theta}_\delta}^z \frac{f(\tilde{z})}{1 + \tilde{z}} d\tilde{z} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1 \right) z \frac{f(z)}{1 + z} \right], \end{aligned}$$

where the last step is obtained by making a change of variables $z = e^{\lambda^1\delta}\theta$. Therefore, if Assumption 2 holds for all $\theta \leq \theta_p$ given f , then it holds for all $\theta \leq \theta_p e^{-\lambda^1\delta}$ given f_δ .

8.6 Proof of Lemma 4

For $x \in \Omega$ and $0 < \epsilon < 1$, the concavity of L implies that

$$\begin{aligned} L(\tilde{x} + \epsilon(x - \tilde{x})) &\geq L(\tilde{x}) + \epsilon(L(x) - L(\tilde{x})) \\ \implies L(x) - L(\tilde{x}) &\leq \frac{1}{\epsilon} (L(\tilde{x} + \epsilon(x - \tilde{x})) - L(\tilde{x})). \end{aligned}$$

As $\epsilon \rightarrow 0+$, the right-hand side of this equation tends toward $\partial L(\tilde{x}; x - \tilde{x})$. Therefore, I obtain

$$L(x) - L(\tilde{x}) \leq \partial L(\tilde{x}; x - \tilde{x}) = \partial L(\tilde{x}; x) - \partial L(\tilde{x}; \tilde{x}).$$

Given that $\partial L(\tilde{x}; \tilde{x}) = 0$ and $\partial L(\tilde{x}; x) \leq 0$ for all $x \in \Omega$, the sign of $\partial L(\tilde{x}; x - \tilde{x})$ is negative. Therefore, $L(x) \leq L(\tilde{x})$ for all $x \in \Omega$.

8.7 Proof of Lemma 5

I first introduce the problem studied in section 8.4 of Luenberger (1969, p. 217) [21]: $\max_{x \in X} Q(x)$ subject to $x \in \Omega$ and $J(x) \in P$, where Ω is a convex subset of the vector space X , $Q : \Omega \rightarrow \mathbf{R}$ and $J : \Omega \rightarrow Z$ are both concave; where Z is a normed vector space, and P is a nonempty positive

cone in Z . To apply Theorem 1 in Luenberger (1969, p. 217) [21], set

$$\begin{aligned} X &= \{w^1, \underline{w}^0 \mid \underline{w} \in \mathbf{R} \text{ and } w^1 : \Theta \rightarrow \mathbf{R}\}, \\ \Omega &= \hat{\Phi}, \\ Z &= \{z \mid z : \Theta \rightarrow \mathbf{R}^2 \text{ with } \sup_{\theta \in \Theta} \|z(\theta)\| < \infty\}, \\ &\text{with the norm } \|z\| = \sup_{\theta \in \Theta} \|z(\theta)\|, \\ P &= \{z \mid z \in Z \text{ and } z(\theta) \geq (0, 0), \forall \theta \in \Theta\}. \end{aligned}$$

I let the objective function in (OBJ) be Q and the left-hand side of (8) and (9) be defined as J . It is easy to verify that both Q and J are concave. This result holds because the hypotheses of Theorem 1 in Luenberger (1969, p. 217) [21] are met.

8.8 Proof of Proposition 5

Let $a, b \in \Theta$ be such that $a < b < \theta_p$ and $1 - G(a) > 1 - G(b)$ (so Assumption 2 does not hold).

It is easy to verify that the Gâteaux differential $\partial L(w_{x_p}^1, \underline{w}_{x_p}^0; w^1, \underline{w}^0 \mid \tilde{\Lambda}^{\text{se}})$ exists for any $(w^1, \underline{w}^0) \in \hat{\Phi}$. I want to show that a necessary condition that $(w_{x_p}^1, \underline{w}_{x_p}^0)$ maximizes $L(w^1, \underline{w}^0 \mid \tilde{\Lambda}^{\text{se}})$ over $\hat{\Phi}$ is that

$$\partial L(w_{x_p}^1, \underline{w}_{x_p}^0; w^1, \underline{w}^0 \mid \tilde{\Lambda}^{\text{se}}) \leq 0, \quad \forall (w^1, \underline{w}^0) \in \hat{\Phi}, \quad (23)$$

$$\partial L(w_{x_p}^1, \underline{w}_{x_p}^0; w_{x_p}^1, \underline{w}_{x_p}^0 \mid \tilde{\Lambda}^{\text{se}}) = 0. \quad (24)$$

