American Economic Review 2016, 106(8): 1969–2008
http://dx.doi.org/10.1257/aer.20141215

Dynamic Delegation of Experimentation†

By Yingni Guo*

I study a dynamic relationship where a principal delegates experimentation to an agent. Experimentation is modeled as a one-armed bandit that yields successes following a Poisson process. Its unknown intensity is 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 does, so prefers more experimentation. The optimal mechanism is a cutoff rule in the belief space: the cutoff gives pessimistic types total freedom but curtails optimistic types' behavior. Pessimistic types overexperiment while the most optimistic ones underexperiment. This delegation rule is time consistent. (*JEL* D23, D82, D83, O30)

Innovation carries great uncertainty. Firms frequently start R&D projects with little knowledge about the likelihood of eventual success. If experimentation goes on but no success occurs, firms grow pessimistic and reduce their resource input or even discontinue the project altogether. In this paper, I study how to manage innovation in a hierarchical organization. Specifically, I study the optimal mechanism by which a principal ("she") delegates experimentation to an agent ("he"), as in the case of a firm delegating an R&D project to an employee.

The current literature on delegation focuses mainly on static problems in which the principal determines a permissible set of actions from which the agent is to choose. Designing this permissible set is a nontrivial task because typically the agent has better information and is biased. On the one hand, the principal wants to give the agent more flexibility in order to use his information; on the other, she may wish to offer a smaller permissible set in order to limit the expression of the agent's bias. In this paper, I consider delegation in a dynamic environment with learning. Unlike static delegation, the two parties learn more about the underlying state over time. As experimentation unfolds and new information arrives, how should the principal adjust the flexibility granted to the agent accordingly? Such a problem arises in

* Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60201 (e-mail: yingni.guo@northwestern.edu). 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 discussions and comments. I also thank three anonymous referees for their constructive comments. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

† Go to http://dx.doi.org/10.1257/aer.20141215 to visit the article page for additional materials and author disclosure statement.

many different situations when contingent transfers are severely restricted or infeasible. For instance, a pharmaceutical firm delegates decision rights over resource allocation to its employee to develop a new drug, which might be more or less profitable than improving on the existing ones. The headquarters delegates its divisional manager to develop a new product which might be a success or failure. Constituents (or government agencies) delegate decision rights to their representatives (or their employees) to experiment with policy reforms.[1]

To understand how to optimally delegate experimentation, I model experimentation as a continuous-time bandit problem similar to that of Presman (1990; Keller, Rady, and Cripps (2005); and Keller and Rady (2010). There is one unit of a perfectly divisible resource per unit of time and the agent continually splits the resource between a risky project (e.g., developing a new drug) and a safe one. In any given time interval, the safe project generates a known flow payoff proportional to the resource allocated to it.[2] The risky project's payoff depends on an unknown binary state. In the benchmark setting, if the state is good, the risky project yields successes at random times. The arrival rate is proportional to the resource allocated to it. If the state is bad, the risky project yields no successes at all. (In the example of developing a new drug, the first breakthrough reveals that the new drug is superior.) After experimentation begins, the agent's actions and the arrivals of successes are publicly observed.

The risky project requires resources from the firm and time from the agent. Both prefer to allocate the resource to the risky project if and only if the state is good. Therefore, if the state were known, there would be no conflict of interest. When the state is unknown, however, the two parties' preferences might differ. They both wish to discontinue the risky project if they become pessimistic enough. However, the agent values the risky project's successes over the safe project's flow payoffs more than the principal does (due to, for instance, the high cost of the principal's resources, the principal's moderate benefit from one project out of her many responsibilities, or the agent's career advancement as an extra benefit). Hence, the agent prefers to keep the risky project alive for a longer time if faced with a prolonged absence of success.[3]

Building on his expertise, the agent has private knowledge at the outset about the prospect of success for the risky project. Formally, the agent knows more precisely the probability that the state is good, with his type being his prior belief of the good state. A promising project warrants longer experimentation. If the principal wishes to take advantage of the agent's information, she has to give him some flexibility over resource allocation. However, their preferences are not perfectly aligned, which curtails the flexibility that she is willing to grant him.

The principal delegates the decision as to how the agent should allocate the resource over time. This decision is delegated 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. The purpose of this paper is

[1] Other examples include a hospital delegating decision rights over resource allocation to its surgeon to test a new procedure that might be more or less effective than the old one, and a university delegating the department chair to experiment with a new academic program.

[2] This flow payoff can be interpreted as the cost saved or the payoff from conducting conventional tasks.

[3] This assumption is later relaxed and the case in which the bias goes in the other direction is also studied.

to solve for the optimal delegation rule. More specifically, it addresses the following questions: in the absence of transfers, what instruments are used by the principal 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, and how much retained by the principal? Will some projects involve overexperimentation and others underexperimentation? Is the optimal delegation rule time consistent?

Solving for the optimal delegation contract directly is difficult, since the policy space is quite large. In particular, this is not a stopping-time problem, since the agent is allowed not only to switch back and forth between the two projects, but also to scale up and down the resource inputs to the projects. Possible policies include: allocating the whole resource to the risky project until a fixed time and then switching all of it to the safe project only if no success has been realized; gradually reducing the resource input to the risky project if no success occurs and allocating the whole resource to it after the first success; allocating the whole resource to the risky project until the first success and then allocating only a fixed fraction to the risky project; always allocating a fixed fraction of the unit resource to the risky project, etc. A key observation is that any policy, in terms of payoffs, can be summarized by a pair of numbers, corresponding to (i) the *total expected discounted resource* allocated to the risky project conditional on the state being good, and (ii) the total expected discounted resource allocated to the risky project conditional on the state being bad. As far as payoffs are concerned, this constitutes a simple, finite-dimensional summary statistic for any given policy. The range of these summary statistics as policies are varied is what I call the feasible set—a subset of the plane. Determining the feasible set is a nontrivial problem in general, but it involves no incentive constraints and therefore reduces to a standard optimization problem. This reduces the delegation problem to a static one. Given that the resulting problem is static, I use Lagrangian optimization methods (similar to those used by Amador, Werning, and Angeletos 2006 and Amador and Bagwell 2013) to solve for the optimal delegation rule.

Under the regularity condition that the type distribution is not too skewed toward lower types, the optimal delegation rule takes a quite 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. This belief drifts downward when no success is observed and jumps to 1 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 how to allocate the resource as long as the updated belief remains above the cutoff. However, if this belief ever reaches the cutoff, the agent is asked to stop experimenting. This rule turns out not to bind for types with low enough priors, who voluntarily stop experimenting when no success occurs, 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 1 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 the whole resource to the risky project after the first success, but is not allowed to experiment past the deadline. When no success occurs, those types with low enough priors stop experimenting before the deadline, whereas a positive measure of types with high enough priors stops at the deadline. In equilibrium, there is no delay in information acquisition since the risky project is operated exclusively until either the first success reveals that the state is good or the agent stops. Within the set of high enough types who are forced to stop when the cutoff (or the deadline) is reached, the highest subset underexperiment even from the principal's point of view. Every other type overexperiments. A prediction of my model is that 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 revise the cutoff rule even if she were given a chance to do so. In particular, if, after the agent experiments for some time, no success has yet been realized, then the principal still finds it optimal to keep the cutoff (or the deadline) at the level set at the beginning. To understand why the cutoff rule is time consistent, note that the cutoff has no impact on the low types who are not constrained by the cutoff. These types implement their preferred policies and stop voluntarily after some time has elapsed without a success. Given that the cutoff creates no distortions on lower types' behaviors, the principal chooses the cutoff by focusing only on those high types who are constrained by the cutoff. The principal chooses the cutoff such that she finds it optimal to stop experimenting when the cutoff becomes binding. Hence, implementing the cutoff rule requires no commitment on the principal's side.

I next show that both the optimality of the cutoff rule and its time consistency generalize to situations in which the risky project generates successes in the bad state as well but less frequently. (Take policy reforms in the public sector as an example. The learning process is noisy in the sense that a bad policy reform generates successes as well but at a lower rate.) The arrival of successes is indicative of the good state but is not conclusive proof of one. When successes are inconclusive, the belief is updated differently than in the benchmark setting. It jumps up upon successes and then drifts downward. Consequently, the cutoff rule cannot be implemented by imposing a deadline. Instead, it can be implemented as a sliding deadline. The principal initially extends some time to the agent to operate the risky project; and, whenever a success is realized, more time is granted. The agent is free to switch to the safe project before he uses up the time granted by the principal. But, if a long stretch of time goes by without any further success, the agent is required to switch to the safe project.[4]

I further extend the analysis to the case in which the agent gains less from the experimentation than the principal does, and therefore tends to underexperiment. This happens when an innovative project 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 as in the previous case, the principal imposes a floor. The agent has no flexibility but to experiment before the lockup period ends, yet has full flexibility afterward. Time consistency is no longer preserved, since whenever the agent stops experimenting voluntarily, he reveals that the principal's optimal experimentation

[4] In online Appendix B3, I show that the optimality of the cutoff rule and its time-consistent property generalizes to Lévy bandits (Cohen and Solan 2013).

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.

These results have important implications for the practical design of delegation rules. When the agent is perceived to gain more from a successful innovation, 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. On the other hand, the agent should have flexibility before the (sliding) deadline is reached. The (sliding) deadline rule has already seen its application in real life scenarios. For instance, in the public sector, policymakers are better informed than their constituents, and tend to prolong the policy experimentation since they gain the most popularity from successful reforms they initiated. They typically are given the flexibility over resource allocation as long as the reforms go well. However, if a reform fails to deliver, it will be discontinued in the form of election defeat, legislative repeal, or automatic expiration with a sunset provision.[5] By contrast, when the agent is biased toward the safe project, the principal should impose a (sliding) lockup period and give the agent freedom only after the (sliding) lockup period ends. Seru (2014) finds evidence that established high-level managers in conglomerates are less willing to put resources in innovative projects for fear of resource reallocation in the case of failure. The headquarters will benefit from imposing a lockup period, forcing these established managers to run the experimentation longer and be more innovative.

My paper contributes to the literature on delegation. This literature addresses the incentive problems in organizations due to hidden information and misaligned preferences. Holmström (1977, 1984) 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 that, for the most part, only interval delegation sets are allowed.[6] Melumad and Shibano (1991) are the first to fully characterize the solution to the delegation problem with no restrictions on the delegation sets. They do so in a model with quadratic loss functions in which the state is uniformly distributed and the preferred decisions are linear in the state. Alonso and Matouschek (2008) and Amador and Bagwell (2013) characterize the optimal delegation set in general environments under certain conditions and provide conditions under which interval delegation is optimal. This paper extends the analysis to the environment with learning, and facilitates the understanding of how to optimally delegate experimentation.

Despite the difference in motivation and setup, I borrow numerous ideas from Amador, Werning, and Angeletos (2006) and Amador and Bagwell (2013) to prove the optimality of interval delegation once the problem is reduced to a static one. Also, the condition and intuition for interval delegation to be optimal are quite

[5] Another example involves funding agencies awarding grants for scholarly research. (Suppose that experimentation has a flow cost. The agent is cash-constrained and his action is contractible. Delegating experimentation means funding his research project.) Funding agencies promise to fund the research for a fixed period, and extend the funding only upon demonstrated results. At the same time, the scholar is free to discontinue the project whenever he gets pessimistic enough simply by not claiming the designated fund.

[6] There is a large literature that followed Holmström's original work and characterized the optimal delegation contract and/or its implementation in various settings. See, for instance, Athey, Atkeson, and Kehoe (2005); Martimort and Semenov (2006); Mylovanov (2008); Armstrong and Vickers (2010); and Frankel (2014).

similar to those in the delegation literature. I discuss the details of the relationship with Alonso and Matouschek (2008), Amador, Werning, and Angeletos (2006), and Amador and Bagwell (2013) in Section IVB after I present the main results.

This paper also contributes to the study of how to optimally allocate formal and real authority within organizations. See, for example, Aghion and Tirole (1997); Dessein (2002); and Grenadier, Malenko, and Malenko (2016). Grenadier, Malenko, and Malenko (2016) examine a model in which a principal decides when to exercise an option and relies on the information of an informed yet biased agent. They also demonstrate that there is an asymmetry in results when the agent has a late versus an early exercise bias due to the irreversibility of time.

