Dead Ends*

Evan Sadler†

October 20, 2017

Abstract

Evidence suggests that innovation benefits from low-powered incentives, yet innovative organizations use widely varying incentive structures. I offer an explanation based on a characteristic feature of creative work: dead ends. To solve a problem, an agent works on successive ideas, each of which may succeed with some probability. At each instant, the agent chooses whether to exert effort. The agent may also abandon his idea, incurring delay to come up with a new one. High rewards for success can slow innovation because the agent is reluctant to incur the cost of delay, spending too much time on unpromising ideas. I apply this framework to study intellectual property rights (IPR) and optimal contracts for innovation. Dead ends provide a new explanation for the inverse U relationship between IPR and innovation, suggesting that "low-hanging fruit" suffers most from strong IPR. In a principal-agent setting with moral hazard, we get front-loading because high continuation values increase the cost of current incentives. Contract structure depends on whether the principal is more or less patient than the agent. Impatient principals impose deadlines, while patient principals grant tenure, using delay rather than the threat of termination to reduce incentive costs.

1 Introduction

Creative work is a major source of economic output. The literature on endogenous growth is predicated on the importance of research and innovation, and firms invest billons of dollars in R&D spending.[1] Whether researching new drugs, designing new products, developing new trading strategies, or starting new businesses, a large share of workers engage in innovative activity as a central part of their jobs. How best to incentivize these workers is a natural question for firms, governments, and other organizations.

*I am especially grateful to Daron Acemoglu, Abhijit Banerjee, Ben Golub, Aditya Kuvalekar, and Elliot Lipnowski for advice and feedback. I also thank Dhruva Bhaskar, Bob Gibbons, Yingni Guo, Harry Pei, Roy Radner, and seminar participants at Harvard and MIT for helpful conversations and suggestions.

†Harvard University – esadler@fas.harvard.edu

[1]See Klette and Kortum (2004); Lentz and Mortensen (2008); Acemoglu et al. (2013) for recent contributions to the endogenous growth literature. The World Bank estimates that the United States spent 2.73% of its GDP on R&D in 2013.

Intuitively, high rewards for innovative success should incentivize effort and lead to more innovation, and indeed there is evidence suggesting that employee stock ownership is associated with greater innovative activity (Garrett, 2010; Chang et al., 2015). However, a central message in studies of innovation and incentives is that innovation benefits from low-powered incentives—payoffs that respond less drastically and less immediately to outcomes. Extensive research, both in the field and the laboratory, finds that tolerance for early failure, long-run performance incentives, and ceding creative control are associated with higher levels of innovation. In a corporate context, long-term incentives for research and development heads (Lerner and Wulf, 2007), longer stock option vesting periods (Yanadori and Marler, 2006), golden parachutes (Francis et al., 2011), and a failure-tolerant culture (Tian and Wang, 2014) are all associated with firms creating more patents and more heavily cited patents. Ederer and Manso (2013) present evidence from a laboratory experiment that long-term incentives and tolerance for failure help motivate innovation, and Azoulay et al. (2011) show how grants that are more failure tolerant, and give researchers more flexibility over what projects to pursue, increase the output of academic scientists.

The incentives that innovative organizations use in practice are highly varied in strength and structure. The routine output of an academic scientist has little if any near-term effect on compensation—long-term effects come from improved job prospects and reputation, which are often difficult to tie to any single piece of work. Traders in financial markets, who engage in highly creative work to develop effective strategies, receive a majority of their compensation through stock options and performance contingent bonuses, and they are often subject to "up or out" promotion policies. Employees in technology startups are typically compensated through equity in the company. Given the apparent efficacy of low-powered incentives, why do we see so much variation, and what effect does this have on the level of innovation across different organizations?

I offer a new perspective based on a distinguishing feature of creative work: dead ends. As any experienced researcher knows, not every new idea works as intended. A serial entrepreneur might suffer several failures and bankruptcies before coming up with a successful business plan. An academic scientist may try multiple experimental designs before convincing evidence emerges for or against a hypothesis. In each case, the worker decides not just how hard to work, but also how long to keep trying before searching for a new approach.

The option to come up with a new idea leads to a non-trivial tradeoff: do I suffer delay to start on a more promising project, or do I hold out just a bit longer hoping for a quick success? This tradeoff offers a new explanation why high-powered incentives discourage creativity. A high reward for success makes delay more costly, causing an individual to work "too hard" on any one idea. In an agency setting, contractual features like tenure and "up or out" emerge depending on whether the prinicpal or the agent is more patient. To model the innovation process, I extend the exponential bandit framework of Keller et al. (2005).

To solve a creative problem, an agent successively comes up with ideas, and the agent works on each idea until either he solves the problem or gives up to start on a new idea. Each idea is good with independent probability p_0. If an idea is good, and the agent exerts effort, a breakthrough arrives at a constant positive rate. If the idea is bad, or the agent makes no effort, a breakthrough never arrives. At any time, the agent may give up and incur delay to come up with a new idea—this option is the key departure from previous work. We could imagine the agent taking a break from work to brainstorm, talk to colleagues, or read about what others are doing until a new approach occurs to him. After formally describing the innovative process, I study two applications in detail. First, for a self-employed agent, I ask how incentives affect the rate of innovation and how dead ends might influence intellectual property rights. I then consider a principal-agent setting with moral hazard, asking how dead ends affect the structure of optimal contracts.

A self-employed agent chooses a constant threshold belief at which he gives up on any idea. This threshold decreases as rewards increase or as the agent becomes less patient—an impatient agent receiving a large reward for success spends more time working and less time searching for ideas. Intuitively, we can think of the option to search as the "safe arm" in a more standard bandit problem—though its value is determined endogenously. Higher rewards and higher discount rates diminish the value of the safe arm relative to continued experimentation. The expected time until success is a U-shaped function of the threshold belief. As a result, impatient agents earning high rewards are slow because they spend a lot of time working on dead ends. Conversely, patient agents earning low rewards are slow because they give up easily. These results highlight that fast innovation is not the same as efficient innovation, and we obtain a nuanced account of when low-powered incentives facilitate faster innovation.

I emphasize that the tradeoff behind these results is fundamentally different than the familiar exploration versus exploitation tradeoff. In the latter case (e.g. Manso, 2011), high rewards for early success make it more costly to gather information, so an agent gravitates towards the safe arm. Discounting the future makes information less valuable still, reinforcing this effect. Here, impatience and high rewards make the safe arm *less attractive*. This occurs because of a switching cost between the two actions, the cost of delay.

Our findings can inform a social planner who designs a regime of intellectual property rights for innovative agents. A standard argument to limit protections is that they are ex-post inefficient: information is non-rival, and intellectual property rights can impede free dissemination and use of an innovation. Assuming the planner is more patient than the agent, I suggest a rationale to limit protections even without any ex-post inefficiency. High rewards cause short-sighted agents to work harder on any one idea than is socially optimal. This result offers an alternative explanation for the inverse U-shaped relationship between intellectual property rights and innovation (Qian, 2007). It also suggests that "low-hanging fruit," innovations with large benefits relative to development costs, suffers most from strong protections.

In a second application, I introduce a principal who earns a payoff from the agent's success and designs a contract to incentivize the agent. The principal provides a flow of resources for the agent to work on his ideas, but the agent can divert these funds for personal gain. The principal observes when the agent starts on a new idea but does not observe the diversion of funds. A technical lemma shows that the principal can implement any desired experimentation policy and specifies the reward function that does so most cheaply. The agent extracts rents from two sources. First, to prevent the agent from diverting funds, the principal must pay the agent at least the value these funds. Second, there is option value from diverting funds today and collecting the intended continuation payoff tomorrow. The principal must further compensate the agent for the value of this option. Since high continuation values increase the cost of current incentives, optimal contracts must use front-loading. Interestingly, the characterization of the optimal reward follows from a first order approach despite the presence of hidden information. Diverting funds increases the chance of success later—since the agent has not tried and failed, his belief about the likelihood of success becomes higher than the principal's. Nevertheless, the agent cannot gain from such a deviation because the reward function front-loads his surplus: the agent loses more from delaying his work than what he can later recoup.

An important feature of this analysis is that the principal and agent may have different discount rates. With access to perfect financial markets, we should expect equal discount rates, but this is hardly realistic for the relevant settings. An agent who diverts funds often does not obtain cash; he might instead finance a pet project or office perks. Likewise, while firms typically discount based on a market-interest rate, it is well-documented that markets for financing innovation are incomplete (Lerner and Hall, 2010). External financing is typically more expensive than internal financing, so the capital structure of a firm affects the required rate of return for R&D investments. For some principals (e.g. academic or government institutions), the value of an innovation includes less tangible benefits like prestige, which are not traded in any market, and hence the relevant discount rate is partly subjective. I call attention to this issue not merely for the sake of generality: the relationship between the players' discount rates affects the qualitative structure of an optimal contract.

When the agent is at least as patient as the principal, optimal contracts take a simple form. The agent works continuously, the promised continuation value declines with each new idea, and if the agent fails to achieve success, he is eventually fired. By imposing a deadline, the principal reduces the option value of delaying work, thereby reducing the cost of incentives early in the relationship. This is clearly inefficient as the principal fails to obtain a solution with positive probability. In contrast, when the prinicpal is more patient than the agent, the contract looks like a tenure arrangement: the agent is employed in perpetuity, and the promised continuation value is bounded above zero. While the problem is sure to get solved, we see a different form of inefficiency in these contracts. To reduce incentive costs, the principal may periodically ask the agent to take a break from working on his ideas.