If $(w_{x_p}^1, \underline{w}_{x_p}^0)$ maximizes $L(w^1, \underline{w}^0 \mid \tilde{\Lambda}^{\text{se}})$, then for any $(w^1, \underline{w}^0) \in \hat{\Phi}$, it must be true that

$$\left. \frac{d}{d\epsilon} L((w_{x_p}^1, \underline{w}_{x_p}^0) + \epsilon((w^1, \underline{w}^0) - (w_{x_p}^1, \underline{w}_{x_p}^0)) \mid \tilde{\Lambda}^{\text{se}}) \right|_{\epsilon=0} \leq 0.$$

Hence, $\partial L(w_{x_p}^1, \underline{w}_{x_p}^0; (w^1, \underline{w}^0) - (w_{x_p}^1, \underline{w}_{x_p}^0) \mid \tilde{\Lambda}^{\text{se}}) \leq 0$. Setting $(w^1, \underline{w}^0) = (w_{x_p}^1, \underline{w}_{x_p}^0)/2 \in \hat{\Phi}$ yields $\partial L(w_{x_p}^1, \underline{w}_{x_p}^0; w_{x_p}^1, \underline{w}_{x_p}^0 \mid \tilde{\Lambda}^{\text{se}}) \geq 0$. By the definition of $(w_{x_p}^1, \underline{w}_{x_p}^0)$, there exists $\epsilon > 0$ sufficiently small such that $(1 + \epsilon)(w_{x_p}^1, \underline{w}_{x_p}^0) \in \hat{\Phi}$. Setting $(w^1, \underline{w}^0) = (1 + \epsilon)(w_{x_p}^1, \underline{w}_{x_p}^0)$ yields $\partial L(w_{x_p}^1, \underline{w}_{x_p}^0; w_{x_p}^1, \underline{w}_{x_p}^0 \mid \tilde{\Lambda}^{\text{se}}) \leq 0$. Together, (23) and (24) obtain.

The last step is to show that there exists no $\tilde{\Lambda}^{\text{se}}$ that satisfies the first-order conditions (23) and (24). Here, I use the same approach as in the proof of Proposition 4 in Amador, Werning, and Angeletos (2006) [4]. The Gâteaux differential $\partial L(w_{x_p}^1, \underline{w}_{x_p}^0; w^1, \underline{w}^0 \mid \tilde{\Lambda}^{\text{se}})$ is similar to equation (16) with θ_p replaced by x_p . Conditions (23) and (24) imply that $\tilde{\Lambda}^{\text{se}}(\underline{\theta}) = 0$. Integrating the

Gâteaux differential by parts yields

$$\partial L(w_{x_p}^1, \underline{w}_{x_p}^0; w^1, \underline{w}^0 \mid \tilde{\Lambda}^{\text{se}}) = \chi(\underline{\theta})w^1(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \chi(\theta)dw^1(\theta), \quad (25)$$

with

$$\chi(\theta) \equiv \eta_\alpha \int_{\underline{\theta}}^{\bar{\theta}} \left[\tilde{\Lambda}^{\text{se}}(\tilde{\theta}) - (1 - G(\tilde{\theta})) \right] d\tilde{\theta} + \eta_\alpha \int_{\max\{x_p, \theta\}}^{\bar{\theta}} (\tilde{\theta} - x_p) d\tilde{\Lambda}^{\text{se}}(\tilde{\theta}).$$