This paper is related to the literature on experimentation in a principal-agent setting with a single agent. Most papers in this literature assume transferable utilities and unobservable actions. Bergemann and Hege (1998, 2005) and Hörner and Samuelson (2013) analyze dynamic moral hazard models in which the principal and the agent learn about the value of the project over time. Halac, Kartik, and Liu (forthcoming b) study long-term contracts for experimentation, with adverse selection about the agent's ability and moral hazard about his effort choice. Manso (2011) derives an optimal contract when the agent chooses between shirking, a safe arm, or a risky arm, and shows that the optimal contract exhibits tolerance for early failure. Klein (forthcoming) studies how to incentivize an agent to conduct innovative activities when the agent has access to a known technology that produces fake successes. Gerardi and Maestri (2012) study optimal sequential testing in which an agent exerts costly effort to acquire information. Both his actions and the realized signals are unobservable.[7]

This paper is mostly closely related to Gomes, Gottlieb, and Maestri (2016) and Garfagnini (2011). Gomes, Gottlieb, and Maestri (2016) study the optimal financing contract in a setting with experimentation and adverse selection. The agent's actions are observable, but he has private information about both the project quality and his cost of effort. Unlike my setting, their agent gains no benefit from the project and outcome-contingent transfers are allowed. They identify necessary and sufficient conditions under which the principal pays rents only for the agent's information about his cost, but not for the agent's information about the project quality. Garfagnini (2011) studies a dynamic delegation model with no 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 precontractual hidden information; transfers are infeasible and the principal is able to commit to long-term contract terms; and the agent gains a direct benefit from experimentation and shares the same preferences as the principal conditional on the state. Agency conflicts arise since

[7] More broadly, my paper is related to the literature on incentives for experimentation. Bolton and Harris (1999); Keller, Rady, and Cripps (2005); Rosenberg, Solan, and Vieille (2007); Keller and Rady (2010, 2015); Strulovici (2010); Klein and Rady (2011); and Murto and Välimäki (2011) analyze the strategic interaction among experimenting agents. Bonatti and Hörner (2011) study experimentation in teams with unobservable actions. Ederer (2013) extends the setting of Manso (2011) to a setting with multiple agents, and derives the optimal contract when agents explore different risky projects. Moroni (2016) studies the optimal contract when the principal contracts with multiple agents who work on the same project which consists of several stages. Halac, Kartik, and Liu (forthcoming a) study the contest design for experimentation.

the agent is inclined to exaggerate the prospects for success in order to prolong the experimentation.

The paper is organized as follows. The model is presented in Section I. Section II considers a single player's decision problem. In Section III, I illustrate how to reduce the delegation problem to a static one. The main results are presented in Section IV. I extend the analysis to more general stochastic processes in Section V and discuss other extensions of the model. Section VI concludes.

I. The Model

Players, Tasks, and States.—Time $t \in [0, \infty)$ is continuous. There are two risk-neutral players $i \in \{\alpha, \rho\}$: a (male) agent and a (female) principal; 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 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 $1 - \pi_t$ is allocated 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 of these data are common knowledge.[8]

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

Conflicts of Interest.—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 that $\lambda^1 h_i > s_i > \lambda^0 h_i$ for $i \in \{\alpha, \rho\}$, and

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

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

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

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

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 1, measures how misaligned the 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 Section VB.)

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 by θ and referred to as the agent's type. The agent's type is drawn from a compact interval $\Theta \equiv \left[\underline{\theta}, \overline{\theta}\right] \subset \mathbb{R}_+$ 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 type distribution. Hence, her prior belief that the state is 1 is given by

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

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 nonanticipative 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 a success occurs at the random time τ_1, then, depending on the value of τ_1, a new function $\pi(t\,|\,\tau_1, 1)$ is chosen, and so forth. The space of all policies, including randomized ones, is denoted as Π (see footnote 10).

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

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

Here, the expectation is taken over the state 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, as shown here:

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

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, it follows that, before the first success of the process N_t, the process p_t satisfies this differential equation:

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

At the first success, p_t jumps to 1.[9]

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 determines 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 $\boldsymbol{\pi} : \Theta \rightarrow \Pi$ to maximize her expected payoff, subject to the agent's incentive compatibility (IC) constraint. Formally, I solve the following:

$$\sup \int_{\Theta} U_{\rho}(\boldsymbol{\pi}(\theta), p(\theta))\, dF(\theta),$$

subject to $\quad U_{\alpha}(\boldsymbol{\pi}(\theta), p(\theta)) \geq U_{\alpha}(\boldsymbol{\pi}(\theta'), p(\theta)) \quad \forall \theta, \theta' \in \Theta,$

over measurable $\boldsymbol{\pi} : \Theta \rightarrow \Pi$.[10]

II. 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 proposition 3.1 in Keller, Rady, and Cripps 2005, for instance), the cutoff belief is

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

[9] In general, subscripts indicate either the time or the player, and superscripts indicate the state. Parentheses contain the type or the policy.

[10] Here, I define randomized policies and stochastic mechanisms, following Aumann (1964). Let $\mathcal{B}_{[0,1]}$ (or $\mathcal{B}^k$) denote the σ-algebra of Borel sets of $[0, 1]$ (or $\mathbb{R}_+^k$) and λ the Lebesgue measure on $[0, 1]$, where k is a positive integer. I denote the set of measurable functions from $(\mathbb{R}_+^k, \mathcal{B}^k)$ 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 $s \in \mathbb{R}_+^k$ 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), I define randomized policies as measurable functions $\hat{\pi} : [0, 1] \rightarrow \Pi^*$. According to $\hat{\pi}$, a value $\varepsilon \in [0, 1]$ is drawn uniformly from $[0, 1]$ and then the pure policy $\hat{\pi}(\varepsilon)$ is implemented. Analogously, I define stochastic mechanisms as measurable functions $\pi : [0, 1] \times \Theta \rightarrow \Pi^*$. A value $\varepsilon \in [0, 1]$ 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.

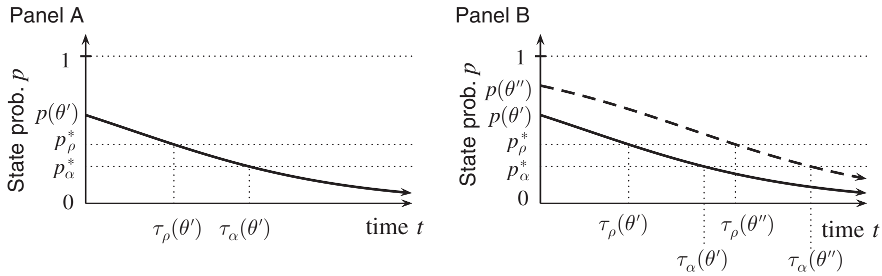

FIGURE 1. THRESHOLDS AND STOPPING TIMES

Note: Parameters are $\eta_\alpha = 3/2, \eta_\rho = 3/4, \lambda^1 = 1, \theta' = 3/2, \theta'' = 4$.

Note that 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^*, since he values *R*'s successes over *S*'s flow payoffs more than the principal does.

Given the law of motion of belief (1) and the cutoff belief (2), player *i*'s preferred policy given prior $p(\theta)$ is a *stopping-time policy* with a fixed 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 the odds ratio $\theta \in \Theta$ is*

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

Figure 1 illustrates the two players' cutoff beliefs and their preferred stopping times associated with two possible odds ratios θ', θ'' (with $\theta' < \theta''$). The prior beliefs are thus $p(\theta')$, $p(\theta'')$ (with $p(\theta') < p(\theta'')$). The x-axis variable is time and the y-axis variable is the posterior belief. On the y-axis are the two players' cutoff beliefs p_ρ^* and p_α^*. The solid and dashed lines depict how posterior beliefs evolve when the whole resource is directed to *R* and no success is realized.

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 hers. 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. For a given θ, the agent prefers to stop later than the principal does and thus has incentives to misreport his type. More specifically, lower types (those types with a lower θ) have incentives to mimic higher types in order to prolong the experimentation.

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 would be a set of stopping times. This is the 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 1, then the two players' preferences are perfectly aligned. Knowing that for any prior the agent's preferred stopping time coincides with hers, the principal 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. In particular, it is unclear whether the principal would still offer a set of stopping times. In fact, as shown in online Appendix B1, there exist examples in which restricting attention to stopping-time policies is suboptimal.[11] For this reason, I am led to consider the space of all policies.

III. A Finite-Dimensional Characterization of the Policy Space

The space of all policies is large. In Section IIIA, for each policy, I associate to it a pair of numbers, called the *total expected discounted resource* (hereafter, *expected resource*) pair, and show that this pair is a sufficient statistic for this policy in terms of both players' payoffs. In Section IIIB, I solve for the set of feasible expected resource pairs, which is a subset of $\mathbb{R}_+^2$ and can be treated as the space of all policies.

This transformation allows me to reduce the dynamic problem to a static one. In Section IIIC, I characterize players' preferences over the feasible pairs and reformulate the delegation problem.

A. *A Policy as a Pair of Numbers*

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

$$(3)\quad W^1(\pi) \equiv E\left[\int_0^\infty re^{-rt}\pi_t\,dt \,|\, \pi, 1\right], \quad W^0(\pi) \equiv E\left[\int_0^\infty re^{-rt}\pi_t\,dt \,|\, \pi, 0\right].$$

The term $W^1(\pi)$ measures the expected discounted sum of the resource allocated to R under π in state 1. I refer to $W^1(\pi)$ as the *expected resource* allocated to R under π in state 1.[12] Similarly, the term $W^0(\pi)$ is the *expected resource* allocated to R under π in state 0. Both $W^1(\pi)$ and $W^0(\pi)$ are in $[0, 1]$ because π takes values in $[0, 1]$. Therefore, (W^1, W^0) defines a mapping from the policy space Π to $[0, 1]^2$.

A policy's resource pair is important because all payoffs from implementing this policy can be written in terms of its resource pair. Conditional on the state, the payoff rate of R is known, so the payoff from implementing this policy depends only on how the resource is allocated between R and S. For instance, conditional on state 1, $(\lambda^1 h_i, s_i)$ are, respectively, the payoff rates of R and S to player i; $(W^1(\pi), 1 - W^1(\pi))$ are the resources allocated to R and S, respectively, in expectation. Their product is

[11] It is shown in online Appendix B1 that, with two types, the low-type policy is a stopping-time one, whereas the high-type policy might be a slack-after-success or delay policy for certain parameters. In these cases, the principal is strictly better off by offering policies other than stopping-time ones.

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

the payoff conditional on state 1.[13] Multiplying the payoffs conditional on the states by the initial state distribution gives the payoff of this policy. I summarize the analysis above 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 follows:*

$$U_i(\pi, p) = p\big(\lambda^1 h_i - s_i\big)W^1(\pi) + (1-p)\big(\lambda^0 h_i - s_i\big)W^0(\pi) + s_i. \tag{4}$$

Lemma 1 shows that $\big(W^1(\pi), W^0(\pi)\big)$ is a sufficient statistic for policy π for the payoffs. Instead of working with a generic policy π, I focus on $\big(W^1(\pi), W^0(\pi)\big)$ without loss of generality.

B. *Feasible Set*

Let Γ denote the image of the mapping $\big(W^1, W^0\big) : \Pi \to [0, 1]^2$. The set Γ, referred to as the *feasible set*, contains all resource pairs that can be achieved by some policy. The feasible set is convex, since the policy space Π is convexified. Therefore, to characterize the feasible set, I need to characterize only its extreme points. The extreme point $\big(w^1, w^0\big) \in \Gamma$ in the direction $(p_1, p_2) \in \mathbb{R}^2$, $\|(p_1, p_2)\| = 1$ can be found by solving for the policy that maximizes $p_1 W^1(\pi) + p_2 W^0(\pi)$. That is,

$$\max_{\pi \in \Pi} p_1 W^1(\pi) + p_2 W^0(\pi). \tag{5}$$

Compared with (4), this maximand can be interpreted as the expected payoff of a single player i whose prior belief of state 1 is $|p_1|/(|p_1| + |p_2|)$ and whose payoff parameters are such that $\lambda^1 h_i - s_i = \text{sgn}(p_1)$, $\lambda^0 h_i - s_i = \text{sgn}(p_2)$. Here $\text{sgn}(\cdot)$ is the sign function. This transforms the problem into a single player's optimization problem. Therefore, Markov policies are sufficient.[14] The following lemma shows that the feasible set is the convex hull of the resource pairs of Markov policies.

LEMMA 2 (Feasible Set): *The feasible set is the convex hull of* $\big\{(W^1(\pi), W^0(\pi)) : \pi \in \Pi^M\big\}$, *where* Π^M *is the set of Markov policies with respect to the posterior belief of state* 1.

Here, I characterize the feasible set of the benchmark setting by finding the extreme points in all directions. In Section VA and online Appendix B3, I characterize the feasible sets of more general stochastic processes.[15]

[13] Recall 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.

[14] An example of non-Markov policies is to let the agent allocate the whole resource to R until the second success and then switch the whole resource to S. Under this policy, the posterior jumps to 1 after the first success, yet the agent takes different actions before and after the second success.

[15] In online Appendix B3, I extend the analysis to Lévy bandits as in Cohen and Solan (2013) and characterize the feasible set when the risky task's payoff is one of the two Lévy processes which have independent and stationary increments. Both the Poisson bandits (Keller, Rady, and Cripps 2005; Keller and Rady 2010) and the Brownian bandits (Bolton and Harris 1999) are special cases of the class of Lévy bandits.

If $p_1 \geq 0, p_2 \geq 0$, the maximum in (5) is achieved by the policy which directs the whole resource to R. The corresponding resource pair is $(1, 1)$. Similarly, if $p_1 \leq 0, p_2 \leq 0$, the maximum is achieved by the policy which directs the whole resource to S. The corresponding resource pair is $(0, 0)$.

If $p_1 > 0, p_2 < 0$, the maximum is achieved by a lower-cutoff Markov policy which directs the whole resource to R if the posterior belief is above a certain cutoff $p^* \in (0, 1)$, and to S if it is below that cutoff:

$$\pi(p_t) = \begin{cases} 1, & \text{if } p_t \geq p^*, \\ 0, & \text{if } p_t < p^*. \end{cases}$$

This is effectively a stopping-time policy. Therefore, the extreme points in directions $p_1 > 0, p_2 < 0$ are characterized by the class of stopping-time policies. Let Γ^{lc} denote the set of resource pairs generated by the lower-cutoff policies defined above (equivalently, stopping-time policies). It is easy to verify the following:

$$\Gamma^{lc} = \left\{ \left(w^1, w^0\right) \middle| w^0 = 1 - \left(1 - w^1\right)^{\frac{r}{r+\lambda^1}}, w^1 \in [0, 1] \right\}.$$

Note that both $(0, 0)$ and $(1, 1)$ are in this set, corresponding to the stopping times at zero and infinity, respectively.