Hitting dead ends is an inherent part of innovation that has implications for the structure of incentives. This feature gives us a new rationale why low-powered incentives facilitate faster innovation. Impatience can motivate an agent to work too hard on a given project because coming up with new ideas entails costly delay. Weaker incentives can counteract this effect. Dead ends also create a role for time preferences in the structure of optimal contracts. The option to wait for a new idea means that high continuation values increase the cost of current incentives. This leads to inefficient front-loading that manifests differently depending on who is more patient. Impatient principals impose deadlines, while patient principals use extra delay. Beyond these applications, the underlying innovation model tractably addresses important limitations of the experimentation literature and can potentially serve as a workhorse model in a variety of contexts.

1.1 Related Work

Economic theory offers two main explanations why low-powered incentives help encourage innovation. First, strong incentives may crowd out an agent's intrinsic motivation. Work in psychology documents that extrinsic rewards can impair an individual's intrinsic motivation to perform a task (Deci et al., 1999), and intrinsic motivation seems particularly important in creative pursuits. Bénabou and Tirole (2003) develop a principal-agent model in which this crowding out effect can appear. In their model, a principal has private information about an agent's ability to complete some task, and she can offer rewards to induce the agent to try. If the principal receives a negatve signal, she may worry that the agent—who also has private information about his ability—is too pessimistic to attempt the task. She can use a strong reward to compensate, but this provides the agent with additional negative information, undermining his self-confidence. While the reward can get the agent to try in the short-run, the agent is more reluctatnt to undertake similar tasks in the future.

A second explanation stems from an agent's aversion to uncertainty. In many cases, an agent can choose whether to innovate or to rely on current knowledge. Innovating entails a chance of failure, so rewards that are contingent upon success make pursuing innovation more risky for the agent. Manso (2011) formalizes this idea using a two-period principal-agent model. In each of the two periods, the agent may exploit an action with a known chance of success, explore a new action with unknown chance of success, or shirk. Exploring the new action leads to learning about its success probability—a success in the first period leads to higher beliefs in the second period, with a failure leading to lower beliefs. The optimal contract that induces exploration may reward failure in the first period while paying nothing for a success. The principal offers high rewards for success in the second period so that the agent finds it valuable to gather information in the first period. Manso explicitly ties this incentive structure to common features of managerial compensation contracts like stock options and golden parachutes. In a related model, Byun (2015) shows that ambiguity about the likelihood of successful innovation exacerbates the effect. When an agent can choose whether or not to innovate, or can choose between radical or incremental innovation, strong performance incentives discourage the agent from choosing the riskier option.

In contrast with both of these frameworks, there is no question here whether the agent will try to innovate—the agent's decision concerns *when* to take another approach. As a result, time preferences play a central role. The strength of incentives and the agent's discount rate jointly determine how the agent perceives the switching cost. When the agent decides to switch determines how fast, not whether, the agent innovates. The tradeoff between working more and waiting for a new idea tells us more about the conditions under which low-powered incentives encourage creativity and has distinct implications for the design of incentives.

Methodologically, I build on a large and growing literature that studies experimentation using exponential bandits. This literature includes work on strategic experimentation (Keller et al., 2005; Keller and Rady, 2015), contest design (Halac et al., 2016a), and delegation (Guo, 2016). The option to quit and come up with a new project is an innovation in this paper. Much of the experimentation literature motivates itself through questions about firm R&D or venture capital financing. A limitation in all of this work is the assumption that there is only one possible project. After exerting enough effort without success, it is clearly optimal to quit, and the game ends. The R&D division cannot reallocate its resources to a different line of research. The venture capitalist cannot redirect his investment to a new start up. Under this assumption, we miss an important tradeoff that should influence decisions about when to halt a project.

Bergemann and Hege (2005), Hörner and Samuelson (2013), and Halac et al. (2016b) are closest in spirit to the present paper; all study dynamic incentives for experimentation in a principal-agent setting. Both Bergemann and Hege (2005) and Hörner and Samuelson (2013) consider an investor financing an entrepreneur. The key contracting frictions are moral hazard and limited liability, and like in this paper, inefficiency arises because the agent can delay experimentation. In order to reduce incentive costs, the principal may ask for less than full effort and terminates the project early. Two key differences in this paper are the addition of commitment power and the availability of new ideas. In my model, when the principal and the agent have the same discount rate, the prinicpal never asks for less than full effort. This is due to commitment power: instead of delaying the agent's effort to reduce incentive costs, the principal commits to fire the agent if success takes too long. Both settings lead to early termination but for different reasons. With a single project, the principal must terminate eventually and does so when the expected payoff cannot compensate both the investment and incentive costs. Here, the basic problem is stationary—without commitment power the principal would never fire the agent—but the principal commits to fire the agent to reduce earlier incentive costs.

Halac et al. (2016b) study a similar model, dispensing with limited liability but adding adverse selection and commitment. In optimal contracts, high-ability agents experiment efficiently, but low-ability agents quit too early. Without adverse selection, the principal can obtain an efficient outcome in their framework—limited liability prevents that here. The option to switch projects, and potentially different discount rates for the principal and the agent, are notable differences in the present paper. While I do not formally analyze a model with adverse selection, I discuss such an extension in section 4.4.

In a departure from these models, Moroni (2015) considers multi-stage projects. The author studies experimentation in an organizational context. A project consists of multiple "milestones," and we can only proceed to the next milestone after obtaining success on the first. Experimentation is divided among multiple agents. An important focus is when the optimal contract uses monetary bonuses versus promotions to incentivize effort. A key finding is that agents who succeed early are rewarded with additional responsibility to experiment in the future, which yields a higher chance at a high monetary reward in later stages. Rather than examine the tradeoff between bonuses and promotions in incentive provision, I explore the optimal structure of monetary incentives.

Akcigit and Liu (2016) offer a different perspective on dead ends. They study a model with a single risky arm and a safe arm, but there are two firms in competition to develop a patent. One firm may realize the risky arm is bad and secretly move to the safe arm, letting the other firm continue to waste effort on the dead end. The authors study the inefficiency due to competition and how to incentivize firms to share information about dead ends. The model here differs in having infinitely many risky arms and just a single researcher. The ever present option to switch creates distinct tradeoffs, and I focus on the structure of optimal contracts.

This paper is also related to the literature on dynamic principal-agent models, and more broadly to the literature on dynamic contracting. Demarzo and Sannikov (2006), Sannikov (2008), and Zhu (2013) provide seminal contributions on the canonical continuous time principal-agent model based on Brownian motion. Here, the delay to come up with a new idea causes the players' discount rates to affect qualitative features of the optimal contract. As a result, the model offers new insights on the understudied role of time preferences in dynamic contract design. Like in Zhu (2013), optimal contracts can involve periods of "shirking," but only if the principal is patient. Unlike the earlier paper, these periods are neither a reward nor a punishment since no output is observed until the end of the game—delay is used purely to reduce the cost of earlier incentives. Opp and Zhu (2015) provide a recent contribution, showing how the opportunity to trade across time leads to oscillating Pareto-optimal contracts—contracts on the Pareto frontier in our setting should behave similarly. That optimal contracts for the principal do not exhibit oscillations is a consequence of limited liability and agency costs.

With hidden information, it is not immediately clear that checking one-shot deviations is sufficient to show that a contract is incentive compatible. As in the models of Prat and Jovanovic (2014) and Di Tella and Sannikov (2016), we must show that the contract obtained through a first order approach is in fact incentive compatible for the agent. In our case, this follows because the contract we obtain through the first order approach front-loads the agent's surplus, and there is a bound on how hard the agent can work at a given instant. The principal must always compensate the agent for the option value of shirking. As the work progresses, this option value decreases, and so does the surplus the agent can earn. While the agent could drive a wedge between his beliefs and the principal's, there is no way for him to profit from doing so.

Similar to the repeated delegation model of Lipnowski and Ramos (2015), the principal-agent relationship in an optimal contract becomes less productive over time. In the delegation model, this occurs as a discrete shift from always delegating to a continuation contract with periods in which the agent cannot act. The threat of moving to the less productive regime incentivizes good behavior during the more productive regime. Here, the agent's continuation value presents a more direct cost to the principal, causing a steady decline in the relationship over time.

Finally, the application of section 3.2 relates to the literature on innovation and intellectual property rights. Recent studies document an inverse U-shaped relationship between the strength of intellectual property rights and the rate of innovation (Qian, 2007; Lerner, 2009). At least in developed countries, there is convincing evidence that stronger intellectual property rights increase R&D spending (Kanwar and Evenson, 2003), but there is conflicting evidence on how this spending affects actual growth (Jones, 1995; Gould and Gruben, 1996). While other models can generate the inverse U relationship (e.g. Horowitz and Lai, 1996; Furukawa, 2010), the present paper uniquely captures this entire constellation of features. Dead ends highlight that not only the level of investment matters: it matters *how* funds are spent. Strong intellectual property rights may lead to inefficient effort on ideas that are best abandoned.

2 The Innovation Process

An agent faces a problem requiring an innovative solution. A successful solution delivers a payoff $y > 0$. To solve the problem, the agent can work on different ideas, which we represent as exponential bandits. Each new idea is either good or bad, independently with probability p_0. Good ideas can lead to success, while bad ideas cannot.