By condition (23), it follows that $\chi(\theta) \leq 0$ for all θ . Condition (24) implies that $\chi(\theta) = 0$ for $\theta \in [\underline{\theta}, x_p]$. It follows that $\tilde{\Lambda}^{\text{se}}(\theta) = 1 - G(\theta)$ for all $\theta \in (\underline{\theta}, x_p]$. This implies that $x_p \leq b$ otherwise the associated multiplier $\tilde{\Lambda}^{\text{se}}$ would be decreasing. Integrating by parts the second term of $\chi(\theta)$, I obtain

$$\chi(\theta) = \int_{\underline{\theta}}^{\bar{\theta}} G(\tilde{\theta})d\tilde{\theta} + (\theta - x_p)(1 - \tilde{\Lambda}^{\text{se}}(\theta)), \quad \forall \theta \geq x_p.$$

By definition of θ_p , there must exist a $\theta \in [x_p, \theta_p)$ such that the first term is strictly positive; since $\tilde{\Lambda}^{\text{se}}(\theta) \leq 1$, the second term is nonnegative. Hence $\chi(\theta) > 0$, contradicting the necessary conditions. This completes the proof.

8.9 Proof of Lemma 6

Given $\gamma = (\gamma^1, \gamma^0) \in \mathbf{R}^2$, I let $K(\gamma)$ and $\mathcal{H}(\gamma)$ denote the supremum score in direction γ and the associated half space. And \mathcal{H} is the intersection of all half spaces. Since $\Gamma \subset \mathcal{H}(\gamma)$ for any γ and Γ is convex, it follows that $\Gamma = \mathcal{H}$. If $\gamma^1 \geq 0, \gamma^0 \geq 0$, $K(\gamma)$ equals $\gamma^1 + \gamma^0$, achieved by the policy which directs all resources to R . If $\gamma^1 \leq 0, \gamma^0 \leq 0$, $K(\gamma)$ equals 0, achieved by the policy which directs all resources to S .

If $\gamma^1 > 0, \gamma^0 < 0$, finding $K(\gamma)$ is equivalent to a decision problem of player i who chooses π to maximize $U_i(\pi, p)$. His prior p is given by $p/(1-p) = -\gamma^1/(\gamma^0\eta_i)$. According to Keller and Rady (2010) [18], the optimum is a cutoff policy under which R is used when the belief is above the cutoff and S is used below. The cutoff belief, denoted p_i^* , satisfies the equation $p_i^*/(1-p_i^*) = \mu/((1+\mu)\eta_i)$ where μ is the positive root of equation $r + \lambda^0 - \mu(\lambda^1 - \lambda^0) = \lambda^0(\lambda^0/\lambda^1)^\mu$. Let $\bar{U}_i(p) \equiv \max_\pi U_i(\pi, p)$. If $p \leq p_i^*$, $\bar{U}_i(p)$ equals s_i . Hence, $K(\gamma) = 0$. If $p > p_i^*$,

$$\bar{U}_i(p) = \lambda(p)h_i + (s_i - \lambda(p_i^*)h_i) \frac{1-p}{1-p_i^*} \left(\frac{(1-p)p_i^*}{p(1-p_i^*)} \right)^\mu, \quad \text{where } \lambda(p) = \lambda^1 p + \lambda^0(1-p).$$

Hence, I obtain

$$K(\gamma) = -\frac{\gamma^0 \left(-\frac{\gamma^0 \mu}{\gamma^1 \mu + \gamma^1} \right)^\mu}{\mu + 1} + \gamma^1 + \gamma^0. \text{ }^{26}$$

It is easy to verify that the southeast boundary of Γ is pinned down by the pairs

$$(w^1, w^0) = \left(1 - \left(-\frac{\gamma^0 \mu}{\gamma^1 \mu + \gamma^1} \right)^{\mu+1}, 1 - \left(-\frac{\gamma^0 \mu}{\gamma^1 \mu + \gamma^1} \right)^\mu \right).$$