If $p_1 < 0, p_2 > 0$, the maximum is achieved by an upper-cutoff Markov policy which directs the whole resource to R if the posterior belief is below a certain cutoff $p^{**} \in (0, 1)$ and to S if it is above that cutoff:

$$\pi(p_t) = \begin{cases} 1, & \text{if } p_t \leq p^{**}, \\ 0, & \text{if } p_t > p^{**}. \end{cases}$$

Since the posterior belief jumps to 1 upon the first success, this upper-cutoff Markov policy takes one of two possible forms. If the prior belief $|p_1|/(|p_1| + |p_2|)$ is higher than the cutoff p^{**}, this policy corresponds to the policy which directs the whole resource to S. If the prior belief $|p_1|/(|p_1| + |p_2|)$ is lower than the cutoff p^{**}, this policy corresponds to the policy which directs the whole resource to R before the first success and the whole resource to S after the first success. Let Γ^{uc} denote the set of the resource pair under this policy:

$$\Gamma^{uc} = \left\{ \left(w^1, w^0\right) \middle| w^0 = 1, w^1 = \frac{r}{r + \lambda^1} \right\}.$$

This is the only extreme point in addition to Γ^{lc} that we need to consider. I have now found the extreme points in all directions. The next lemma states that the feasible set is the convex hull of the set of all extreme points.

LEMMA 3 (Conclusive News—Feasible Set): *The feasible set is* $\text{co}\left\{\Gamma^{lc} \cup \Gamma^{uc}\right\}$, *the convex hull of* Γ^{lc} *and* Γ^{uc}.

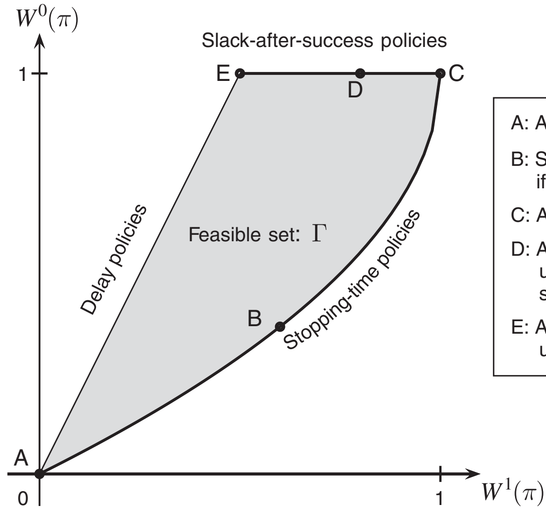

Figure 2. Feasible Set and Example Policies $\left(r/\lambda^1 = 1\right)$

Figure 2 depicts the feasible set when r/λ^1 equals 1. The $\left(w^1, w^0\right)$ pairs on the southeast boundary correspond to stopping-time policies. Point E corresponds to the unique resource pair in Γ^{uc}. The points between C and E can be generated by what I call *slack-after-success policies*: allocate the whole resource to R until the first success; then allocate only a fixed fraction to R after the first success. The points between A and E can be generated by what I call *delay policies*: allocate the whole resource to S for a fixed period, then allocate the whole resource to R until the first success, and then switch the whole resource back to S forever. (Hence, delay policies involve not only delay in experimentation but also no exploitation of success at all.) From now on, I also refer to the union of the slack-after-success and delay policies as the northwest boundary of Γ.

For future reference, a $\left(w^1, w^0\right) \in \Gamma$ pair is also called a *bundle*. I write the feasible set as $\Gamma = \left\{\left(w^1, w^0\right) \mid \beta^{se}\left(w^1\right) \leq w^0 \leq \beta^{nw}\left(w^1\right), w^1 \in [0, 1]\right\}$, where β^{se}, β^{nw} are functions from $[0, 1]$ to $[0, 1]$, characterizing the southeast and northwest boundaries of the feasible set, as follows:

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

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

Note that the shape of the feasible set depends only on the ratio r/λ^1. As demonstrated in Figure 3, the feasible set expands as r/λ^1 decreases. Intuitively, if future payoffs are discounted to a lesser extent, a player has more time to learn about the state. As a result, he is better able to direct resources to R in one state while avoiding wasting resources on R in the other state.

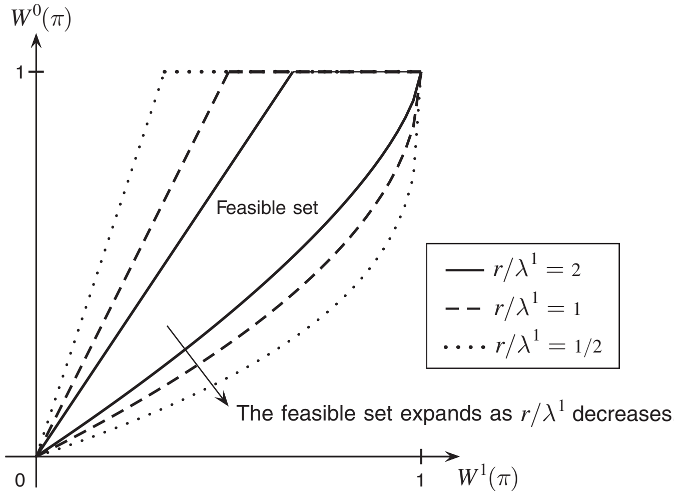

FIGURE 3. FEASIBLE SETS AS r/λ^1 VARIES

C. *Delegation Problem Reformulated*

According to Lemma 1, player i's payoff, for a fixed policy π and a fixed odds ratio θ, is

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

Therefore, player i's preferences over $(w^1, w^0) \in \Gamma$ are characterized by upward-sloping indifference curves, with slope $\theta \cdot \eta_i$. The more confident a player is of state 1, or the more he values R's successes over S's flow payoffs, the steeper his indifference curves are. Player i's preferred bundle given θ, denoted by $\left(w_i^1(\theta), w_i^0(\theta)\right)$, is the point at which his indifference curve is tangent to the southeast boundary of Γ (see Figure 4).

Based on Lemma 1 and 3, I reformulate the delegation problem by replacing Π with Γ. The principal offers a direct mechanism $\left(w^1, w^0\right) : \Theta \to \Gamma$, called a contract, such that

$$\max_{w^1, w^0} \int_\Theta \left(\frac{\theta}{1+\theta}\,\eta_\rho w^1(\theta) - \frac{1}{1+\theta}\,w^0(\theta)\right) dF(\theta), \tag{6}$$

$$\text{subject to} \quad \theta\eta_\alpha w^1(\theta) - w^0(\theta) \geq \theta\eta_\alpha w^1(\theta') - w^0(\theta'), \quad \forall \theta, \theta' \in \Theta. \tag{7}$$

The IC constraint (7) ensures that the agent reports his type truthfully. The data relevant to this problem include: (i) two payoff parameters η_α, η_ρ; (ii) the feasible set parametrized by r and λ^1; and (iii) the type distribution F. The solution to this problem, called the optimal contract, is denoted by $\left(w^{1*}(\theta), w^{0*}(\theta)\right)$.[16]

[16] Since both players' payoffs are linear in $\left(w^1, w^0\right)$, the optimal mechanism is deterministic.

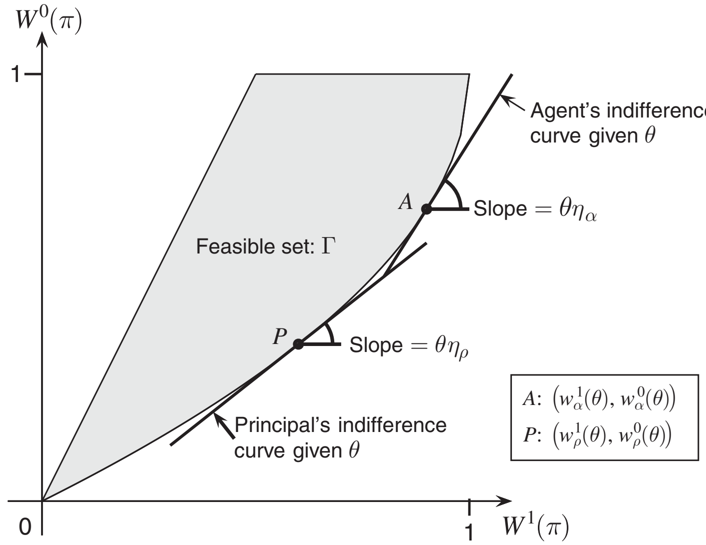

Figure 4. Indifference Curves and Preferred Bundles

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

IV. Main Results

Given a direct mechanism $\left(w^1(\theta),\, w^0(\theta)\right)$, let $U_\alpha(\theta)$ denote the payoff that the agent of type θ obtains by maximizing over his report. Since 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

$$\theta\eta_\alpha w^1(\theta) - w^0(\theta) = \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1\left(\tilde{\theta}\right) d\tilde{\theta} + \underline{\theta}\eta_\alpha w^1(\underline{\theta}) - w^0(\underline{\theta}). \tag{8}$$

The incentive compatibility of $\left(w^1,\, w^0\right)$ also requires w^1 to be a nondecreasing function of θ: higher types are willing to spend more resource on R in state 0 for a given increase in w^1 than lower types. Thus, condition (8) and the monotonicity of w^1 are necessary for incentive compatibility. As is standard, these two conditions are also sufficient.

Substituting (8) into both (6) and the feasibility constraint, and integrating by parts allows me to eliminate $w^0(\theta)$ from the problem except for 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 $w^1 : \Theta \to [0,\, 1]$ and a scalar $\underline{w}^0$ that solves the following:

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

subject to

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

$$\beta^{nw}\left(w^1(\theta)\right) \geq \theta\eta_\alpha w^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1\left(\tilde{\theta}\right) d\tilde{\theta} - \underline{\theta}\eta_\alpha w^1(\underline{\theta}) + \underline{w}^0, \quad \forall\, \theta \in \Theta, \tag{10}$$

where $\Phi \equiv \left\{w^1, \underline{w}^0 | w^1 : \Theta \to [0, 1], w^1 \text{ is nondecreasing}, \underline{w}^0 \in [0, 1]\right\}$, and

$$G(\theta) = \left(1 - \frac{H(\theta)}{H\left(\overline{\theta}\right)}\right) + \left(\frac{\eta_\rho}{\eta_\alpha} - 1\right) \frac{\theta h(\theta)}{H\left(\overline{\theta}\right)},$$

where $h(\theta) = \dfrac{f(\theta)}{1+\theta}$, $H(\theta) = \int_{\underline{\theta}}^{\theta} h\left(\tilde{\theta}\right) d\tilde{\theta}$.

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

I denote this problem by $\mathcal{P}$. The set Φ is convex and includes the monotonicity constraint. A contract is admissible if $\left(w^1, \underline{w}^0\right) \in \Phi$, and (9) and (10) are satisfied.[17]

A. *A Robust Result: Pooling at the Top*

I first show that types above some threshold are offered the same $\left(w^1, w^0\right)$ bundle. Intuitively, types at the very top prefer to experiment more than the principal would like 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 $\overline{\theta}$, so $G\left(\overline{\theta}\right)$ is negative. The principal finds it optimal to pool those types at the very top.

Let θ_p be the lowest value in Θ such that

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

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

PROPOSITION 1 (Pooling at the 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* (9) *or* (10) *to hold with equality at* θ_p.

To understand the definition of θ_p, consider an auxiliary delegation problem in which only bundles on the southeast boundary of Γ are allowed. This restriction

[17] The problem $\mathcal{P}$ is convex because the objective function is linear in $\left(w^1, \underline{w}^0\right)$ and the constraint set is convex.

reduces the action space to a one-dimensional line segment and allows me to compare the definition here with those in the related literature, especially Alonso and Matouschek (2008). I identify a bundle on the southeast boundary with the slope of the tangent line at that point. The set of possible slopes is denoted by $Y = [(\beta^{se})'(0), \infty]$. Since there is a one-to-one mapping between the southeast boundary of Γ and Y, I let Y be the action space and refer to $y \in Y$ as an action. For a given type θ, player i's preferred action is denoted by $y_i(\theta) = \eta_i \theta$. The principal's preferred action if she believes that the agent's type is above θ is

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

Following Alonso and Matouschek (2008), I define the *forward bias* as

$$S(\theta) \equiv \left(1 - \frac{H(\theta)}{H(\overline{\theta})}\right)\left(\eta_\alpha \theta - \eta_\rho \frac{\int_\theta^{\overline{\theta}} \tilde{\theta} h(\tilde{\theta})\, d\tilde{\theta}}{H(\overline{\theta}) - H(\theta)}\right);$$

this measures the difference between the agent's preferred decision given θ and the principal's preferred decision if she believes that the agent's type is *above* θ. It is easy to verify the following equality:

$$S(\theta) = -\eta_\alpha \int_\theta^{\overline{\theta}} G(\tilde{\theta})\, d\tilde{\theta}.$$

Therefore, θ_p is the lowest value in Θ such that for any $\hat{\theta} \geq \theta_p$, the forward bias $S(\hat{\theta})$ is positive. If θ_p is interior, the forward bias at θ_p must equal zero. The principal finds it optimal to pool the top types until θ_p.