Working on an idea requires costly effort. At each instant t the agent chooses an action $x(t) \in \{0, 1\}$, incurring a flow cost $cx(t)$ with $c > 0$. If the idea is good, and the agent exerts effort, a success arrives at a constant rate $\lambda > 0$. If the idea is bad, or the agent shirks, success never arrives. At any point in time, the agent can choose to abandon the current idea to come up with a new one. Getting a new idea incurs a time delay, following an exponential distribution with parameter $\gamma > 0$. The agent discounts the future at a rate $r > 0$. Figure 1 summarizes the innovation process.

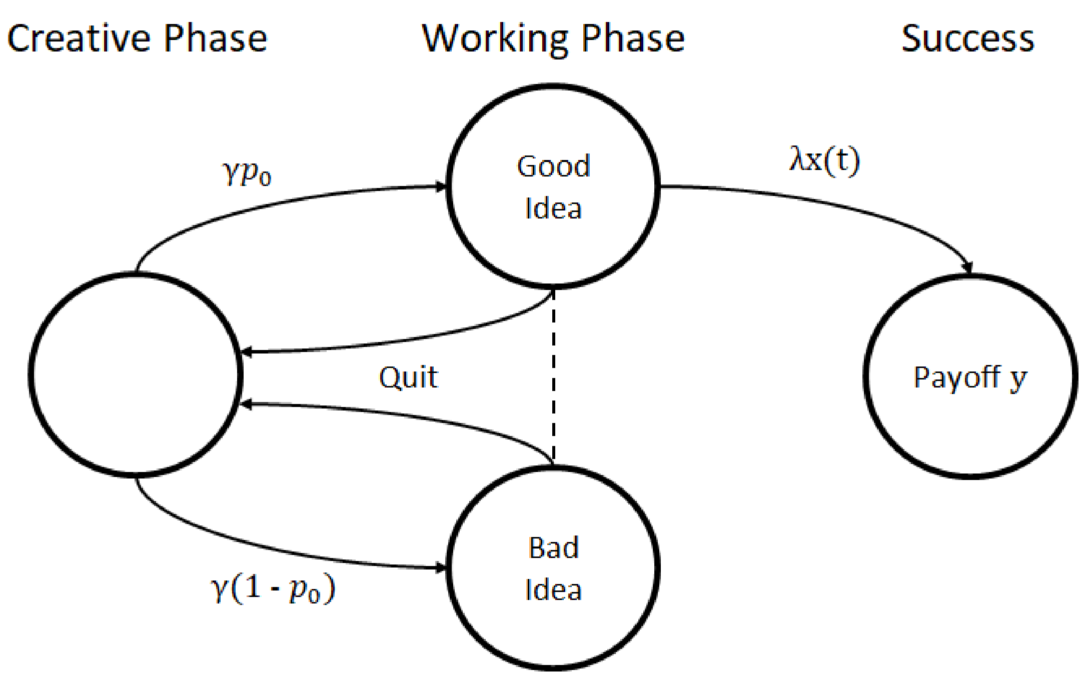

Figure 1: The process of innovation.

Innovation follows a Markov process with transitions that depend on the agent's choices. Success is an absorbing state. We can distinguish two phases that may occur prior to success. The creative phase corresponds to the delay necessary to come up with a new idea. During this phase, the agent transitions to the working phase at the rate γ, reaching a good idea a fraction p_0 of the time. During the working phase, the agent transitions to success at a rate $\lambda x(t)$ if and only if his idea is good. At any instant, the agent can voluntarily return to the creative phase, which intuitively should occur when he is sufficiently convinced that his idea is bad. Assume the agent begins in the creative phase. The agent chooses an effort and quitting policy to maximize his expected discounted payoff.

In subsequent sections, I analyze two versions of the innovation problem. In the first, a self-employed agent incurs the flow cost of any effort and earns the full value of success when it arrives. In the second, a principal employs the agent, providing a flow of resources to carry out the work and retaining the value of success—if the agent shirks at time t, he pockets the flow value $c dt$ that the principal provides. In this case, the principal designs a contract to align incentives.

2.1 A Few Remarks

By varying how we interpret its components, this framework can serve a wide range of applications. If we think of the agent as a scientist, an idea might represent a hypothesis or an experimental design. In this case, we can interpret a creative phase as time spent reading and brainstorming, while a working phase corresponds to actually conducting experiments and analysis. If we think of the agent as an entrepreneur, an idea might correspond to a particular product or marketing strategy. In addition to brainstorming, a creative phase would include activities like procuring materials and equipment or hiring support personnel. A working phase would correspond to the actual execution of the strategy. The two core elements are dead ends and the delay to start a new idea. Key results are robust to many embellishments of the basic model as long as these elements are present.

There are two components to the agent's ability. A higher λ means the agent is more efficient at working on existing ideas—the agent is more productive. In contrast, a higher γ means the agent is more adept at coming up with new ideas—the agent is more creative. Both attributes are valuable, but they have different effects on the agent's decisions. A more creative agent (higher γ) switches ideas more readily because doing so is less costly, while a more productive agent (higher λ) learns faster while working and is therefore more willing to work on ideas with a low chance of success.

The agent incurs the flow cost c only during the working phase, not in the creative phase. Qualitative results would remain unchanged if we instead had separate flow costs $c_w > c_c$ for effort in the working and creative phases respectively. One way to interpret this is that the working phase depends on physical resources that the creative phase does not. For instance, activities like executing a business plan or conducting experiments clearly require capital and equipment that are unnecessary for coming up with ideas. The principal-agent model of section 4 is closely tied to this interpretation of effort. When the agent shirks, we interpret this as diverting the principal's funds for personal gain.

In some contexts, we could also interpret the flow cost as psychological effort. A large literature documents the role of "incubation" in creative problem solving—taking a break from conscious work on a problem helps generate new insights.[2] If we interpret the working phase as consisting mainly of conscious effort, and the creative phase more as incubation periods, our assumption on relative costs seems sensible. Solving the problem is similar under the reverse assumption, but key comparative statics get reversed.

Finally, the framework implicitly assumes that the agent never returns to an old idea once abandoned. Given our other assumptions, this is without loss: new ideas are always available and no news is bad news, so if it is ever optimal to quit and search for a new idea, the agent always prefers a new idea over returning to an old one. This feature of the model seemingly conflicts with experience as researchers often return to old ideas. However, the return to a previously abandoned project typically follows a novel insight, one that this framework would model as a new idea of its own.

3 The Self-Employed Agent

In this section, a self-employed agent incurs the effort cost c and earns the full value of success y. We could imagine an independent firm investing in R&D projects, an entrepreneur self-funding a series of innovative business ventures, or an academic researcher seeking a major breakthrough. A few observations simplify our analysis. First, if y is high enough to ever induce effort, the agent always exerts effort 1. When the agent shirks, it delays success while yielding no information about the idea's quality. Hence, if the agent can obtain a positive payoff from making effort, any such delay is strictly suboptimal. Consequently, the agent's policy amounts to choosing when to cease working on the current idea.

While the agent works without success, his belief p about whether the idea is good steadily decreases. We can characterize the optimal strategy via a threshold belief $\underline{p} \leq p_0$ at which the agent abandons an idea. If we write $u(p)$ for the agent's expected continuation payoff given current belief p, we recognize this as a modified version of the exponential bandit problem of Keller et al. (2005). The key difference is that the value of the "safe" arm is determined endogenously as the value of incurring delay to come up with a new idea.

The first result of this section characterizes the threshold belief $\underline{p}$ at which the agent quits working on the current idea. For notational convenience, I define the agent's **quitting ratio**

$$\rho = \frac{\underline{p}(1-p_0)}{p_0(1-\underline{p})}. \tag{1}$$

The agent's quitting ratio measures how much bad news the agent can receive before giving up. The lower the quitting ratio, the more negative information it takes to get the agent to quit. The higher the quitting ratio, the more easily the agent gives up on the current idea. We can also think of the quitting ratio as an inverse measure of how hard the agent works.

[2] For recent contributions in this area, see Sio and Rudowicz (2007), Hélie and Sun (2010), and Gilhooly (2016).

When ρ is small, the agent spends more time working on each idea, and hence spends a higher fraction of his time making costly effort. I write T for the amount of time the agent spends on an idea before giving up, and note that Bayes' rule implies $e^{-\lambda T} = \rho$.

Proposition 1. *The agent works if and only if $y > \frac{c}{\lambda p_0}$. In this case, the equation*

$$\frac{\gamma}{r}\rho\left(1 - \frac{\lambda \rho^{\frac{r}{\lambda}}}{\lambda + r}\right) + r\rho\frac{\lambda y - c}{\lambda(ry + c)(1 - p_0)} = \frac{\gamma}{\lambda + r} + \frac{cr}{p_0\lambda(ry + c)} \tag{2}$$

uniquely characterizes the belief at which the agent abandons an idea.

Proof. The proof entails solving a Bellman equation for $u(p)$ and explicitly computing $u(p_0)$ to obtain the endpoints. See Appendix for details. □

Equation (2) makes it straightforward to derive comparative statics using implicit differentiation.

Proposition 2. *The agent's quitting ratio ρ has the following comparative statics.*

(a) The quitting ratio is increasing in the cost c and decreasing in the value y.