The functional form of the southeast boundary is

$$\beta^{\text{se}}(w^1) = 1 - (1 - w^1)^{\frac{\mu}{\mu+1}}, w^1 \in [0, 1].$$

If $\gamma^1 < 0, \gamma^0 > 0$, finding $K(\gamma)$ is equivalent to a decision problem of player i who chooses π to *minimize* $U_i(\pi, p)$. His prior p is given by $p/(1-p) = -\gamma^1/(\gamma^0 \eta_i)$. As shown by Keller and Rady (2013) [19], the optimum is a cutoff policy under which S is used if the belief is above the cutoff and R is used if below. Let p_i^{**} denote this cutoff belief and $\underline{U}_i(p) \equiv \min_\pi U_i(\pi, p)$. The function $\underline{U}_i(p)$ is continuous, concave, and non-decreasing. Except for a kink at p_i^{**} , $\underline{U}_i(p)$ is once continuously differentiable. When $\gamma^1 < 0, \gamma^0 > 0$, $K(\gamma)$ equals $(\underline{U}_i(p) - s_i)(\gamma^1/\eta_i - \gamma^0)/(s_i - \lambda^0 h_i) \geq 0$. Hence, the northwest boundary of Γ is concave, nondecreasing, once continuously differentiable, with the end points being $(0, 0)$ and $(1, 1)$.

8.10 Proof of Proposition 7

Suppose that a success occurs at time t . Before the success, the principal's belief about the type distribution is denoted f_{t-} . Without loss of generality, suppose that the support is $[\underline{\theta}, \bar{\theta}]$. Here, I calculate the updated belief after the success, denoted f_t . Let $Q(\theta, dt)$ be the probability that a success occurs in an infinitesimal interval $[t, t + dt)$ given type θ

$$\begin{aligned} Q(\theta, dt) &= p(\theta)(1 - e^{-\lambda^1 dt}) + (1 - p(\theta))(1 - e^{-\lambda^0 dt}) \\ &= (1 - e^{-\lambda^0 dt}) \left(p(\theta) \frac{1 - e^{-\lambda^1 dt}}{1 - e^{-\lambda^0 dt}} + 1 - p(\theta) \right). \end{aligned}$$

²⁶If $\gamma^1 > 0, \gamma^0 < 0$, $K(\gamma)$ equals $(\bar{U}_i(p) - s_i)(\gamma^1/\eta_i - \gamma^0)/(s_i - \lambda^0 h_i) \geq 0$.

Given a success at time t , the principal's updated belief about the agent's type distribution, in terms of his type at time $t-$, is

$$\begin{aligned} f_t^*(\theta) &= \lim_{dt \rightarrow 0} \frac{f_{t-}(\theta)Q(\theta, dt)}{\int_{\underline{\theta}}^{\bar{\theta}} f_{t-}(\theta)Q(\theta, dt)d\theta} \\ &= \frac{f_{t-}(\theta) [p(\theta)\lambda^1 + (1 - p(\theta))\lambda^0]}{\int_{\underline{\theta}}^{\bar{\theta}} f_{t-}(\theta) [p(\theta)\lambda^1 + (1 - p(\theta))\lambda^0] d\theta}, \quad \forall \theta \in [\underline{\theta}, \bar{\theta}]. \end{aligned}$$

After the success, the agent of type θ updates his belief about the state to the odds ratio $\theta\lambda^1/\lambda^0$. Therefore, the principal's belief about the agent's type distribution, in terms of his type at time t , is

$$f_t(\theta) = \begin{cases} f_t^*(\theta\lambda^0/\lambda^1) \lambda^0/\lambda^1 & \text{if } \theta \in [\underline{\theta}\lambda^1/\lambda^0, \bar{\theta}\lambda^1/\lambda^0], \\ 0 & \text{otherwise.} \end{cases}$$

Given the distribution f_t , I define h_t, H_t, G_t as follows:

$$\begin{aligned} h_t(\theta) &= \frac{f_t(\theta)}{1 + \theta}, \quad \text{and} \quad H_t(\theta) = \int_{\underline{\theta}\lambda^1/\lambda^0}^{\theta} h_t(\tilde{\theta})d\tilde{\theta}, \\ G_t(\theta) &= \frac{H_t(\bar{\theta}\lambda^1/\lambda^0) - H_t(\theta)}{H_t(\bar{\theta}\lambda^1/\lambda^0)} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1 \right) \theta \frac{h_t(\theta)}{H_t(\bar{\theta}\lambda^1/\lambda^0)}. \end{aligned}$$