For any $\theta \geq \overline{\theta}\eta_\rho/\eta_\alpha$, the agent's preferred action $y_\alpha(\theta)$ is greater than $\overline{\theta}\eta_\rho$, which is the principal's preferred action if she believes that the agent's type is equal to $\overline{\theta}$. Therefore, for any $\theta \geq \overline{\theta}\eta_\rho/\eta_\alpha$, the forward bias is positive. It thus follows that the pooling threshold θ_p has to be below $\overline{\theta}\eta_\rho/\eta_\alpha$. I summarize this observation in the following corollary.

COROLLARY 1: *The threshold of the top pooling segment θ_p is below $\overline{\theta}\eta_\rho/\eta_\alpha$.*

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 $\overline{\theta}$. 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, $\theta_p = \overline{\theta}$ if $\eta_\rho/\eta_\alpha = 1$, and there exists $z^* \in [0, 1)$ such that $\theta_p = \underline{\theta}$ if $\eta_\rho/\eta_\alpha \leq z^*$.*

If $\theta_p = \underline{\theta}$, all types are pooled. The optimal contract consists of the principal's preferred bundle given her prior belief. From now on, I focus on the more interesting case in which $\theta_p > \underline{\theta}$.

B. *Imposing a Cutoff*

To make progress, I assume for the rest of this section and Section VA that the type distribution satisfies Assumption 1. In online Appendix B2, I examine how results change when Assumption 1 fails.[18]

ASSUMPTION 1: *For all* $\theta \leq \theta_p$, $1 - G(\theta)$ *is nondecreasing.*

When the density function f is differentiable, Assumption 1 is equivalent to the following condition:

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

This condition is satisfied for all density functions that are nondecreasing, as well as for the exponential, the log-normal, the Pareto, and the Gamma distributions for a subset of their parameters. It is also 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 2, shows that under Assumption 1 the optimal contract takes a quite simple form. To describe it formally, I introduce the following definition:

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

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

Under the cutoff rule, types below θ_p are offered their preferred bundles $\left(w_\alpha^1(\theta), w_\alpha^0(\theta)\right)$ whereas types above θ_p are pooled at $\left(w_\alpha^1(\theta_p), w_\alpha^0(\theta_p)\right)$. I denote the cutoff rule by $\left(w_{\theta_p}^1, \underline{w}_{\theta_p}^0\right)$. Figure 5 shows the delegation set corresponding to the cutoff rule. 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 $\overline{\theta}\eta_\rho$ while the agent's ranges from $\underline{\theta}\eta_\alpha$ to $\overline{\theta}\eta_\alpha$. The delegation set is the interval between $\underline{\theta}\eta_\alpha$ and $\theta_p\eta_\alpha$. According to Corollary 1, the upper bound of the delegation set $\theta_p\eta_\alpha$ is smaller than $\overline{\theta}\eta_\rho$, the principal's preferred bundle given $\overline{\theta}$.

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

[18] In online Appendix B2, I show that Assumption 1 is necessary for a cutoff rule to be optimal.

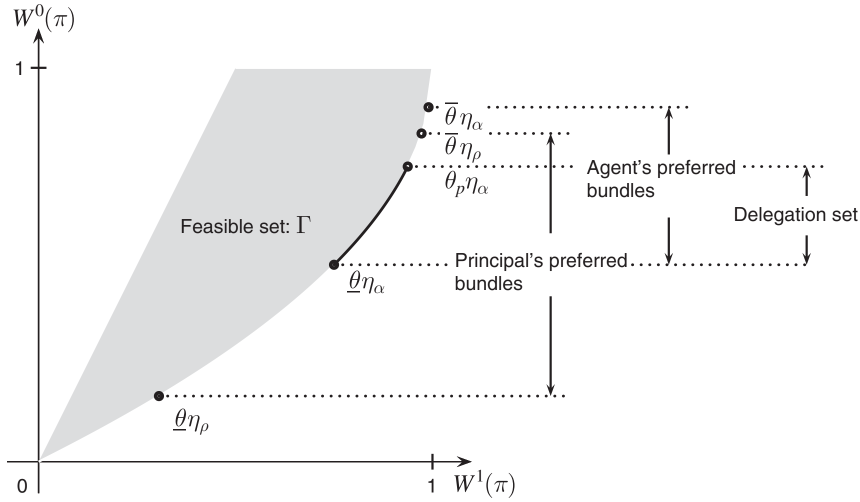

Figure 5. Delegation Set under Cutoff Rule

Notes: Parameters are $\eta_\alpha = 3/5$, $\eta_\rho = 3/5$, $r/\lambda^1 = 1$, $\underline{\theta} = 1$, $\overline{\theta} = 5$ with θ being uniformly distributed. The pooling threshold is $\theta_p \approx 1.99$.

To better interpret Assumption 1, I revisit the auxiliary delegation problem in which only bundles on Γ^{se} are allowed. Following Alonso and Matouschek (2008), I define the *backward bias* for a given type θ as

$$T(\theta) \equiv \frac{H(\theta)}{H(\overline{\theta})}\left(\eta_\alpha \theta - \eta_\rho \frac{\int_{\underline{\theta}}^{\theta} \tilde{\theta} h(\tilde{\theta})\, d\tilde{\theta}}{H(\theta)}\right), \tag{11}$$

which 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. It is easy to verify the following equality:

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

Therefore, Assumption 1 is equivalent to requiring that the backward bias be convex when $\theta \leq \theta_p$. This is the condition that Alonso and Matouschek (2008) find for the interval delegation set to be optimal in their setting. Intuitively, Assumption 1 requires that the type distribution cannot be too skewed toward lower types. If some lower types were very probable, the principal could remove actions from the interval delegation set, force these types to choose bundles with less experimentation, and benefit overall because of the concentration of the lower types. Assumption 1 ensures that the density is sufficiently greater for higher types, so that the principal finds it optimal to fill in the "holes" in the delegate set.[19]

[19] The discussion so far also suggests that Assumption 1 is necessary for a cutoff rule to be optimal. In online Appendix B2, I show that no x_p-cutoff contract is optimal for any $x_p \in \Theta$ if Assumption 1 does not hold. The

It is important to emphasize that this is not a proof of the optimality of the cutoff rule, because considering only the bundles on the southeast boundary might be restrictive. For example, with two types, the low-type policy is a stopping-time one. Yet, the high-type policy might be a slack-after-success or delay policy for certain parameters. The high-type policy involves excessive experimentation, constrained exploitation of success, and possibly delay in experimentation, all of which deter a low type from mimicking the high type. The principal gains from effectively shortening the low type's experimentation (see online Appendix B1). With a continuum of types, there also exist examples such that the principal is strictly better off by offering policies other than stopping-time ones. By using Lagrangian methods, I prove that the cutoff rule is indeed optimal under Assumption 1.

To prove Proposition 2, I borrow numerous ideas from Amador, Werning, and Angeletos (2006) and Amador and Bagwell (2013). The problem $\mathcal{P}$ is similar to that in Amador, Werning, and Angeletos (2006). Amador and Bagwell (2013) extend the methods in Amador, Werning, and Angeletos (2006) to prove the optimality of interval delegation in more general settings, with and without money burning. Conceptually, problem $\mathcal{P}$ can be thought of as a one-dimensional delegation problem with money burning: moving away from the southeast boundary reduces both players' payoffs, which can be interpreted as money burning. However, my approach is not an application of Amador and Bagwell (2013), as the current payoff is not written as a special case of theirs. Nonetheless, the reason that the optimal contract involves only bundles on the southeast boundary is similar to that in Amador and Bagwell (2013): when the principal's payoff function is at most as concave as the agent's, the sufficient conditions for interval delegation to be optimal with and without money burning coincide.

The detailed proof is relegated to the Appendix. In Section IVC, I first introduce an indirect implementation of the cutoff rule and then prove that this indirect mechanism is time consistent.

C. *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 $p(\theta_p)$ 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 $\dot{p}_t = -\lambda^1 \pi_t p_t (1 - p_t)$, where π_t is the agent's action. Upon the first success this belief jumps to 1.

The principal imposes a cutoff at p^*_α. As long as the constructed belief stays above the cutoff, the agent can decide how to allocate the resource. 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 it does constrain those types above θ_p, who are forced by the principal to stop by the principal.

x_p-cutoff contract is defined as $\left(w^1(\theta), w^0(\theta)\right) = \left(w^1_\alpha(\theta), w^0_\alpha(\theta)\right)$ for $\theta < x_p$, and $\left(w^1(\theta), w^0(\theta)\right) = \left(w^1_\alpha(x_p), w^0_\alpha(x_p)\right)$ for $\theta \geq x_p$.

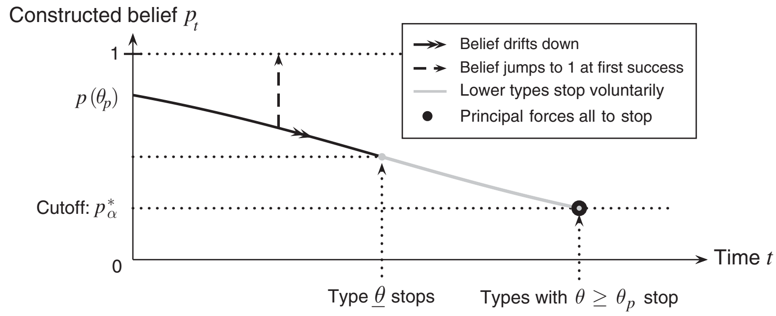

FIGURE 6. IMPLEMENTING THE CUTOFF RULE

Figure 6 illustrates how the constructed belief evolves over time. The solid arrow shows that the belief is downgraded when the whole resource is directed to *R* and no success has been realized. The dashed arrow shows that the belief jumps to 1 at the first success. The gray area shows that those types below θ_p stop voluntarily when their posterior beliefs drop to the agent's cutoff p^*_α. As illustrated by the black dot, a mass of higher types is 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 $p(\theta_p \eta_\alpha / \eta_\rho)$, in which case the principal imposes a cutoff at p^*_ρ. What matters is that the prior belief and the cutoff are chosen simultaneously to ensure that exactly those types below θ_p are given the freedom to decide when to switch to *S*.

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 explore 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 offer a new contract at any time, would she benefit from doing so? As my next result shows, the cutoff rule is indeed time consistent.

PROPOSITION 3 (Time Consistency): *If Assumption* 1 *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; and (iii) no success has occurred and the agent has not stopped. Clearly, the principal has no incentive 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 direct the whole resource to *R* 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 has already overexperimented. Hence, she has no incentive to ask the agent to use *R* any more.

What remains to be shown is that the principal finds the cutoff set at time 0 to be optimal if the agent has not stopped experimenting and there is no success. To clarify the intuition, it is useful to note that the cutoff rule separates the type space into two segments: low types that are not constrained by the cutoff, and high types that are constrained. The cutoff rule does not distort the low types' behaviors at all since they implement their preferred policies. Hence, by setting a cutoff, the principal gets no benefit from these low types who are not constrained by the cutoff. Given that the cutoff has no effect on their behaviors, the principal chooses the cutoff by looking at only those high types who will be constrained by the cutoff. The principal chooses θ_p such that she finds it optimal to stop experimenting when the cutoff becomes binding. This explains why commitment is not required even if the principal learns more about the agent's type as time passes.

I now sketch how to prove time consistency. First, I calculate the principal's updated belief about the type distribution, given that no success has occurred and that the agent has not stopped. By continuing to experiment, 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. The detailed proof is relegated to the Appendix.

Over- and Underexperimentation.—The result of Proposition 2 can also be represented by a delegation rule mapping types into stopping times since 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 $\overline{\theta}$. The dotted line represents the agent's preferred stopping time and the dashed line represents the principal's. 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 $\left[\theta_p, \overline{\theta}\right]$ where the stopping time is independent of the agent's report (i.e., the pooling segment). To implement the rule, the principal simply imposes a deadline at $\tau_\alpha(\theta_p)$.

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 perspective, 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.

V. Discussion

A. *More General Stochastic Processes*

In the benchmark setting, the risky task generates no successes in state 0, so the first success reveals that the state is 1. Here, I examine the case in which $0 < \lambda^0 < \lambda^1$. The risky task generates successes in state 0 as well, so players never fully learn the state. This happens, for instance, when a bad project generates successes but at a lower intensity. Successes are indicative of a good project but are not conclusive proof of one. The analysis is readily extended to incorporate this setting. I first show how to reduce the policy space to a two-dimensional feasible set. Then I prove the optimality of the cutoff rule and its time-consistent property. In online Appendix B3, I extend the analysis to the general class of Lévy bandits (Cohen and Solan 2013): the risky task's payoff is one of the two Lévy processes which have independent and stationary increments. Both Poisson bandits and Brownian bandits are special cases

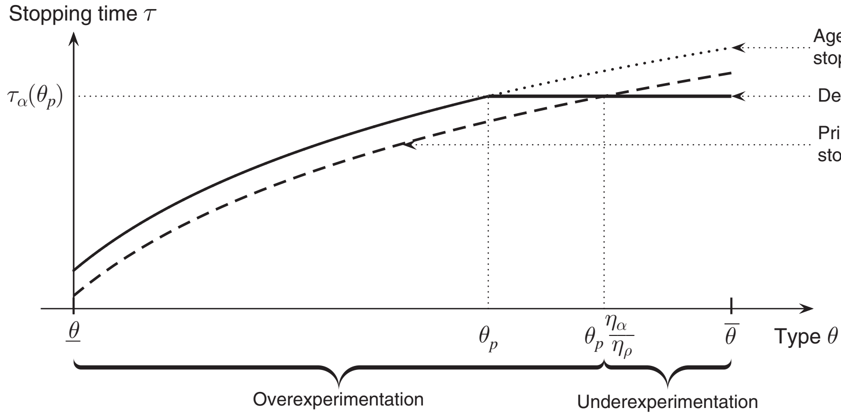

Figure 7. Equilibrium Stopping Times

Note: Parameters are $\eta_\alpha = 6/5, \eta_\rho = 1, r/\lambda^1 = 1, \underline{\theta} = 1, \overline{\theta} = 5$ with θ being uniformly distributed.

of the class of Lévy bandits, so the main results also apply to Brownian bandits as seen, for instance, in Bolton and Harris (1999).