(b) The quitting ratio is decreasing in the probability p_0 that an idea is good, increasing in creativity γ, and decreasing in productivity λ.

(c) The quitting ratio is decreasing in the agent's discount rate r.

Proof. See Appendix. □

Parts (a) and (b) of Proposition 2 are fairly intuitive. If you lower the cost of effort, or increase the benefit, the agent works harder. If an idea is good with higher probability, the agent finds it worthwhile to work more. If the agent is more creative and comes up with new ideas faster, switching is less costly, and he spends less time on any one idea. If the agent is more productive, it takes more bad news for him to quit, but since information arrives faster, this may take more or less time depending on other parameters.

Part (c) is more subtle. Coming up with a new idea is relatively costly for an impatient agent because the delay entails a larger discount to his reward. The option to switch functions like the "safe" arm in a more standard experimentation problem. When the agent has a higher discount rate r, this effectively reduces the value of the safe arm, which makes experimentation (i.e. working on the current project) more attractive. As a result, the impatient agent will work harder for lower y than a more patient agent.

3.1 The Rate of Innovation

This section demonstrates how low-powered incentives (i.e. a reduction in y) can lead to faster innovation. We first find the experimentation policy that maximizes the rate of innovation. Let μ denote the expected that time passes before the agent succeeds on some idea. There is a unique quitting ratio ρ^* that minimizes the expected time μ until success. For $\rho > \rho^*$, the agent gives up too easily, resulting in slower innovation. For $\rho < \rho^*$, the agent spends too much time working on dead ends, resulting in slower innovation.

Proposition 3. *The expected time until success μ achieves its minimum at the quitting ratio ρ^*, which solves*

$$\gamma(1-p_0)\left[1-\rho(1-\ln\rho)\right]-\lambda\rho=0. \tag{3}$$

For $\rho > \rho^*$, the time μ is strictly increasing in ρ; for $\rho < \rho^*$, the time μ is strictly decreasing in ρ.

Proof. If T is the time spent on an idea before giving up, we can compute μ recursively as

$$\begin{aligned}\mu &= \frac{1}{\gamma}+p_0\int_0^T \lambda t e^{-\lambda t}\mathrm{d}t+\left(1-p_0(1-e^{-\lambda T})\right)(T+\mu)\\ &= \frac{1}{\gamma}+\frac{p_0}{\lambda}(1-e^{-\lambda T})+(1-p_0)T+\left(1-p_0(1-e^{-\lambda T})\right)\mu.\end{aligned}$$

Solving yields

$$\mu=\frac{\lambda+p_0\gamma(1-e^{-\lambda T})+\gamma\lambda(1-p_0)T}{p_0\gamma\lambda(1-e^{-\lambda T})}=1+\frac{\lambda-\gamma(1-p_0)\ln\rho}{p_0\gamma\lambda(1-\rho)}.$$

This is a convex function of ρ on $(0,1)$. Taking first order conditions, we find a unique minumum characterized by (3). □

The quitting ratio that the agent chooses is a function of the discount rate r and the value of success y. While higher y induces the agent to work more, the agent's patience limits how far this can go. From (2), we see that as y approaches infinity, or as c approaches 0, the quitting ratio approaches the solution to

$$\frac{\gamma}{r}\rho\left(1-\frac{\lambda\rho^{\frac{r}{\lambda}}}{\lambda+r}\right)+\frac{\rho}{1-p_0}=\frac{\gamma}{\lambda+r}. \tag{4}$$

I refer to the solution $\overline{\rho}$ of (4) as the agent's **quitting floor**. No matter how valuable success is, or how little effort costs, the agent never works beyond the time that corresponds to $\overline{\rho}$—it is always better to think up a new idea at this point. The proof of Proposition 2 implies that the agent's quitting floor satisfies the same comparative statics as his quitting ratio. In particular, the quitting floor is decreasing in the agent's discount rate r. This means that there are some effort levels that only very impatient agents ever exert: a more patient agent will never work as hard no matter how large rewards are.

Building from the comparative statics in Proposition 2, I show that for y sufficiently high, we have $\rho < \rho^*$: the agent works harder than the rate maximizing amount. Lower rewards can therefore increase the rate of innovation, particularly when r is high.

Proposition 4. *The agent's quitting floor $\overline{\rho}$ is always less than ρ^*, so the agent works harder than the rate maximizing amount for sufficiently high y. Moreover, the floor $\overline{\rho}$ is strictly decreasing in r with*

$$\lim_{r\to 0}\overline{\rho}=\rho^*.$$

Proof. Taking $y \to \infty$ in part (c) of Proposition 2 implies $\overline{\rho}$ is decreasing in r. We can rewrite (4) as

$$\gamma(1-p_0)\left[1-\rho\left(1+\frac{\lambda}{r}(1-\rho^{\frac{r}{\lambda}})\right)\right]-(\lambda+r)\rho=0.$$

Taking the limit as $r \to 0$ yields (3). Since $\overline{\rho}$ is decreasing in r, this also implies that $\overline{\rho} < \rho^*$ as claimed. □

Together with Proposition 2, Proposition 4 offers a detailed account of how the rate of innovation varies with the agent's incentives. Figure 2 shows the success time μ as a function of the agent's quitting ratio. Since the agent's quitting floor $\overline{\rho}$ is lower than ρ^*, high-powered incentives can induce the agent to keep working beyond the point that maximizes the innovation rate. Fixing the discount rate r, the expected time μ needed to achieve success on some idea is a U-shaped function of an innovation's value y. When the value of success is low, the agent spends most of his time trying to come up with fresh ideas. When the value is high, the agent works too hard on each individual idea, wasting effort on dead ends.

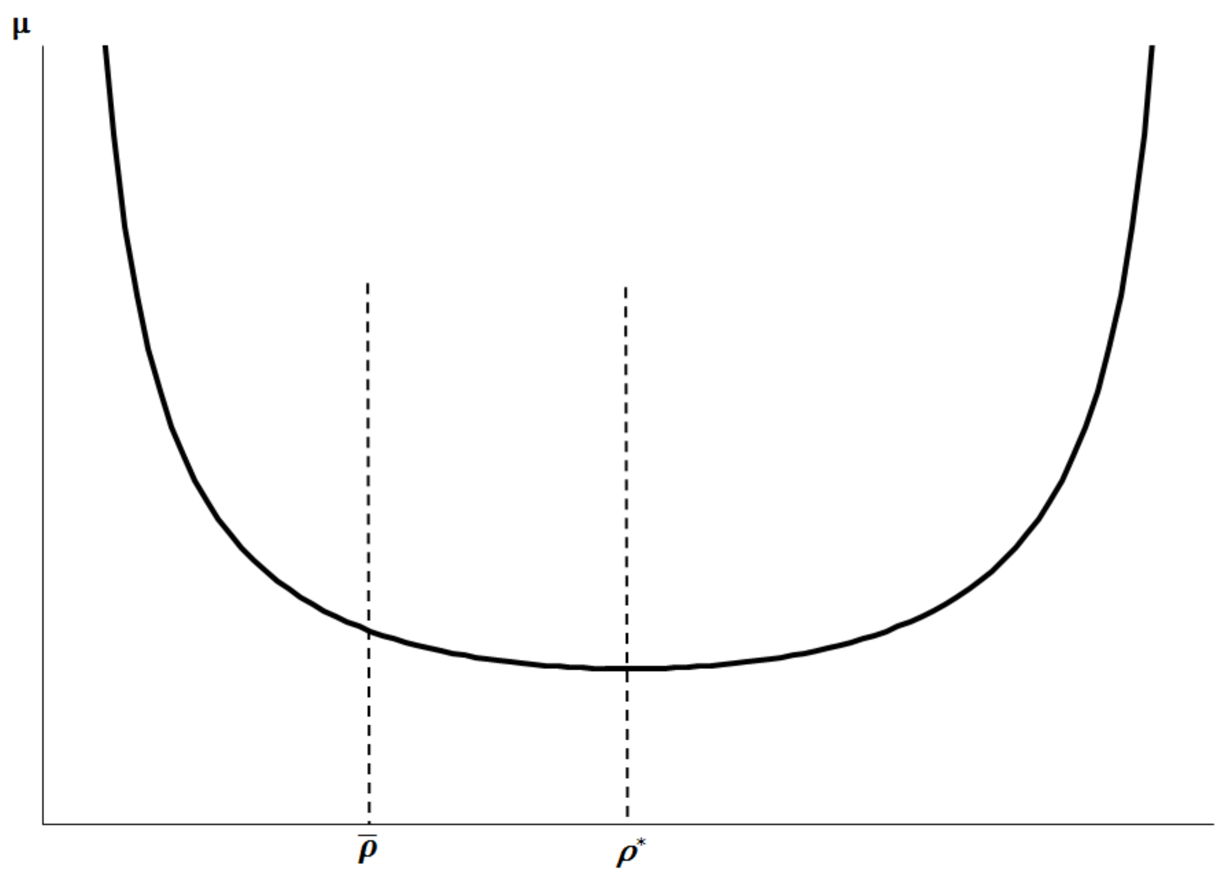

Figure 2: The expected success time as a function of the quitting ratio.