Substituting $f_t(\theta) = f_t^*(\theta\lambda^0/\lambda^1) \lambda^0/\lambda^1$ into $h_t(\theta)$ and simplifying, I obtain

$$h_t(\theta) = \frac{f_{t-}(\theta\lambda^0/\lambda^1)}{C(1 + \theta\lambda^0/\lambda^1)}, \quad \text{where } C = \frac{\lambda^1}{(\lambda^0)^2} \int_{\underline{\theta}}^{\bar{\theta}} f_{t-}(\theta) [p(\theta)\lambda^1 + (1 - p(\theta))\lambda^0] d\theta.$$

Following the same argument as in Subsection 8.5, I can show that (i) if the threshold of the pooling segment given f_{t-} is θ_p , then the threshold of the pooling segment given f_t is $\theta_p\lambda^1/\lambda^0$; (ii) if Assumption 2 holds for $\theta \leq \theta_p$ given f_{t-} , then it holds for $\theta \leq \theta_p\lambda^1/\lambda^0$ given f_t . This completes the proof.

8.11 Proof of Lemma 7

Given $\gamma = (\gamma^1, \gamma^0) \in \mathbf{R}^2$, I let $K(\gamma)$ and $\mathcal{H}(\gamma)$ denote the supremum score in direction γ and the associated half space. And \mathcal{H} is the intersection of all half spaces. Since $\Gamma \subset \mathcal{H}(\gamma)$ for any γ and Γ is convex, it follows that $\Gamma = \mathcal{H}$. If $\gamma^1 \geq 0, \gamma^0 \geq 0$, $K(\gamma)$ equals $\gamma^1 + \gamma^0$, achieved by the policy which uses R exclusively. If $\gamma^1 \leq 0, \gamma^0 \leq 0$, $K(\gamma)$ equals 0, achieved by the policy which uses S exclusively.

If $\gamma^1 > 0, \gamma^0 < 0$, finding $K(\gamma)$ is equivalent to a decision problem of player i who chooses

π to maximize $U_i(\pi, p)$. His prior p is given by $p/(1-p) = -\gamma^1/(\gamma^0\eta_i)$. According to Cohen and Solan (2013) [11], the optimum is a cutoff policy under which R is used when the belief is above the cutoff and S is used below. The cutoff belief, denoted p_i^* , satisfies the equation $p_i^*/(1-p_i^*) = a^*/((1+a^*)\eta_i)$, where a^* is the positive root of equation 6.1 in Cohen and Solan (2013) [11].

Let $\bar{U}_i(p) \equiv \max_{\pi} U_i(\pi, p)$. If $p \leq p_i^*$, $\bar{U}_i(p)$ equals s_i . Hence, $K(\gamma) = 0$. If $p > p_i^*$,

$$\bar{U}_i(p) = p\mu^1 + (1-p)\mu^0 + [s_i - (p_i^*\mu^1 + (1-p_i^*)\mu^0)] \frac{1-p}{1-p_i^*} \left(\frac{(1-p)p_i^*}{p(1-p_i^*)} \right)^{a^*}.$$

Hence, I obtain

$$K(\gamma) = -\frac{\gamma^0 \left(-\frac{\gamma^0 a^*}{\gamma^1 a^* + \gamma^1} \right)^{a^*}}{a^* + 1} + \gamma^1 + \gamma^0. \text{ } ^{27}$$

It is easy to verify that the southeast boundary of Γ is pinned down by the pairs

$$(w^1, w^0) = \left(1 - \left(-\frac{\gamma^0 a^*}{\gamma^1 a^* + \gamma^1} \right)^{a^*+1}, 1 - \left(-\frac{\gamma^0 a^*}{\gamma^1 a^* + \gamma^1} \right)^{a^*} \right).$$

The functional form of the southeast boundary is

$$\beta^{\text{se}}(w^1) = 1 - (1 - w^1)^{\frac{a^*}{a^*+1}}, w^1 \in [0, 1].$$

²⁷If $\gamma^1 > 0, \gamma^0 < 0$, $K(\gamma)$ equals $(\bar{U}_i(p) - s_i)(\gamma^1/\eta_i - \gamma^0)/(s_i - \mu^0) \geq 0$.