For the rest of this subsection, I use the same notation as in the benchmark setting. Recall that $\left(W^1(\pi), W^0(\pi)\right)$ denotes the *expected resource* allocated to *R* under π, conditional on state 1 and state 0. The feasible set Γ is the image of the mapping $\left(W^1, W^0\right) : \Pi \to [0, 1]^2$ and is convex. As in the benchmark setting, the extreme point $\left(w^1, w^0\right) \in \Gamma$ in the direction $(p_1, p_2) \in \mathbb{R}^2, \| (p_1, p_2) \| = 1$ can be found by solving for the policy π that maximizes $p_1 W^1(\pi) + p_2 W^0(\pi)$. This maximand can be interpreted as a single player's expected payoff (with properly adjusted prior belief and payoff parameters). Therefore, it suffices to focus on lower-cutoff and upper-cutoff Markov policies since a single player's optimal policy must be either a lower-cutoff or an upper-cutoff policy. The following lemma generalizes Lemma 3 to the inconclusive success case.

LEMMA 4 (Inconclusive News—Feasible Set): *There exist two functions* $\beta^{se}, \beta^{nw} : [0, 1] \to [0, 1]$ *such that* $\Gamma = \left\{\left(w^1, w^0\right) \mid \beta^{se}\left(w^1\right) \leq w^0 \leq \beta^{nw}\left(w^1\right), w^1 \in [0, 1]\right\}$. *The southeast boundary is given by* $\beta^{se}\left(w^1\right) = 1 - \left(1 - w^1\right)^{a^*/(1+a^*)}$, *for some constant* $a^* > 0$. *The northwest boundary* β^{nw} *is concave, nondecreasing, and once continuously differentiable, with end points* $(0, 0)$ *and* $(1, 1)$.

Figure 8 depicts the feasible set with inconclusive successes.[20] In the benchmark setting, the northwest boundary is piecewise linear. Recall that the northwest boundary consists of upper-cutoff Markov policies that direct the resource to *R* only when the belief is below a certain cutoff. When successes are conclusive, the posterior belief drifts downward without a success and gets absorbed at 1 after the

[20]The functional form for β^{se} when $\lambda_0 > 0$ is the same as that in the benchmark setting $\lambda_0 = 0$. This is because the payoff of a single player who wants to direct the resource to *R* only in state 1 takes the same functional form in both the conclusive- and inconclusive-success cases.

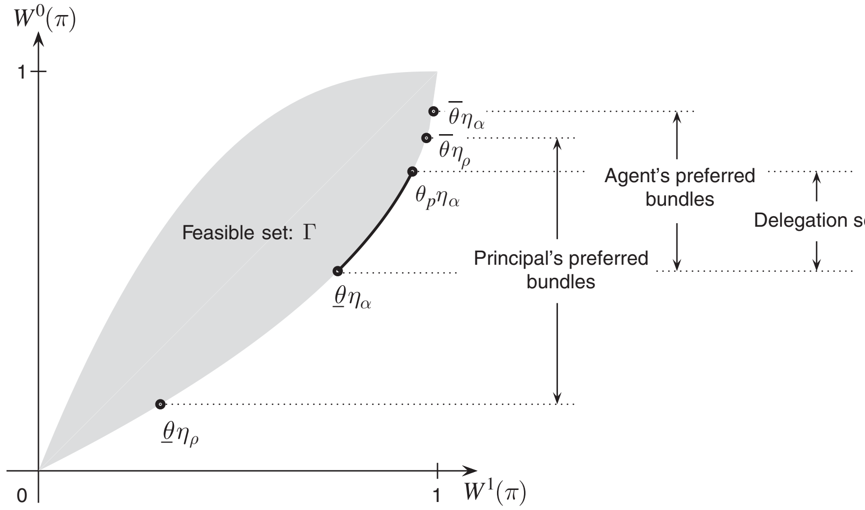

FIGURE 8. DELEGATION SET UNDER CUTOFF RULE: INCONCLUSIVE NEWS CASE

Notes: Parameters are $r = 1/5$, $\lambda^1 = 2/5$, $\lambda^0 = \left(2 - \sqrt{2}\right)/5$. All other parameters are the same as Figure 5.

first success. Therefore, the set of the upper-cutoff policies contains essentially one resource pair, which directs the whole resource to R before the first success and to S after it. In the general setting, the northwestern boundary is differentiable, since the set of the upper-cutoff policies contains a range of resource pairs as the cutoff or the prior belief is adjusted. Moreover, for a fixed cutoff belief, the single player's payoff is concave and differentiable in his prior belief for any prior that is below the cutoff.

My next result shows that the cutoff rule as defined in Definition 1 is optimal under Assumption 1. This is the case because the southeast boundary here takes the same functional form as in the benchmark setting. The proof of Proposition 2, which relies only on the properties of the southeast boundary, applies directly to the current setting.

PROPOSITION 4 (Inconclusive News—Sufficiency): *If Assumption* 1 *holds*, *the cutoff rule is optimal when* $0 < \lambda^0 < \lambda^1$.

The implementation of the cutoff rule, similar to its implementation in the benchmark setting, is achieved by calibrating a constructed belief which starts with the prior belief $p(\theta_p)$. This belief is updated as follows: (i) if no success occurs, the belief drifts downward according to the differential equation $\dot{p}_t = -\left(\lambda^1 - \lambda^0\right)\pi_t p_t(1 - p_t)$; 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_α^*. If this belief stays above the cutoff, the agent freely decides how to allocate the resource, but he is required to allocate the whole resource to S when it drops to the cutoff. The gray areas in Figure 9 show that lower

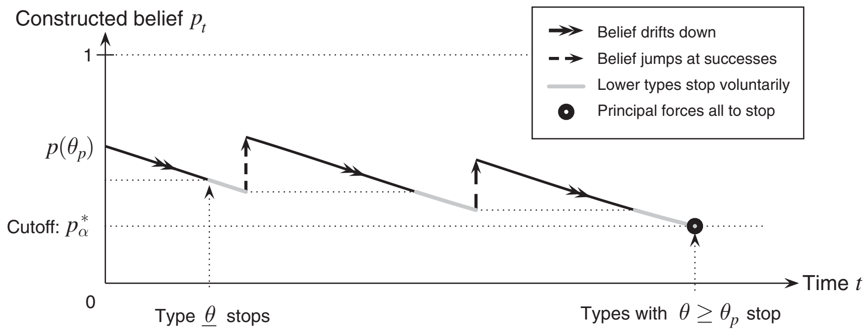

Figure 9. Implementing the Cutoff Rule: Inconclusive News

types stop voluntarily when 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, in which the belief jumps to 1 after the first success, and the inconclusive news setting, in which the belief jumps up upon successes and then drifts downward. 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. Whenever a success is realized, more time is granted. The agent is also allowed to give up his time voluntarily. However, after a long stretch of time without success, the agent must switch to *S*.

The cutoff rule is time consistent when successes are inconclusive. The intuition is the same as in the benchmark setting: since the cutoff creates no distortion on low types' behaviors, the principal chooses the optimal cutoff conditional on the agent being constrained by the cutoff. The proof is slightly different than for the benchmark setting since successes never fully reveal the state. I need to show that the principal finds it optimal not to adjust the cutoff upon the arrival of any success. The detailed proof is relegated to the Appendix.

PROPOSITION 5 (Inconclusive News—Time Consistency): *If Assumption* 1 *holds, the cutoff rule is time consistent when* $0 < \lambda^0 < \lambda^1$.

B. *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 (i.e., the 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 (i.e., the principal). Another possibility is that the agent does not perceive the risky task to be generating significant career opportunities compared with alternative activities. In both cases, the agent prefers to stop experimenting earlier. I derive the optimal delegation rule. The results here apply to both conclusive and inconclusive successes.

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 benchmark setting, I first show in Proposition 6 that types below some threshold are pooled. Then I show in Proposition 7 that under a regularity condition, the optimal contract is the *reversed-cutoff rule* (as defined in Definition 2): There exists a threshold type such that those types who are more optimistic than the threshold type are offered their preferred policies, whereas those types who are less optimistic are pooled at the threshold type's preferred policy.

Formally, let θ_p^s be the highest value in Θ such that

$$\int_{\underline{\theta}}^{\hat{\theta}} (1 - G(\theta))\, d\theta \leq 0, \quad \text{for any } \hat{\theta} \leq \theta_p^s.$$

Therefore, θ_p^s is the highest value in Θ such that for any $\hat{\theta} \leq \theta_p^s$, the backward bias $T(\hat{\theta})$ as defined in (11) is negative. If θ_p^s is less than $\overline{\theta}$, the backward bias at θ_p^s equals zero. Thus, the agent's preferred action given θ_p^s is equal to the principal's preferred action if she believes that the agent's type is below θ_p^s. The following proposition shows that types with $\theta \in [\underline{\theta}, \theta_p^s]$ 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.

PROPOSITION 6 (Pooling at the Bottom): *An optimal contract* (w^{1*}, w^{0*}) *satisfies* $(w^{1*}(\theta), w^{0*}(\theta)) = (w^{1*}(\theta_p^s), w^{0*}(\theta_p^s))$ *for* $\theta \leq \theta_p^s$. *It is optimal for* $(w^{1*}(\theta_p^s), w^{0*}(\theta_p^s))$ *to be on the boundary of the feasible set.*

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

DEFINITION 2: *The reversed-cutoff rule is the contract* (w^1, w^0) *such that*

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

Under the reversed-cutoff rule, types below θ_p^s are pooled at θ_p^s's preferred bundle, and types above θ_p^s are assigned their preferred bundles. Figure 10 illustrates the delegation set corresponding to the reversed-cutoff rule. The next result gives a sufficient condition under which the reversed-cutoff rule is optimal.

PROPOSITION 7 (Sufficiency—The Reversed-Cutoff Rule): *The reversed-cutoff rule* $(w_{\theta_p^s}^1, \overline{w}_{\theta_p^s}^0)$ *is optimal, if* $1 - G(\theta)$ *is nondecreasing when* $\theta \geq \theta_p^s$.

Similar to the previous results, interval delegation is optimal when the backward bias is convex for $\theta \geq \theta_p^s$. If $f(\theta)$ is differentiable, $1 - G(\theta)$ being nondecreasing for $\theta \geq \theta_p^s$ is equivalent to requiring that

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

This condition is satisfied if the type distribution is not too skewed toward higher types. For any density f with $\theta f'/f$ bounded from above, the condition is satisfied

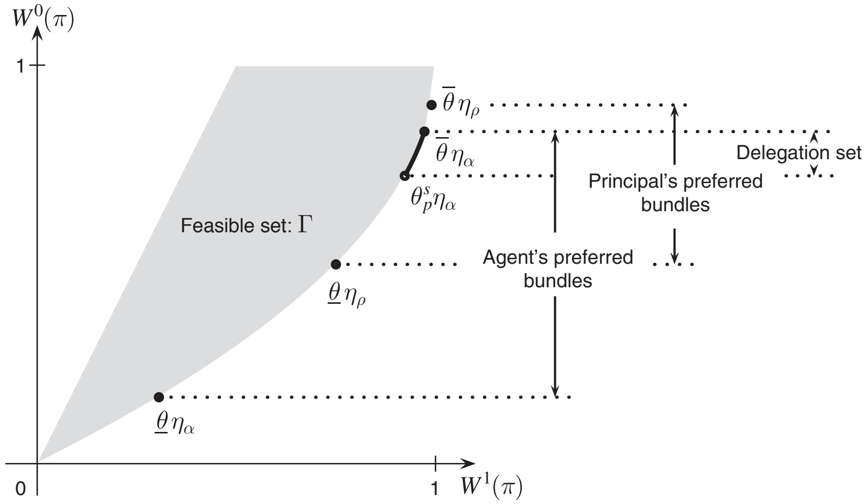

Figure 10. Delegation Set under Lockup Rule

Notes: Parameters are $\eta_\alpha = 3/5$, $\eta_\rho = 1$, $r/\lambda^1 = 1$, $\underline{\theta} = 1$, $\overline{\theta} = 5$ with θ being uniformly distributed. The pooling threshold is $\theta_p^s \approx 3.42$.

when η_ρ/η_α is sufficiently close to 1, or equivalently, when two players' preferences are sufficiently aligned. The proof is relegated to the Appendix.