Fixing the value y, the time μ is a similarly U-shaped function of the discount rate r. When the agent is impatient, the cost of delay motivatives him to work too hard. When the agent is patient, he makes less effort and too readily gives up on each new idea. As the agent becomes more patient, it gets harder to induce a quitting ratio below ρ^*. In the limit as the agent becomes infinitely patient, the quitting floor is exactly equal to ρ^*: an infinitely patient agent with no cost of effort would choose the quitting ratio that maximizes the rate of innovation.

It may be tempting to interpret these findings as a slight tweak of the classic exploration-exploitation tradeoff, and one would then conclude that the rationale given for low-powered incentives is not so different from the one that Manso (2011) offers. However, upon closer inspection, we can see that the two accounts of low-powered incentives are fundamentally different. The story here revolves around intertemporal tradeoffs, not the value of information, and there is no question of whether the agent will innovate, only when. Note the natural analog to the action with known payoffs, the "safe arm," is the option to quit and come up with a new idea—the current idea is always less likely to succeed. In contrast with models of exploration versus exploitation, increasing the reward for success makes the agent *less inclined* to use the safe arm. While Manso studies a model without discounting, if one were to discount second period payoffs in that model, rewards in an optimal contract would need to increase. Here, an impatient agent requires less incentive to work. Moreover, the account of when weaker incentives increase the innovation rate—when the value y is high or the discount rate r is high—is a distinctive prediction of this model.

3.2 Innovation and Intellectual Property Rights

Intellectual property rights help innovators appropriate the value of their work, thereby strengthening incentives to innovate. The standard argument to limit intellectual property rights is that information is non-rival and essentially costless to replicate, so these protections are ex-post inefficient. The previous section suggests a second argument to limit protections: strong incentives may reduce innovation if they cause innovators to waste time on dead ends. The present section explores the merits of this argument.

A benevolent social planner designs a regime of intellectual property rights for our agent. Let $\alpha \in [0, 1]$ index the strength of intellectual property rights. The social value of success is given by a (weakly) decreasing function $w(\alpha)$, while the value that the agent can appropriate is a strictly increasing function $y(\alpha)$ with $y(\alpha) < w(\alpha)$, and $y(0) = 0$. Assume that w and y are continuously differentiable. We can think of α as a reduced-form representation for a collection of laws related to patents, copyrights, trade secrets, etc. Stronger protections allow the agent to obtain a larger reward, but these protections also reduce the overall social value of an innovation.

The social planner discounts the future at a rate $\tilde{r}$, not necessarily equal to the agent's discount rate r. For simplicity, I assume the social planner chooses α to maximize the expected discounted value $w(\alpha) - y(\alpha)$ of success to the rest of society. That is, the social planner ignores the value the agent derives from success as well as the cost of the agent's effort. This is a reasonable approximation if the social value of an innovation is much larger than what the agent can appropriate; moreover, the qualitative features of the planner's optimal choice are insensitive to whether we account for the agent's welfare.

Proposition 5. *The planner's payoff is*

$$\pi(\alpha) = (w(\alpha) - y(\alpha)) \frac{\lambda p_0(\gamma + \tilde{r})(1 - \rho^{1+\frac{\tilde{r}}{\lambda}})}{(\lambda + \tilde{r}) \left(\tilde{r} + \gamma(1 - \rho^{\frac{\tilde{r}}{\lambda}}) + \gamma p_0 \rho^{\frac{\tilde{r}}{\lambda}}(1 - \rho)\right)}, \tag{5}$$

where ρ is obtained by taking $y = y(\alpha)$ in (2).

Proof. Let T denote the amount of time the agent spends on each idea. Following a calculation similar to that for the agent's utility, we have

$$\begin{aligned}\pi(\alpha) &= p_0 (w(\alpha) - y(\alpha)) \int_0^T \lambda e^{-(\tilde{r}+\lambda)t} \mathrm{d}t + (1 - p_0 + p_0 e^{-\lambda T}) e^{-\tilde{r}T} \frac{\gamma}{\gamma + \tilde{r}} \pi(\alpha) \\ &= \frac{\lambda p_0}{\lambda + \tilde{r}} (w(\alpha) - y(\alpha)) (1 - e^{-(\tilde{r}+\lambda)T}) + (1 - p_0 + p_0 e^{-\lambda T}) e^{-\tilde{r}T} \frac{\gamma}{\gamma + \tilde{r}} \pi(\alpha).\end{aligned}$$

Solving this yields

$$\begin{aligned}\pi(\alpha) &= (w(\alpha) - y(\alpha)) \frac{\lambda p_0(\gamma + \tilde{r})(1 - e^{-(\tilde{r}+\lambda)T})}{(\lambda + \tilde{r}) (\tilde{r} + \gamma(1 - e^{-\tilde{r}T}) + \gamma p_0 e^{-\tilde{r}T}(1 - e^{-\lambda T}))} \\ &= (w(\alpha) - y(\alpha)) \frac{\lambda p_0(\gamma + \tilde{r})(1 - \rho^{1+\frac{\tilde{r}}{\lambda}})}{(\lambda + \tilde{r}) \left(\tilde{r} + \gamma(1 - \rho^{\frac{\tilde{r}}{\lambda}}) + \gamma p_0 \rho^{\frac{\tilde{r}}{\lambda}}(1 - \rho)\right)},\end{aligned}$$

giving (5). ☐

Taking the first order condition, the planner should choose α such that

$$0 = \pi'(\alpha) = \frac{\pi(\alpha)}{w(\alpha) - y(\alpha)} (w'(\alpha) - y'(\alpha)) + \frac{\partial \pi}{\partial \rho} \frac{\partial \rho}{\partial \alpha}. \tag{6}$$

The first term, which is strictly negative, represents the reduction in social value from stronger protections, while the second term gives the effect of stronger incentives. Proposition 2 implies that $\frac{\partial \rho}{\partial \alpha}$ is negative, so at an optimal α we must also have $\frac{\partial \pi}{\partial \rho} < 0$. As a function of ρ, the planner's payoff looks like that of an agent with no effort cost. The derivation of (4) implies that $\frac{\partial \pi}{\partial \rho} < 0$ whenever $\rho > \tilde{\rho}$, defined by

$$\frac{\gamma}{r} \tilde{\rho} \left(1 - \frac{\lambda \tilde{\rho}^{\frac{r}{\lambda}}}{\lambda + r}\right) + \frac{\tilde{\rho}}{1 - p_0} = \frac{\gamma}{\lambda + r}. \tag{7}$$

We can use this to establish limit results for the planner's optimal policy.

For each $k > 0$, define

$$w_k(\alpha) = k w(\alpha) \quad \text{and} \quad y_k(\alpha) = k y(\alpha).$$

We should think of k as a measure of an innovation's value relative to its effort cost c. High k innovations deliver signficiant value with relatively little investment. If the social value of success is given by $w_k(\alpha)$, and the value the agent appropriates is $y_k(\alpha)$, write $\pi_k(\alpha)$ for the associated planner's objective, and let α_k^* denote the planner's optimal policy. In the proof of the next result, I use $\pi(w,y)$ to denote a version of the planner's payoff function in which we replace $w(\alpha)$ and $y(\alpha)$ with w and y respectively.

Proposition 6. *We have*

(a) $\lim_{k\to\infty} \alpha_k^* = 0$.

(b) If $\tilde{r} \geq r$, *then* $\lim_{k\to\infty} y_k(\alpha_k^*) = \infty$.

(c) If $\tilde{r} < r$, *then* $\lim_{k\to\infty} y_k(\alpha_k^*) = y^* < \infty$.

Proof. We can rewrite the first order condition as

$$1 = -\frac{w_k'(\alpha) - y_k'(\alpha)}{w_k(\alpha) - y_k(\alpha)} \cdot \frac{\pi_k(\alpha)}{\frac{\partial \pi_k}{\partial \rho}\frac{\partial \rho}{\partial \alpha}} = -\frac{w_1'(\alpha) - y_1'(\alpha)}{w_1(\alpha) - y_1(\alpha)} \cdot \frac{\pi_1(\alpha)}{\frac{\partial \pi_1}{\partial \rho}\frac{\partial \rho}{\partial \alpha}}.$$

Fixing α, observe that the only terms that change with k are the partial derivatives $\frac{\partial \pi_1}{\partial \rho}$ and $\frac{\partial \rho}{\partial \alpha}$ in the denominator. For any fixed α, the value of $y_k(\alpha)$ grows without bound, so $\rho(\alpha)$ approaches the agent's quitting floor. This implies that $\frac{\partial \pi_1}{\partial \rho}$ approaches the derivative at $\overline{\rho}$. Similarly, the derivative $\frac{\partial \rho}{\partial \alpha}$ approaches zero, implying that the right hand side blows up for large k if we bound α away from zero. We conclude that α_k^* converges to zero.

If $\tilde{r} \geq r$, then $\tilde{\rho} < \overline{\rho}$, and $\frac{\partial \pi_k}{\partial \rho} > 0$ for any ρ the planner can induce the agent to choose. If we hold $y_k(\alpha) = \overline{y}$ fixed, and hence we hold ρ fixed, then for large k, the right hand side of the first order condition approaches

$$-\frac{w_1'(0) - y_1'(0)}{w_1(0) - y_1(0)} \cdot \frac{\pi_1(w, \overline{y})}{\frac{\partial \pi_1}{\partial \rho}\frac{\partial \rho}{\partial y_k}(\overline{y})\frac{\partial y_k}{\partial \alpha}(0)}.$$

The term $\frac{\partial \rho}{\partial y_k}$ evaluated at $\overline{y}$ is fixed by assumption, so the only term that varies with k is $\frac{\partial y_k}{\partial \alpha}$. This clearly grows linearly in k, implying that the right hand side of the first order condition converges to zero. For large enough k, the planner wants to choose a higher α, so we conclude that $y_k(a_k^*)$ must grow without bound.