References

- [1] Aghion, Philippe and Jean Tirole. 1997. Formal and Real Authority in Organizations. *Journal of Political Economy*, 105(1): 1–29.
- [2] Alonso, Ricardo and Niko Matouschek. 2008. Optimal Delegation. *Review of Economic Studies*, 75(1): 259–293.
- [3] Amador, Manuel and Kyle Bagwell. 2013. The Theory of Optimal Delegation with an Application to Tariff Caps. *Econometrica*, 81(4): 1541–1599.
- [4] Amador, Manuel, Iván Werning and George-Marios Angeletos. 2006. Commitment vs. Flexibility. *Econometrica*, 74(2): 365–396.
- [5] Ambrus, Attila and Georgy Egorov. 2013. Comment on “Commitment vs. Flexibility”. *Econometrica*, 81(5): 2113–2124.
- [6] Aumann, Robert J.. 1964. Mixed and Behavior Strategies in In Extensive Games. *Advances in Game Theory*, Princeton University Press, Princeton, NJ, 627–650.
- [7] Bergemann, Dirk and Ulrich Hege. 1998. Venture Capital Financing, Moral Hazard, and Learning. *Journal of Banking and Finance*, 22(6): 703–735.
- [8] Bergemann, Dirk and Ulrich Hege. 2005. The Financing of Innovation: Learning and Stopping. *RAND Journal of Economics*, 36(4): 719–752.
- [9] Bolton, Patrick and Christopher Harris. 1999. Strategic Experimentation. *Econometrica*, 67(2): 349–374.
- [10] Coase, Ronald H.. 1937. The Nature of the Firm. *Economica*, 4: 386–405.
- [11] Cohen, Asaf and Eilon Solan. 2013. Bandit Problems with Lévy Processes. *Mathematics of Operations Research*, 38(1): 92–107.
- [12] Garfagnini, Umberto. 2011. Delegated Experimentation. *Working paper*.
- [13] Gomes, Renato, Daniel Gottlieb and Lucas Maestri. 2013. Experimentation and Project Selection: Screening and Learning. *Working paper*.
- [14] Halac, Marina, Navin Kartik and Qingmin Liu. 2013. Optimal Contracts for Experimentation. *Working paper*.
- [15] Holmström, Bengt. 1977. On Incentives and Control in Organizations. *PhD dissertation*, Stanford University.
- [16] Holmström, Bengt. 1984. On the Theory of Delegation. in *Bayesian Models in Economic Theory*, New York: North-Holland, 115–141.
- [17] Hörner, Johannes, and Larry Samuelson. 2013. Incentives for Experimenting Agent. *RAND Journal of Economics*, forthcoming.

- [18] Keller, Godfrey, and Sven Rady. 2010. Strategic Experimentation with Poisson Bandits. *Theoretical Economics*, 5(2): 275–311.
- [19] Keller, Godfrey, and Sven Rady. 2013. Breakdowns. *Working paper*.
- [20] Keller, Godfrey, Sven Rady and Martin Cripps. 2005. Strategic Experimentation with Exponential Bandits. *Econometrica*, 73(1): 39–68.
- [21] Luenberger, David G.. 1969. *Optimization by Vector Space Method*. Series in Decision and Control. New York: Wiley.
- [22] Milgrom, Paul and Ilya Segal. 2002. Envelope Theorems for Arbitrary Choice Sets. *Econometrica*, 70(2): 583–601.
- [23] Presman, Ernst L.. 1990. Poisson Version of the Two-Armed Bandit Problem with Discounting. *Theory of Probability and Its Applications*, 35: 307–317.
- [24] Strulovici, Bruno. 2010. Learning While Voting: Determinants of Collective Experimentation. *Econometrica*, 78(3): 933–971.
- [25] Williamson, Oliver E.. 1975. *Markets and Hierarchies*. New York: Free Press.