This optimal contract 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 allocate the whole resource to *R*. As soon as it drops to the cutoff, the principal keeps her hands off the project and lets the agent decide how to allocate the resource. 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. In contrast to the setting with $\eta_\alpha > \eta_\rho$, the agent here has no flexibility until the cutoff belief is reached.

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

The reversed-cutoff rule is not time consistent. Whenever those high types voluntarily stop experimenting, they perfectly reveal their types. The principal learns that these high types underexperiment compared with her optimal experimentation policy, so she is tempted to command the agent to experiment further. Therefore, there is tension on the principal's ability to commit.

C. *Optimal Contract with Transfers*

In this subsection, I examine the optimal contract when the principal can make transfers to the agent. I assume conclusive successes $\left(\text{i.e., } \lambda^0 = 0\right)$ and a larger agent's return (i.e., $\eta_\alpha > \eta_\rho$) in this discussion.

From both the theoretical and empirical points of view, it is important to know how the results change when transfers are allowed. I assume that the principal has full commitment power, that is, she can write a contract specifying both an experimentation policy π and a transfer scheme c at the outset of the game. I also assume that the agent is protected by limited liability, i.e., only nonnegative transfers from the principal to the agent are allowed. An experimentation policy π is defined in the same way as before. A transfer scheme c is a nonnegative, nondecreasing process $\{c_t\}_{t \geq 0}$, which depends on the history of events up to t, where c_t denotes the cumulative transfers the principal has made to the agent up to time t (inclusive).[21] Under perfect commitment, the revelation principle applies. A direct mechanism specifies for each type an experimentation policy and transfer scheme pair. Again, due to the large contract space, it is difficult to solve for the optimal contract directly.

In Section III, I represent an experimentation policy by a summary statistic, characterize the range of all summary statistics, and then reduce the dynamic contract problem to a static one. This method is not restricted to the delegation model but can be generalized to the case with transfers. Each policy and transfer scheme pair, in terms of payoffs, can be represented by a summary statistic of four numbers. The first two numbers are the expected resources directed to the risky task conditional on states 1 and 0. The last two numbers are the expected transfers made to the agent conditional on states 1 and 0, summarizing the transfers that the agent expects to receive conditional on the state. Given that both players are risk-neutral, their payoffs can be written as linear functions of these four numbers. The range of these summary statistics is a subset of $\mathbb{R}^4_+$, and can be considered as the new contract space. This reduces the dynamic contracting problem to a static one. Given that the problem is now static, I utilize the optimal control methods (similar to those used by Krishna and Morgan 2008) to solve for the optimal contract. The detailed analysis can be found in online Appendix B4. Here, I briefly discuss the results (under certain regularity conditions), focusing on when and how the transfers are made.

When the principal can both use monetary incentives and restrict the agent's behavior, she uses both tools. For very low types, the principal "bribes" them to experiment less than those types' preferred policies.[22] However, paying higher types to experiment less makes it more expensive to "bribe" lower types due to the agent's bias to overexperiment. Therefore, when the agent's type is sufficiently high, the principal stops using transfers. For the highest types, the principal simply caps the agent's choice. The policy chosen becomes unresponsive to the agent's information.

To understand the optimal contract with transfers, let us first consider an auxiliary problem in which every type's outside option is his preferred policy. This is referred to as a *transfer-only* problem in the sense that the principal has no authority over the policy taken and can only use transfers to affect the agent's behavior. When transfers are allowed, the efficient policy for a given θ is the one that maximizes the sum of the principal's and the agent's payoffs. Since upward IC constraints potentially bind, the principal makes the highest payment to the lowest type and assigns him

[21] Formally, the number of successes achieved up to time t (inclusive) defines the point process $\{N_t\}_{t \geq 0}$. Let $\mathcal{F} \equiv \{\mathcal{F}_t\}_{t \geq 0}$ denote the filtration generated by the process π and N_t. The process $\{c_t\}_{t \geq 0}$ is $\mathcal{F}$-adapted.

[22] If the principal asks the lowest type to experiment slightly less, the cost to the agent is approximately zero, since the lowest type was implementing his optimal policy. But the benefit to the principal is of the first order and strictly positive. Therefore, the principal always uses monetary incentives.

the efficient policy. In other words, efficiency is achieved at the bottom. For any type above $\underline{\theta}$, reducing his experimentation length improves efficiency. However, the principal has to increase all lower types' information rents. As a result, there is distortion toward the agent's preferred policy except at the lowest type. There exists a type $\theta^* \in \left(\underline{\theta}, \overline{\theta}\right)$ at which the marginal cost of the information rent is equal to the marginal benefit from marginally reducing type θ's experimentation length. The principal stops making payments to types above θ^*, letting them implement their preferred policies.

One interesting observation is that transfers are made only after the state is revealed to be 1, even though the agent is biased toward more experimentation. Intuitively, lower types have incentives to mimic higher ones. Also, lower types are less optimistic about the state being 1 than higher types are. Therefore, if transfers are made only if the state is revealed to be 1, then lower types perceive higher types' transfer schemes to be less lucrative than higher types perceive them to be. This allows the principal to save on information rents. So, if the agent is biased toward more experimentation, the agent receives positive payment when he has claimed a short period of experimentation and a success has been realized. This completes the solution to the *transfer-only* problem.

The solution to the original problem is different in the sense that the principal can cap the experimentation length of the most optimistic types. The exact contract form depends on whether θ^* is smaller or larger than the pooling threshold θ_p in the delegation contract.

When the bias term η_α/η_ρ is small, θ^* is smaller than θ_p. The principal makes transfers to types below θ^*, "bribing" them to stop experimentation earlier. The lowest type's experimentation policy is the efficient one, and he receives the highest level of transfer payment. As the type increases, the experimentation policy shifts toward the agent's preferred one, with a corresponding decrease in the transfer payments. The principal makes no transfers to types above θ^*. The optimal contract for types above θ^* coincides with the delegation solution: those types between θ^* and θ_p implement their preferred policies, and those types above θ_p are pooled. Figure 11 illustrates the equilibrium policy and transfer payments. On the left-hand side, the dotted and sparsely dashed lines correspond to the agent's and the principal's preferred policies, respectively. The densely dashed line illustrates the equilibrium policy in the transfer-only problem, whereas the solid line illustrates the equilibrium policy in the original problem (in which the principal caps types above θ_p). On the right-hand side is the equilibrium transfer payment. The equilibrium policy in the delegation problem is illustrated in Figure 7.

When the bias η_α/η_ρ is large enough, θ^* is larger than θ_p. At the optimum, there exists a type $\tilde{\theta}_p \in (\theta_p, \theta^*)$, such that the optimal contract consists of two segments, $\left[\underline{\theta}, \tilde{\theta}_p\right]$ and $\left[\tilde{\theta}_p, \overline{\theta}\right]$. Types below $\tilde{\theta}_p$ receive positive transfers and are separated. Types above $\tilde{\theta}_p$ are pooled, and implement type $\tilde{\theta}_p$'s preferred policy. There is no segment in which the agent implements his preferred bundle. Note that the principal does not extend the transfer region up to θ^*. If she pooled only the types above θ^* and let them implement type θ^*'s preferred policy, then the types above θ^* would overexperiment from the principal's perspective. This is because for any type θ above θ_p, type θ's preferred action is higher than the principal's preferred action

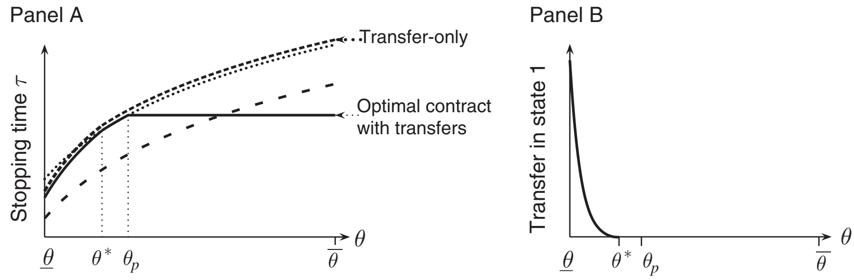

FIGURE 11. EQUILIBRIUM STOPPING TIMES AND TRANSFERS

Notes: Parameters are $\eta_\alpha = 1, \eta_\rho = 3/5, r/\lambda^1 = 1, \underline{\theta} = 1, \overline{\theta} = 5$ with θ being uniformly distributed. The thresholds are $\theta^* \approx 1.80$ and $\theta_p \approx 1.99$.

if she believes that the agent's type is above θ. Therefore, extending the positive transfer region up to θ^* imposes negative externalities on the allocation above θ^*, and thus is suboptimal. The optimal threshold $\tilde{\theta}_p$ balances the benefit of screening and the cost of overexperimentation.

VI. Concluding Remarks

In this paper, I study how to optimally manage innovation in hierarchical organizations. Specifically, I characterize how a principal optimally delegates experimentation to an informed yet biased agent, and how the agent's flexibility over resource allocation should be adjusted over time. Under a regularity condition, interval delegation is optimal. If the agent is biased toward more experimentation, he is constrained by a sliding deadline, and must achieve a success before the next deadline to keep the project alive. On the contrary, if the agent is biased toward less experimentation, he is constrained by a lockup period which extends whenever a success arrives. The agent has no freedom but to experiment during the lockup period. He freely decides how to allocate the resource after the lockup ends.

This paper suggests several promising directions for future research. It is assumed that the agent's private information is one-dimensional. I expect that the optimality of the interval delegation generalizes when the agent's private information is multiple-dimensional. (For instance, the agent's bias level is also private information.) Also, in addition to ex ante private information, the agent might acquire private information about the project's prospect during the experimentation process. I expect that under certain conditions the optimality of the (sliding) deadline/lockup rule can be extended to the setting with asymmetric learning.

Appendix

PROOF OF LEMMA 1:

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

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

PROOF OF PROPOSITION 1:

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

$$\eta_\alpha w^1(\theta_p) \int_{\theta_p}^{\overline{\theta}} G(\theta)\, d\theta + \eta_\alpha \int_{\theta_p}^{\overline{\theta}} \int_{\theta}^{\overline{\theta}} G\left(\tilde{\theta}\right) d\tilde{\theta}\, dw^1(\theta). \tag{A1}$$

The first term depends only on $w^1(\theta_p)$. The integrand of the second term, $\int_{\theta}^{\overline{\theta}} G\left(\tilde{\theta}\right) d\tilde{\theta}$, is negative for $\theta \in \left[\theta_p, \overline{\theta}\right]$. Thus, it is optimal to have $dw^1(\theta) = 0$, $\forall\, \theta \in \left[\theta_p, \overline{\theta}\right]$. If $\theta_p = \underline{\theta}$, all types are pooled, and implement the principal's preferred uninformed bundle on Γ^{se}. If $\theta_p > \underline{\theta}$, $\int_{\theta_p}^{\overline{\theta}} G(\theta)\, d\theta$ equals zero. Adjusting $w^1(\theta_p)$ does not affect (A1), so $w^1(\theta_p)$ can be increased until either (9) or (10) binds. ■

PROOF OF COROLLARIES 1 AND 2:

Substituting $G(\theta)$ and integrating by parts yields

$$\int_{\hat{\theta}}^{\overline{\theta}} G(\theta)\, d\theta = \frac{1}{H\left(\overline{\theta}\right)} \int_{\hat{\theta}}^{\overline{\theta}} \left(\frac{\eta_\rho}{\eta_\alpha}\theta - \hat{\theta}\right) h(\theta)\, d\theta.$$

For any $\hat{\theta} \in \left[\overline{\theta}\eta_\rho/\eta_\alpha, \overline{\theta}\right]$, the integrand $\left(\theta\eta_\rho/\eta_\alpha - \hat{\theta}\right)h(\theta)$ is weakly negative for any $\theta \in \Theta$. Therefore, $\theta_p \leq \overline{\theta}\eta_\rho/\eta_\alpha$.