Finally, if $\tilde{r} < r$, there exists y^* such that the agent chooses $\rho = \tilde{\rho}$ if $y_k(\alpha) = y^*$. The planner clearly never lets the agent appropriate more than y^*, and an argument essentially identical to that in the previous paragraph shows that $y_k(\alpha_k^*)$ must converge to y^*. □

Proposition 6 establishes a connection between time preferences and the optimal intellectual property regime. An impatient planner allows the agent's value to grow without bound as the social value of success increases, while a patient planner caps the value the agent can appropriate. Regardless of who is more patient, the optimal level of protection converges to zero as innovations become more valuable. This might seem counterintuitive at first, but recall that we held the effort cost fixed. Many highly valuable innovations—like new pharmaceuticals—are also very expensive to develop. In general, as the ratio $\frac{y}{c}$ increases (i.e. the social value becomes high relative to the effort cost), the optimal level of protection declines.

The proposition is valid if $w(\alpha)$ is only weakly decreasing, providing a rationale to limit intellectual property rights even if they do not decrease the ex-post social value of innovation. When the planner is more patient than the agent, and $\frac{w(\alpha)}{c}$ is high, strong protections incentivize wasted effort on dead ends. From a patient planner's perspective, it is better for the agent to give up on his ideas more quickly. Regardless of the planner's discount rate, the results in the last section imply that agents with high $\frac{y(\alpha)}{c}$ innovate more slowly as we increase α.

Our reasoning yields a novel explanation for the inverse U-shaped relationship between intellectual property rights and innovation. Suppose the policy α applies to a large population of agents with a distribution over k. An increase in α has two effects. First, potential innovators with low k become willing to invest effort, creating more innovations. On the other hand, innovators with high k innovate more slowly. As α increases, the latter effect becomes more pronounced, and the new innovators are more marginal.

This analysis suggests at least two additional predictions that call for empirical validation. First, rightward shifts in the distribution of k should correspond to lower optimal protections. We might expect less developed countries, ones with more "low-hanging fruit," to benefit less from intellectual property rights—this is broadly consistent with studies of patent laws in developing countries. Second, innovation could benefit from policies that target specific segments of the distribution of k. The most useful protections apply to innovations that require a large amount of investment relative to their social value. Outside the pharmaceutical industry, there is relatively little evidence that patents encourage innovation; the high relative cost of drug development offers an appealing explanation.

3.3 Discussion

The simplicity of exponential bandits makes the model amenable to many tweaks and extensions. Exploring the effects of small changes can help illuminate how robust key findings are. First, I note that representing the switching cost as a delay is crucial for the results on the rate of innovation. Suppose instead that the agent can switch instantaneously to a new idea after paying a "setup cost" s. This leads to fundamentally different comparative statics for the agent's behavior given a fixed reward.

In the adjusted model, switching to a new idea always leads to faster success—if $s = 0$, the agent always switches instantaneously and always works on an idea that is viable with probability p_0. The agent's decision whether to switch hinges on whether getting a success sooner (in expectation) can justify the cost s. When the value of success is higher, the agent more readily switches, contrary to the comparative statics in the model with delay. An impatient agent places *more* value on getting a success sooner, so an impatient agent switches projects *more rapidly*, also contrary to the model with delay. Success occurs more quickly when the agent switches more quickly, so impatient agents are always faster, and high-powered incentives unambiguously leads to faster innovation.

In reality, switching entails both a delay and a setup cost—new ideas require both time and resources to develop. The relative importance of each cost should determine whether high-powered incentives lead to more or less switching and faster or slower innovation. When coming up with new ideas takes time, and this delay represents the main switching cost, we should expect less frequent switching in response to higher rewards and more benefit from low-powered incentives.

Complicating the agent's decision problem can result in subtle changes. In the basic set up, the agent's decision is very simple because shirking never produces a reward. One way we can extend the analysis is to assume that there are different rates of discovery $\lambda_1 > \lambda_0 > 0$ corresponding to effort levels 1 and 0 respectively. The agent now has a more substantive choice whether to exert effort and earn his payoff faster or shirk while collecting a reward more slowly.

One observation is that the agent has a stronger direct incentive to work at more optimistic beliefs. As long as the agent is continuing on a given idea, which is currently good with probability p, the expected flow payment from effort is

$$\lambda_1 py - c,$$

while the expected flow payment from shriking is

$$\lambda_0 py.$$

Since effort is more informative about idea quality than shirking, it is clear that the agent makes effort when

$$p \geq \frac{c}{y(\lambda_1 - \lambda_0)}.$$

At pessimistic beliefs, where the expected flow payment cannot justify the cost of effort, the agent may consider shirking if the value of information is low and the cost of delay is sufficiently high.[3]

As a consequence, we get a more nuanced effect on behavior as the agent gets less patient: impatient agents still exert more effort on each idea, but they also are more likely to shirk before starting a new one. Mirroring the analysis in the proof of Proposition 1 (see the Appendix), we can obtain the belief level at which the agent abandons a project. If the agent shirks just before abandoning, this belief level is

$$\underline{p}_0 = \frac{1}{\lambda_0 y} \frac{\gamma r}{\gamma + r} u(p_0), \tag{8}$$

and if the agent exerts effort before abandoning, it is

[3]This effect is similar to intrinsic motivation as formalized by Bénabou and Tirole (2003): the agent is more willing to incur the cost of effort because he is more optimistic about the chance of success.

$$\underline{p}_1 = \frac{1}{\lambda_1 y}\left(\frac{\gamma r}{\gamma + r} u(p_0) + c\right). \tag{9}$$

Recall that $u(p_0)$ is the agent's expected payoff at the moment he starts working on a new project. There is a region in which the agent shirks if $\underline{p}_0 < \underline{p}_1$.

Observe that $u(p_0)$ converges to zero as the agent becomes more and more impatient, and both belief thresholds $\underline{p}_0$ and $\underline{p}_1$ are decreasing in r. As before, the agent spends more time working on each idea when he is impatient. However, the constant cost term in $\underline{p}_1$ implies that for sufficiently high r, we are sure to have a region in which the agent shirks. Impatient agents spend more time working on each idea, but they also spend more time shirking after they become pessimistic.

4 A Principal-Agent Model

This section studies the innovation process in an agency setting. A principal hires an agent to develop a successful idea, and the agent works according to the process described in section 2. The principal enjoys the value of success y when it arrives and designs a contract to incentivize the agent. Imagine a venture capitalist who funds an entrepreneur to implement a series of business ideas, or a firm designing incentives for research employees. I assume that the principal incurs the flow cost c, providing necessary resources for the agent's work, but the agent can choose to divert these funds for personal gain.

The principal can observe when a success arrives as well as transitions between the creative and working phases in the innovation process. The principal cannot observe the agent's effort choice during the working phase (i.e. whether the agent diverts funds). The principal discounts the future at a rate $\tilde{r}$ that is not necessarily equal to the agent's discount rate r. The agent has limited liability—the principal can never demand a payment—and all transfers must conclude the instant a success arrives.[4]

At the beginning of the game, the principal commits to a contract, and we assume the agent always accepts—equivalently, assume the agent's outside option is zero, so it is always incentive compatible to accept. A contract specifies a sequence of times, experimentation policies, and reward functions $\{T_i, x_i(t), w_i(t)\}_{i\in\mathbb{N}}$. The value T_i is the length of the ith working phase, the value $x_i(t)$ is the prescribed effort choice on the ith idea after working on it for time $t \leq T_i$, and $w_i(t)$ is an immediate payment to the agent if success occurs on the ith idea after working on it for time t. This structure is without loss of generality since any non-contingent payment does not contribute to effort incentives, and by assumption we do not allow the principal to defer payment after success arrives. Note the principal can terminate the agent's employment by setting $x_i(t) = w_i(t) = 0$ for all t and all sufficiently high i. Given the contract, the agent chooses a sequence of experimentation policies $\{\hat{x}_i(t)\}_{\substack{i\in\mathbb{N}\\ t\in[0,T_i]}}$ to maximize his discounted expected payoff.

[4]This precludes an impatient principal from reducing incentive costs by offering enormous payments at some far off date in the future.

At the start of a new working phase, we can recursively express the contract as a quadruple $\zeta = (T, x(t), w(t), u)$. The first three parts prescribe behavior in the current working phase. The value u denotes the continuation utility that the principal offers the agent starting from the *next* working phase. This captures all the information we need to analyze the agent's incentives. I write Z for the set of incentive compatible quadruples—that is, the set for which the agent's best response is to follow the prescribed effort choice $x(t)$. Each working phase functions as a stage game in a repeated interaction, and we can use methods from dynamic programming to characterize the optimal contract. A novel feature is that the principal controls the length of each stage game and therefore the discount rate applied to later stages.

4.1 Two Lemmas

The first step towards understanding optimal contracts is to compute the players' value functions. In addition to the effort level $x(t)$ at time t, we need to keep track of the agent's cumulative effort

$$X(t) = \int_0^t x(s)\mathrm{d}s.$$

Conditional on a good idea, the probability that the agent has *not* succeeded by time t is $e^{-\lambda X(t)}$. Therefore, conditional on starting the current working phase, the probability that the agent has not yet succeeded as of time t is

$$g(t) \equiv 1 - p_0 + p_0 e^{-\lambda X(t)}.$$

After time t, the probability $p(t)$ that the idea is good is

$$p(t) = \frac{p_0 e^{-\lambda X(t)}}{1 - p_0 + p_0 e^{-\lambda X(t)}} = \frac{p_0 e^{-\lambda X(t)}}{g(t)}.$$

Consequently, if the agent follows the contract $\zeta = (T, w(t), x(t), u)$, he obtains a value

$$\begin{aligned} u(\zeta) &= \int_0^T \lambda p(t) w(t) x(t) g(t) e^{-rt}\mathrm{d}t + g(T)\frac{\gamma e^{-rT}}{\gamma + r}u \\ &= \int_0^T \lambda p_0 w(t) x(t) e^{-\lambda X(t) - rt}\mathrm{d}t + g(T)\frac{\gamma e^{-rT}}{\gamma + r}u. \end{aligned} \tag{10}$$

I say that a pair (π, v) of values is **implementable** if there exists an incentive compatible contract such that the principal earns π in profit and the agent earns v. I also define the principal's **maximum profit function** $\pi^*(v)$ as the maximum profit the principal can earn through an incentive compatible contract that delivers utility v to the agent. The set of implementable values is convex, so the function $\pi^*(v)$ is concave.[5] The maximum value the

[5] I assume that a public randomization device is available, so the principal can commit to any convex combination of contracts.

principal can obtain from the contract ζ is then

$$\begin{aligned}\pi(\zeta) &= \int_0^T \left[\lambda p(t)\left(y - w(t)\right) - c\right] x(t) g(t) e^{-\tilde{r}t} \mathrm{d}t + g(T)\frac{\gamma e^{-\tilde{r}T}}{\gamma + \tilde{r}}\pi^*(u) \\ &= \int_0^T x(t) e^{-\tilde{r}t}\left[\lambda p_0 e^{-\lambda X(t)}\left(y - w(t)\right) - c g(t)\right]\mathrm{d}t + g(T)\frac{\gamma e^{-\tilde{r}T}}{\gamma + \tilde{r}}\pi^*(u). \end{aligned} \tag{11}$$

We can recursively define $\pi^*(v)$ as

$$\pi^*(v) = \max_{\substack{\zeta \in Z \\ v(\zeta) = v}} \pi(\zeta).$$

An important question is whether the principal can induce the agent to follow a given experimentation policy. We say that $w(t)$ **implements** $(T, x(t), u)$ if it is incentive compatible for the agent to follow $x(t)$ given $w(t)$ and u. Since the principal can refrain from providing the flow of resources c when $x(t) = 0$, it is clear that for $w(t)$ constant and sufficiently high, the agent is happy to follow the prescribed policy.

The real question is then: how cheaply can the principal implement $x(t)$? The main constraint is the agent's ability to delay effort. As the agent works, he becomes more pessimistic about whether the idea is good. This necessitates offering higher rewards later in the working phase to overcome the opportunity cost of diverting funds. However, doing so increases the cost of incentives early in the working phase: the higher rewards are late in the working phase, the greater the option value of delaying effort. At the end of the working phase, the agent can choose to divert current funds and earn the continuation payoff u. The reward at the end of the working phase must pay at least the value of this option, which in turn increases the cost of incentives throughout the working phase. The optimal implementation delivers the agent exactly the flow value of diverting funds plus the option value of delay.

Lemma 1 (Implementation)**.** *In an optimal contract, the reward function $w_{x,u}(t)$ satisfies*

$$w_{x,u}(t) = \frac{\gamma e^{-r(T-t)}}{\gamma + r}u + \frac{c g(t) e^{\lambda X(t)}}{\lambda p_0} + \int_t^T \frac{c g(s)}{p_0} x(s) e^{\lambda X(s) - r(s-t)} \mathrm{d}s. \tag{12}$$

Proof. See Appendix. □

Each term in the optimal reward function has a clear interpretation. The first term in equation (12) compensates the agent for the option value of diverting funds until the next working phase. The second term compensates for the instantaneous benefit of diverting funds at time t. The last term accounts for the additional option value of delaying effort within the current working phase—success today eliminates the chance to divert funds tomorrow.

Increasing the continuation payoff u directly increases the reward the principal must pay at each point in time. The third term shows that increasing experimentation has a similar effect: the principal faces increasing marginal costs of experimentation because each additional unit of the agent's effort increases the cost of incentives for earlier effort. Minimizing incentive costs means reducing the agent's continuation value, but doing so means less experimentation, which also reduces the principal's value. The principal therefore faces a tradeoff between lowering the cost of incentives early in the relationship and lowering the continuation payoff later in the relationship.

The proof of Lemma 1 uses a first order approach, checking that the agent cannot benefit from deviation at each instant. The presence of hidden information means it is not immediately obvious that this is sufficient because the agent can drive a persistent wedge between his beliefs about the idea and the principal's beliefs. Diverting funds now both gains the flow payment $c\mathrm{d}t$ and increases the likelihood of success later in the working phase. However, the agent is unable to translate this into an improvement because the reward function w front-loads his surplus: after taking discounting into account, the third term in (12) drops too quickly.

A key implication of Lemma 1 is that an impatient agent is good for the principal. When the agent becomes less patient, his continuation value shrinks. When the principal incentivizes effort later in the working phase, this translates to a lower increase in incentive costs early in the working phase. As a result, the principal can implement any experimentation policy more cheaply when the agent is less patient.

Corollary 1 (Impatience is a Virtue)**.** *Impatient agents are cheaper to incentivize: the cost of implementing any experimentation policy is decreasing in the agent's discount rate r.*

Proof. On differentiating $w_{x,u}(t)$ with respect to r, this is immediate from (12). □

Corollary 1 reflects that the main constraint on the principal's costs is the agent's willingness to wait. If the agent is less patient, the option value of waiting goes down, reducing incentive costs at present. I emphasize this is not simply a property of optimal policies: no matter what experimentation plan the principal wishes to implement, she can do so more cheaply if the agent is less patient. Regardless of the other model parameters, a principal will always prefer to hire agents who are impatient.[6]

Using Lemma 1, the agent's value in an optimal contract is

$$\begin{aligned} v(\zeta) &= \int_0^T \lambda p_0 u x(t) \frac{\gamma e^{-\lambda X(t) - rT}}{\gamma + r} + c x(t) g(t) e^{-rt} + \lambda x(t) e^{-\lambda X(t)} \int_t^T c g(s) x(s) e^{\lambda X(s) - rs} \mathrm{d}s \mathrm{d}t \\ &= \int_0^T c x(t) g(t) e^{-rt} + c x(t) g(t) e^{\lambda X(t) - rt} \int_0^t \lambda x(s) e^{-\lambda X(s)} \mathrm{d}s \mathrm{d}t + \frac{\gamma e^{-rT}}{\gamma + r} u \\ &= \int_0^T c g(t) x(t) e^{\lambda X(t) - rt} \mathrm{d}t + \frac{\gamma e^{-rT}}{\gamma + r} u. \end{aligned} \tag{13}$$

The agent earns the full flow of funds the principal provides, inflated by the factor $e^{\lambda X(t)}$, plus the discounted continuation value. Similarly, we can rewrite the principal's maximum

[6] Acemoglu et al. (2016) present evidence that firms with young managers are more innovative, stressing that a culture open to disruption plays a key role both in promoting innovation and allowing young managers to rise in the ranks. To the extent that young people are less patient, as suggested in some studies (Harrison et al., 2002; Bishai, 2004), the lower cost of providing incentives offers an alternative account of this pattern.

value as

$$\pi(\zeta) = \int_0^T x(t)e^{-\tilde{r}t}\left(\lambda p_0 y e^{-\lambda X(t)} - 2cg(t)\right)\mathrm{d}t + g(T)\frac{\gamma e^{-\tilde{r}T}}{\gamma+\tilde{r}}\pi^*(u)$$
$$- \int_0^T \lambda x(t)e^{-\lambda X(t)-(\tilde{r}-r)t}\left(\frac{p_0\gamma e^{-rT}}{\gamma+r}u + c\int_t^T g(s)x(s)e^{\lambda X(s)-rs}\mathrm{d}s\right)\mathrm{d}t. \tag{14}$$

Equation (14) already reveals an agency cost in our setting. The expected flow payment to the principal can only be positive if $y > \frac{2c}{\lambda p_0}$. It is efficient to pursue a solution as long as $y > \frac{c}{\lambda p_0}$, and indeed the self-employed agent always works in this case. Because the principal must provide the funds *and* compensate the agent for not diverting them, the principal only finances projects worth at least twice this amount. For a range of y, agency costs prevent the pursuit of profitable projects.

If the principal is not more patient than the agent, we can further simplify the problem.