Let $G(\theta, z)$ be the value of $G(\theta)$ if η_ρ/η_α equals $z \in [0, 1]$. Let $\theta_p(z)$ be the lowest value in Θ such that $\int_{\hat{\theta}}^{\overline{\theta}} G(\theta, z)\, d\theta \leq 0$ for any $\hat{\theta} \geq \theta_p(z)$. Because $G(\theta, z)$ increases in z for a fixed θ, $\theta_p(z)$ also increases in z. If

$\theta_p(1) < \overline{\theta}$, it holds that $\int_{\theta_p(1)}^{\overline{\theta}} G(\theta, 1)\, d\theta = \int_{\theta_p(1)}^{\overline{\theta}} (\theta - \theta_p(1))h(\theta)\, d\theta / H(\overline{\theta}) \leq 0$, and thus $h(\theta) = 0, \forall\, \theta \in \left[\theta_p(1), \overline{\theta}\right]$. A contradiction. Therefore, $\theta_p(1) = \overline{\theta}$. For any $\hat{\theta} \in \Theta$ and $z \in [0, 1]$, I have $\int_{\hat{\theta}}^{\overline{\theta}} G(\theta, z)\, d\theta = \left(z \int_{\hat{\theta}}^{\overline{\theta}} \theta h(\theta)\, d\theta - \hat{\theta} \int_{\hat{\theta}}^{\overline{\theta}} h(\theta)\, d\theta\right) / H(\overline{\theta})$. Let z^* be $\min_{\hat{\theta}\in\Theta} \left(\hat{\theta} \int_{\hat{\theta}}^{\overline{\theta}} h(\theta)\, d\theta\right) / \left(\int_{\hat{\theta}}^{\overline{\theta}} \theta h(\theta)\, d\theta\right)$. If $z \leq z^*$, $\int_{\hat{\theta}}^{\overline{\theta}} G(\theta, z)\, d\theta \leq 0$ for any $\hat{\theta} \in \Theta$, and thus $\theta_p(z) = \underline{\theta}$. ∎

PROOF OF PROPOSITION 2:

For some $\hat{w}^1$ in $(w_\alpha^1(\overline{\theta}), 1)$, I extend β^{se} to $\mathbb{R}$ in the following way:[23]

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

The function $\hat{\beta}$ is C^1, convex, and below β^{se} on $[0, 1]$. I then define a problem $\hat{\mathcal{P}}$ which differs from $\mathcal{P}$ in two aspects: (i) (10) is dropped, (ii) (9) is replaced with the following:

$$\text{(A2)} \quad \theta \eta_\alpha w^1(\theta) - \eta_\alpha \int_{\underline{\theta}}^{\theta} w^1(\tilde{\theta})\, d\tilde{\theta} - \underline{\theta} \eta_\alpha w^1(\underline{\theta}) + \underline{w}^0 - \hat{\beta}(w^1(\theta)) \geq 0, \quad \forall\, \theta \in \Theta.$$

Therefore, $\hat{\mathcal{P}}$ is a relaxation of $\mathcal{P}$. If the solution to $\hat{\mathcal{P}}$ is admissible, it is also the solution to $\mathcal{P}$. Define the Lagrangian associated with $\hat{\mathcal{P}}$ as

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

where Λ is the multiplier associated with (A2). Fixing a nondecreasing Λ, the Lagrangian is a concave functional on Φ, because all terms are linear in $\left(w^1, \underline{w}^0\right)$ except $\int_{\underline{\theta}}^{\overline{\theta}} - \hat{\beta}\left(w^1(\theta)\right) d\Lambda$ which is concave in w^1. Without loss of generality, I set $\Lambda(\overline{\theta}) = 1$. Integrating the Lagrangian by parts yields

$$\text{(A3)} \quad \hat{L}(w^1, \underline{w}^0 \mid \Lambda) = \left(\underline{\theta} \eta_\alpha w^1(\underline{\theta}) - \underline{w}^0\right) \Lambda(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}} \left(\theta \eta_\alpha w^1(\theta) - \hat{\beta}\left(w^1(\theta)\right)\right) \mathrm{d}\Lambda$$
$$+ \eta_\alpha \int_{\underline{\theta}}^{\overline{\theta}} w^1(\theta) [\Lambda(\theta) - (1 - G(\theta))]\, d\theta.$$

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

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

LEMMA 5 (Langrangian—Sufficiency): *A contract* $\left(\tilde{w}^1, \underline{\tilde{w}}^0\right) \in \Phi$ *solves* $\hat{\mathcal{P}}$ *if* (A2) *holds with equality and there exists a nondecreasing* $\tilde{\Lambda}$ *such that* $\hat{L}\left(\tilde{w}^1, \underline{\tilde{w}}^0 \mid \tilde{\Lambda}\right) \geq \hat{L}\left(w^1, \underline{w}^0 \mid \tilde{\Lambda}\right), \forall\left(w^1, \underline{w}^0\right) \in \Phi$.

PROOF:

I first introduce the problem studied in Luenberger (1997, Section 8.4, p. 220): $\max_{x \in X} Q(x)$ subject to $x \in \Omega$ and $J(x) \in P$, where Ω is a subset of the vector space X, $Q : \Omega \to \mathbb{R}$ and $J : \Omega \to Z$; where Z is a normed vector space, and P is a nonempty positive cone in Z. To apply Theorem 1 in Luenberger (1997, p. 220), set

(A4)

$$X = \left\{w^1, \underline{w}^0 \mid w^1 : \Theta \to \mathbb{R} \quad \text{and} \quad \underline{w}^0 \in \mathbb{R}\right\}, \quad \Omega = \Phi,$$

$$Z = \{z \mid z : \Theta \to \mathbb{R} \text{ with } \sup_{\theta \in \Theta} |z(\theta)| < \infty\}, \quad \text{with the norm } \|z\| = \sup_{\theta \in \Theta} |z(\theta)|,$$

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

I let the objective in (OBJ) be Q and the left-hand side of (A2) be J. This result holds because the hypotheses of Theorem 1 in Luenberger (1997, p. 220) are met. ∎

To apply Lemma 5 and show that $\left(\tilde{w}^1, \underline{\tilde{w}}^0\right)$ maximizes $\hat{L}\left(w^1, \underline{w}^0 \mid \tilde{\Lambda}\right)$ for some $\tilde{\Lambda}$, I modify Lemma 1 in Luenberger (1997, p. 227) which concerns the maximization of a concave functional in a convex cone. Note that Φ is not a convex cone, so Lemma 1 in Luenberger (1997, p. 227) does not apply directly in the current setting.

LEMMA 6 (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* $\partial L(\tilde{x}; x - \tilde{x}) = L(\tilde{x}; x) - L(\tilde{x}; \tilde{x})$.[24] *A sufficient condition that* $\tilde{x} \in \Omega$ *maximizes* L *over* Ω *is that* $\partial L(\tilde{x}; x) \leq 0, \forall x \in \Omega$ *and* $\partial L(\tilde{x}; \tilde{x}) = 0$.

PROOF:

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

$$L(\tilde{x} + \epsilon(x - \tilde{x})) \geq L(\tilde{x}) + \epsilon\big(L(x) - L(\tilde{x})\big)$$

$$\Rightarrow L(x) - L(\tilde{x}) \leq \frac{1}{\epsilon}\big(L(\tilde{x} + \epsilon(x - \tilde{x})) - L(\tilde{x})\big).$$

[24] Let X be a vector space, Y a normed space, and $D \subset X$. Given a transformation $T : D \to Y$, if for $\tilde{x} \in D$ and $x \in X$ the limit $\lim_{\epsilon \to 0}\big(T(\tilde{x} + \epsilon x) - T(\tilde{x})\big)/\epsilon$ 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 $\tilde{x}$.

As $\epsilon \to 0+$, 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 $x \in \Omega$. ∎

I next prove Proposition 2. To apply Lemma 6, I let X and Ω be the same as in (A4). It follows from Lemma A.1 in Amador, Werning, and Angeletos (2006)—hereafter, AWA—that $\partial \hat{L}(w^1_{\theta_p}, \underline{w}^0_{\theta_p}; w^1, \underline{w}^0 | \Lambda)$ and $\partial \hat{L}\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}; \left(w^1, \underline{w}^0\right) - \left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}\right) | \Lambda\right)$ exist for any $\left(w^1, \underline{w}^0\right) \in \Phi$.[25] The linearity condition is also satisfied:

$$\partial \hat{L}\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}; \left(w^1, \underline{w}^0\right) - \left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}\right) | \Lambda\right)$$

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

So the hypotheses of Lemma 6 are met. The Gâteaux differential at $\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}\right)$ is

$$\text{(A5)} \quad \partial \hat{L}\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}; w^1, \underline{w}^0 | \Lambda\right) = \left(\underline{\theta}\eta_\alpha w^1(\underline{\theta}) - \underline{w}^0\right)\Lambda(\underline{\theta}) + \eta_\alpha \int_{\theta_p}^{\overline{\theta}} (\theta - \theta_p) w^1(\theta)\, d\Lambda$$

$$+ \eta_\alpha \int_{\underline{\theta}}^{\overline{\theta}} w^1(\theta)[\Lambda(\theta) - (1 - G(\theta))]\, d\theta, \quad \forall \left(w^1, \underline{w}^0\right) \in \Phi.$$

Next, I construct a nondecreasing multiplier $\tilde{\Lambda}$ in a similar manner as in proposition 3 in AWA. Let $\tilde{\Lambda}(\underline{\theta}) = 0$, $\tilde{\Lambda}(\theta) = 1 - G(\theta)$ for $\theta \in (\underline{\theta}, \theta_p]$, and $\tilde{\Lambda}(\theta) = 1$ for $\theta \in \left(\theta_p, \overline{\theta}\right]$. 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 $\tilde{\Lambda}$ into (A5) yields

$$\partial \hat{L}(\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}; w^1, \underline{w}^0 | \tilde{\Lambda}\right) = \eta_\alpha \int_{\theta_p}^{\overline{\theta}} w^1(\theta) G(\theta)\, d\theta$$

$$= \eta_\alpha \int_{\theta_p}^{\overline{\theta}} \left(\int_{\theta}^{\overline{\theta}} G\left(\tilde{\theta}\right) d\tilde{\theta}\right) dw^1(\theta),$$

where the last equality follows by integrating by parts. This Gâteaux differential is zero at $\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}\right)$ and, by the definition of θ_p, it is nonpositive for all w^1 nondecreasing. It follows that the first-order conditions $\partial \hat{L}\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}; w^1, \underline{w}^0 | \tilde{\Lambda}\right) \leq 0$ and $\partial \hat{L}\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}; w^1_{\theta_p}, \underline{w}^0_{\theta_p} | \tilde{\Lambda}\right) = 0$ are satisfied for any $\left(w^1, \underline{w}^0\right) \in \Phi$. By Lemma 6, $\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}\right)$ maximizes $\hat{L}\left(w^1, \underline{w}^0 | \tilde{\Lambda}\right)$ over Φ. By Lemma 5, $\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}\right)$ solves $\hat{\mathcal{P}}$. Because $\left(w^1_{\theta_p}, \underline{w}^0_{\theta_p}\right)$ is admissible, it solves $\mathcal{P}$. ∎

[25] The observations made by Ambrus and Egorov (2013) do not apply to the results of AWA that my proof relies on.

PROOF OF PROPOSITION 3:

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 odds ratio to $\theta e^{-\lambda^1 \delta}$, referred to as his type at time δ. Let $\underline{\theta}_\delta = \max\left\{\underline{\theta}, \theta_\alpha^* e^{\lambda^1 \delta}\right\}$. 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

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

Here, $1 - p(\theta)\left(1 - e^{-\lambda^1 \delta}\right)$ 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

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

I prove that continuing the cutoff rule is optimal by showing two things. First, given f_δ at time δ, the threshold of the pooling segment is $\theta_p e^{-\lambda^1 \delta}$. Second, if Assumption 1 holds for $\theta \leq \theta_p$ under f, it holds for $\theta \leq \theta_p e^{-\lambda^1 \delta}$ under f_δ. Given f_δ over $\theta \in \left[\underline{\theta}_\delta e^{-\lambda^1 \delta}, \overline{\theta} e^{-\lambda^1 \delta}\right]$, I define h_δ, H_δ, G_δ as follows:

$$h_\delta(\theta) = \frac{f_\delta(\theta)}{1 + \theta}, \quad \text{and} \quad H_\delta(\theta) = \int_{\underline{\theta}_\delta e^{-\lambda^1 \delta}}^{\theta} h_\delta\left(\tilde{\theta}\right) d\tilde{\theta},$$

$$G_\delta(\theta) = \frac{H_\delta\left(\overline{\theta} e^{-\lambda^1 \delta}\right) - H_\delta(\theta)}{H_\delta\left(\overline{\theta} e^{-\lambda^1 \delta}\right)} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1\right) \theta \frac{h_\delta(\theta)}{H_\delta\left(\overline{\theta} e^{-\lambda^1 \delta}\right)}.$$

Substituting $f_\delta(\theta) = f_\delta^0\left(\theta e^{\lambda^1 \delta}\right) e^{\lambda^1 \delta}$ into $h_\delta(\theta)$ and simplifying, I obtain

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

By simplifying and making a change of variables $z = e^{\lambda^1 \delta} \theta$, I obtain the following

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

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 1 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 $f(\theta)$ by making a change of variables $z = e^{\lambda^1 \delta}\theta$

$$1 - G_\delta(\theta) = \frac{1}{C}\left[\int_{\theta_\delta}^{z} \frac{f(\tilde{z})}{1+\tilde{z}}\, d\tilde{z} + \left(\frac{\eta_\rho}{\eta_\alpha} - 1\right) z \frac{f(z)}{1+z}\right],$$

if Assumption 1 holds for all $\theta \leq \theta_p$ given f, it holds for all $\theta \leq \theta_p e^{-\lambda^1 \delta}$ given f_δ. ∎

PROOF OF LEMMA 4:

If $p_1 \geq 0, p_2 \geq 0$ (or $p_1 \leq 0, p_2 \leq 0$), the maximum in (5) is achieved by the policy which directs the unit resource to R (or S). If $p_1 > 0, p_2 < 0$, the maximum is achieved by a lower-cutoff Markov policy which directs the unit resource to R if the posterior is above the cutoff and to S if below (see Keller and Rady 2010). This cutoff belief, denoted by p^*, is given by $p^*/(1-p^*) = a^*/(1+a^*)$, where a^* is the positive root of equation $r + \lambda^0 - a^*(\lambda^1 - \lambda^0) = \lambda^0(\lambda^0/\lambda^1)^{a^*}$. Let $K(p) \equiv \max_{w\in\Gamma}(p_1, p_2)\cdot w$. Then, $K(p)$ equals zero, if the prior is below p^*, and equals $-p_2\left(\frac{-p_2 a^*}{p_1(1+a^*)}\right)^{a^*}/(a^*+1) + p_1 + p_2$, if above. It follows that the southeast boundary is characterized by $\beta^{se}(w^1) = 1 - (1-w^1)^{\frac{a^*}{a^*+1}}$. If $p_1 < 0, p_2 > 0$, the maximum is achieved by a upper-cutoff Markov policy which directs the unit resource to R if the posterior is below the cutoff and to S if below (see Keller and Rady 2015). Let p^{**} denote this cutoff. The function $K(p)$ is continuous, convex, and nonincreasing in the prior $|p_1|/(|p_1| + |p_2|)$. Except for a kink at p^{**}, $K(p)$ is once continuously differentiable. Hence, the northwest boundary is concave, nondecreasing, and once continuously differentiable, with the end points being (0, 0) and (1, 1). ∎

PROOF OF PROPOSITION 5:

Suppose that a success occurs at time t. Before the success, the type distribution is denoted f_{t-}. Without loss of generality, suppose that the support is $[\underline{\theta}, \overline{\theta}]$. Let $Q(\theta, dt)$ be the probability that a success occurs in an infinitesimal interval $[t, t+dt)$ given type θ: $Q(\theta, dt) = p(\theta)\left(1 - e^{-\lambda^1 dt}\right) + (1-p(\theta))\left(1 - e^{-\lambda^0 dt}\right)$. 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

$$f_t^*(\theta) = \lim_{dt\to 0} \frac{f_{t-}(\theta)Q(\theta, dt)}{\int_{\underline{\theta}}^{\overline{\theta}} f_{t-}(\theta)Q(\theta, dt)\, d\theta} = \frac{f_{t-}(\theta)\left[p(\theta)\lambda^1 + (1-p(\theta))\lambda^0\right]}{\int_{\underline{\theta}}^{\overline{\theta}} f_{t-}(\theta)\left[p(\theta)\lambda^1 + (1-p(\theta))\lambda^0\right] d\theta}.$$

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

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

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

PROOF OF PROPOSITION 6:

By integrating the envelope condition $U'_\alpha(\theta) = \eta_\alpha w^1(\theta)$, one obtains the standard integral condition

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

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

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

subject to

$$\text{(A6)} \qquad \beta^{se}\big(w^1(\theta)\big) \leq \theta\eta_\alpha w^1(\theta) + \int_\theta^{\overline{\theta}} \eta_\alpha w^1(\tilde{\theta})\, d\tilde{\theta} - \overline{\theta}\eta_\alpha w^1(\overline{\theta}) + \overline{w}^0, \quad \forall \theta \in \Theta,$$

$$\text{(A7)} \qquad \beta^{nw}\big(w^1(\theta)\big) \geq \theta\eta_\alpha w^1(\theta) + \int_\theta^{\overline{\theta}} \eta_\alpha w^1(\tilde{\theta})\, d\tilde{\theta} - \overline{\theta}\eta_\alpha w^1(\overline{\theta}) + \overline{w}^0, \quad \forall \theta \in \Theta,$$

where $\Phi^s \equiv \{w^1, \overline{w}^0 \,|\, w^1 : \Theta \to [0, 1], w^1 \text{ is nondecreasing}, \overline{w}^0 \in [0, 1]\}$. I denote this problem by $\mathcal{P}^s$.

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

$$\text{(A8)} \quad -\eta_\alpha \int_{\underline{\theta}}^{\theta_p^s} w^1(\theta)(1 - G(\theta))\, d\theta = -\eta_\alpha w^1(\theta_p^s) \int_{\underline{\theta}}^{\theta_p^s} (1 - G(\theta))\, d\theta + \eta_\alpha \int_{\underline{\theta}}^{\theta_p^s} \int_{\underline{\theta}}^{\theta} \big(1 - G(\tilde{\theta})\big)\, d\tilde{\theta}\, dw^1(\theta).$$

The first term depends only on $w^1(\theta_p^s)$. The integrand of the second term, $\int_{\underline{\theta}}^{\theta} \big(1 - G(\tilde{\theta})\big)\, d\tilde{\theta}$, is negative for all $\theta \in [\underline{\theta}, \theta_p^s]$. Thus, it is optimal to set $dw^1(\theta) = 0, \forall \theta \in [\underline{\theta}, \theta_p^s]$. If $\theta_p^s = \overline{\theta}$, all types are pooled, and implement the principal's preferred uninformed bundle on Γ^{se}. If $\theta_p^s < \overline{\theta}$, $\int_{\underline{\theta}}^{\theta_p^s}(1 - G(\theta))\, d\theta$ equals zero. Adjusting $w^1(\theta_p^s)$ does not affect (A8), so $w^1(\theta_p^s)$ can be decreased until either (A6) or (A7) binds. ■

PROOF OF PROPOSITION 7:

I define a problem $\hat{\mathcal{P}}^s$ which differs from $\mathcal{P}^s$ in two aspects: (i) (A7) is dropped, and (ii) (A6) is replaced with the following:

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

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

$$\hat{L}^s(w^1, \overline{w}^0 | \Lambda) = \overline{\theta}\eta_\alpha w^1(\overline{\theta}) - \overline{w}^0 - \eta_\alpha \int_{\underline{\theta}}^{\overline{\theta}} w^1(\theta)(1 - G(\theta))\, d\theta$$

$$+ \int_{\underline{\theta}}^{\overline{\theta}} \left(\theta\eta_\alpha w^1(\theta) + \int_\theta^{\overline{\theta}} \eta_\alpha w^1(\tilde{\theta})\, d\tilde{\theta} - \overline{\theta}\eta_\alpha w^1(\overline{\theta}) + \overline{w}^0 - \hat{\beta}(w^1(\theta))\right) d\Lambda.$$

Based on Lemma 5 and 6, it suffices to construct a nondecreasing Λ such that the first-order conditions $\partial\hat{L}^s\left(w^1_{\theta_p^s}, \overline{w}^0_{\theta_p^s}; w^1, \overline{w}^0 | \Lambda\right) \leq 0$ and $\partial\hat{L}^s\left(w^1_{\theta_p^s}, \overline{w}^0_{\theta_p^s}; w^1_{\theta_p^s}, \overline{w}^0_{\theta_p^s} | \Lambda\right) = 0$ are satisfied for any $(w^1, \overline{w}^0) \in \Phi^s$. Let $\Lambda(\theta) = 0$ for $\theta \in [\underline{\theta}, \theta_p^s)$, $\Lambda(\theta) = 1 - G(\theta)$ for $\theta \in [\theta_p^s, \overline{\theta})$, and $\Lambda(\overline{\theta}) = 1$. The jump at θ_p^s is nonnegative according to the definition of θ_p^s. The jump at $\overline{\theta}$ is nonnegative because $1 - G(\theta) \leq 1$ for all θ. If $1 - G(\theta)$ is nondecreasing for $\theta \in [\theta_p, \overline{\theta}]$, $\Lambda(\theta)$ is nondecreasing. ■

REFERENCES

Aghion, Philippe, and Jean Tirole. 1997. “Formal and Real Authority in Organizations.” *Journal of Political Economy* 105 (1): 1–29.

Alonso, Ricardo, and Niko Matouschek. 2008. “Optimal Delegation.” *Review of Economic Studies* 75 (1): 259–93.

Amador, Manuel, and Kyle Bagwell. 2013. “The Theory of Optimal Delegation with an Application to Tariff Caps.” *Econometrica* 81 (4): 1541–99.

Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. “Commitment vs. Flexibility.” *Econometrica* 74 (2): 365–96.

Ambrus, Attila, and Georgy Egorov. 2013. “Comment on ‘Commitment vs. Flexibility.’” *Econometrica* 81 (5): 2113–24.

Armstrong, Mark, and John Vickers. 2010. “A Model of Delegated Project Choice.” *Econometrica* 78 (1): 213–44.

Athey, Susan, Andrew Atkeson, and Patrick J. Kehoe. 2005. “The Optimal Degree of Discretion in Monetary Policy.” *Econometrica* 73 (5): 1431–75.

Aumann, Robert J. 1964. “Mixed and Behavior Strategies in Infinite Extensive Games.” In *Advances in Game Theory*, edited by Melvin Dresher, Lloyd S. Shapley, and Albert William Tucker, 627–50. Princeton: Princeton University Press.

Bergemann, Dirk, and Ulrich Hege. 1998. “Venture Capital Financing, Moral Hazard, and Learning.” *Journal of Banking & Finance* 22 (6-8): 703–35.

Bergemann, Dirk, and Ulrich Hege. 2005. “The Financing of Innovation: Learning and Stopping.” *RAND Journal of Economics* 36 (4): 719–52.

Bolton, Patrick, and Christopher Harris. 1999. “Strategic Experimentation.” *Econometrica* 67 (2): 349–74.

Bonatti, Alessandro, and Johannes Hörner. 2011. "Collaborating." *American Economic Review* 101 (2): 632–63.
Cohen, Asaf, and Eilon Solan. 2013. "Bandit Problems with Lévy Processes." *Mathematics of Operations Research* 38 (1): 92–107.
Dessein, Wouter. 2002. "Authority and Communication in Organizations." *Review of Economic Studies* 69: 811–38.
Ederer, Florian. 2013. "Incentives for Parallel Innovation." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2309664 (accessed June 24, 2016).
Frankel, Alexander. 2014. "Aligned Delegation." *American Economic Review* 104 (1): 66–83.
Garfagnini, Umberto. 2011. "Delegated Experimentation." Unpublished.
Gerardi, Dino, and Lucas Maestri. 2012. "A Principal-Agent Model of Sequential Testing." *Theoretical Economics* 7 (3): 425–63.
Gomes, Renato, Daniel Gottlieb, and Lucas Maestri. 2016. "Experimentation and Project Selection: Screening and Learning." *Games and Economic Behavior* 96: 145–69.
Grenadier, Steven, Andrey Malenko, and Nadya Malenko. 2016. "Timing Decisions in Organizations: Communication and Authority in a Dynamic Environment." Stanford GSB Working Paper 3049.
Halac, Marina, Navin Kartik, and Qingmin Liu. Forthcoming a. "Contests for Experimentation." *Journal of Political Economy.*
Halac, Marina, Navin Kartik, and Qingmin Liu. Forthcoming b. "Optimal Contracts for Experimentation." *Review of Economic Studies*.
Holmström, Bengt. 1977. "On Incentives and Control in Organizations." PhD diss. Stanford University.
Holmström, Bengt. 1984. "On the Theory of Delegation." In *Bayesian Models in Economic Theory*, edited by Marcel Boyer and Richard Kihlstrom, 115–41. New York: North-Holland.
Hörner, Johannes, and Larry Samuelson. 2013. "Incentives for Experimenting Agents." *RAND Journal of Economics* 44 (4): 632–63.
Keller, Godfrey, and Sven Rady. 2010. "Strategic Experimentation with Poisson Bandits." *Theoretical Economics* 5 (2): 275–311.
Keller, Godfrey, and Sven Rady. 2015. "Breakdowns." *Theoretical Economics* 10 (1): 175–202.
Keller, Godfrey, Sven Rady, and Martin Cripps. 2005. "Strategic Experimentation with Exponential Bandits." *Econometrica* 73 (1): 39–68.
Klein, Nicolas. Forthcoming. "The Importance of Being Honest." *Theoretical Economics*.
Klein, Nicolas, and Sven Rady. 2011. "Negatively Correlated Bandits." *Review of Economic Studies* 78 (2): 693–732.
Krishna, Vijay, and John Morgan. 2008. "Contracting for Information under Imperfect Commitment." *RAND Journal of Economics* 39 (4): 905–25.
Luenberger, David G. 1997. *Optimization by Vector Space Methods.* New York: Wiley.
Manso, Gustavo. 2011. "Motivating Innovation." *Journal of Finance* 66 (5): 1823–60.
Martimort, David, and Aggey Semenov. 2006. "Continuity in Mechanism Design without Transfers." *Economics Letters* 93 (2): 182–89.
Melumad, Nahum D., and Toshiyuki Shibano. 1991. "Communication in Settings with No Transfers." *RAND Journal of Economics* 22 (2): 173–98.
Milgrom, Paul, and Ilya Segal. 2002. "Envelope Theorems for Arbitrary Choice Sets." *Econometrica* 70 (2): 583–601.
Moroni, Sofia. 2016. "Experimentation in Organizations." https://drive.google.com/file/d/0B5GUFW9n2v0VcnNibDZpXzJPSEU/view?usp=sharing (accessed June 24, 2016).
Murto, Pauli, and Juuso Välimäki. 2011. "Learning and Information Aggregation in an Exit Game." *Review of Economic Studies* 78 (4): 1426–61.
Mylovanov, Tymofiy. 2008. "Veto-based Delegation." *Journal of Economic Theory* 138 (1): 297–307.
Presman, E. L. 1990. "Poisson Version of the Two-Armed Bandit Problem with Discounting." *Theory of Probability and Its Applications* 35 (2): 307–17.
Rosenberg, Dinah, Eilon Solan, and Nicolas Vieille. 2007. "Social Learning in One-Arm Bandit Problems." *Econometrica* 75 (6): 1591–1611.
Seierstad, Atle, and Knut Sydsæter. 1987. *Optimal Control Theory with Economic Applications.* Amsterdam: Elsevier Science.
Seru, Amit. 2014. "Firm Boundaries Matter: Evidence from Conglomerates and R&D Activity." *Journal of Financial Economics* 111 (2): 381–405.
Strulovici, Bruno. 2010. "Learning While Voting: Determinants of Collective Experimentation." *Econometrica* 78 (3): 933–71.