Lemma 2 (No Delay)**.** *Suppose $\tilde{r} \geq r$. Prior to termination, we always have $x(t) = 1$ in an optimal contract.*

Proof. By delaying the agent's effort, the principal gains from reduced incentive costs early in the relationship, and loses value from delaying her own continuation payoff. The gains from reduced incentive costs are proportional to the reduction in the agent's continuation value. Since $\tilde{r} \geq r$, delay reduces the principal's continuation value at least in proportion to the reduction in the agent's continuation value. Instead of delaying effort at time t, the principal could instead fire the agent. This reduces the two players' continuation values at the same rate. If $\tilde{r} > r$, the principal is strictly better off firing the agent, and if $\tilde{r} = r$, the principal is indifferent between firing and delay. Hence, the principal can achieve the optimal profit in a contract with no delay in the agent's work. □

An impatient principal is always better off firing the agent than delaying his efforts. In the next subsection, this allows a relatively complete characterization of optimal contracts when the principal is not more patient than the agent.

4.2 Optimal Contracts with an Impatient Principal

In this subsection, I restrict attention to the case with $\tilde{r} \geq r$—the principal is no more patient than the agent. Under this assumption, Lemma 2 permits a simple representation of contracts and value functions. We can express a contract recursively as a pair (T, u) in which T denotes the length of the current working phase, and u is the promised continuation utility starting with the next working phase. The value functions become

$$V(T,u) = c\int_0^T g(t)e^{(\lambda-r)t}\mathrm{d}t + \frac{\gamma e^{-rT}}{\gamma+r}u, \text{ and} \tag{15}$$

$$\pi(T,u) = \int_0^T \left(\lambda p_0 y e^{-\lambda t} - 2cg(t)\right) e^{-\tilde{r}t}\mathrm{d}t + g(T)\frac{\gamma e^{-\tilde{r}T}}{\gamma+\tilde{r}}\pi^*(u)$$
$$- \int_0^T \lambda e^{-(\lambda+\tilde{r}-r)t} \left(\frac{p_0\gamma e^{-rT}}{\gamma+r}u + c\int_t^T g(s)e^{(\lambda-r)s}\mathrm{d}s\right)\mathrm{d}t, \tag{16}$$

where $g(t) = 1 - p_0 + p_0 e^{-\lambda t}$ since $X(t) = t$. Equation (15) implicitly defines a function $T(u,v)$ such that $V(T(u,v),u) = v$. The function π^* is then simply

$$\pi^*(v) = \max_u \pi\left(T(u,v),u\right). \tag{17}$$

One way to understand the dynamics of a contract is through the sequence of continuation values it offers the agent. Let u_0 denote the agent's value in the optimal contract for the principal. This value is supported through a contract (T_1, u_1), where T_1 is the length of the first working phase, and u_1 is the promised continuation value. We can likewise define (T_i, u_i) as the length and promised continuation value of the ith working phase in the optimal contract. Let $S(u)$ denote the agent continuation value that supports the value pair $(\pi^*(u), u)$. The equation $u_{i+1} = S(u_i)$ characterizes the trajectory of the principal-agent relationship over time.

Figure 3 illustrates the set of implementable value pairs (π, v). This set is necessarily convex since we allow randomization. The maximum profit function $\pi^*(v)$ traces the boundary of this set. It should be clear that in the optimal contract, the elements of the sequence $\{u_i\}$ correspond to points on the upward sloping part of the curve—if not, the principal can obtain the same profit while offering the agent a lower value, which means lower incentive costs in earlier working phases. This reveals another agency cost. The downward sloping part of the curve captures the frontier of constrained-efficient outcomes, each of which is supported through a continuation contract on the (inefficient) upward sloping part.[7]

Differentiating $v = V\left(T(u,v),u\right)$ with respect to v and u respectively gives

$$1 = \frac{\partial V}{\partial T}\frac{\partial T}{\partial v}, \qquad 0 = \frac{\partial V}{\partial T}\frac{\partial T}{\partial u} + \frac{\partial V}{\partial u}.$$

Substituting yields

$$0 = \frac{\frac{\partial T}{\partial u}}{\frac{\partial T}{\partial v}} + \frac{\partial V}{\partial u} \implies \frac{\partial T}{\partial u} = -\frac{\partial V}{\partial u}\frac{\partial T}{\partial v}.$$

The first order condition for the maximization in (17) gives

$$0 = \frac{\partial \pi}{\partial T}\frac{\partial T}{\partial u} + \frac{\partial \pi}{\partial u},$$

[7]The notion of efficiency is potentially unclear when the principal and agent have different discount rates. I say that a contract is efficient if the corresponding pair of net present values (π, v) is on the Pareto frontier over all feasible and incentive compatible contracts.

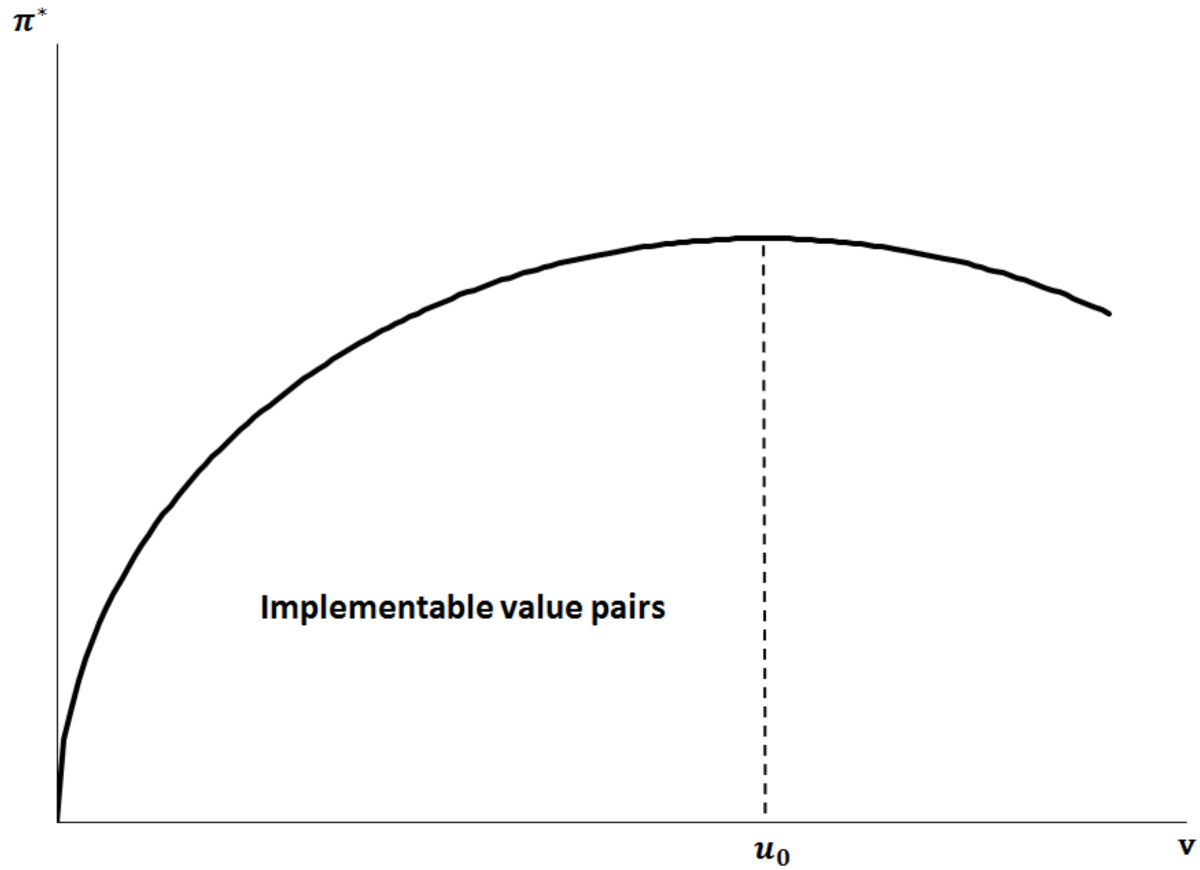

Figure 3: The set of implementable value pairs (π, v).

and substituting from above gives

$$\frac{\partial\pi}{\partial u} = \frac{\partial\pi}{\partial T}\frac{\partial T}{\partial v}\frac{\partial V}{\partial u} = \frac{\frac{\partial\pi}{\partial T}}{\frac{\partial V}{\partial T}}\frac{\partial V}{\partial u}.$$

Finally, the envelope theorem tells us that at $u = S(v)$, we have

$$(\pi^*)'(v) = \frac{\partial\pi}{\partial T}\frac{\partial T}{\partial v},$$

implying that at $u = S(v)$ we have

$$\frac{\partial\pi}{\partial u} = (\pi^*)'(v)\frac{\partial V}{\partial u}.$$

Equations (15) and (16) allow us to compute the partial derivatives in the equations

$$\frac{\partial\pi}{\partial u}\frac{\partial V}{\partial T} = \frac{\partial\pi}{\partial T}\frac{\partial V}{\partial u} \qquad \text{and} \qquad \frac{\partial\pi}{\partial u}\left(S(v)\right) = (\pi^*)'(v)\frac{\partial V}{\partial u}, \tag{18}$$

from which we can obtain the following.

Proposition 7. *Suppose $\tilde{r} \geq r$. In an optimal contract $\{(T_i, u_i)\}$:*

(a) The continuation values u_i are decreasing in i

(b) The principal eventually fires an unsuccessful agent.

Proof. We compute

$$\frac{\partial \pi}{\partial u}(S(v)) = g(T)\frac{\gamma e^{-\tilde{r}T}}{\gamma + \tilde{r}}(\pi^*)'(S(v)) - \frac{p_0 \gamma e^{-rT}}{\gamma + r}\int_0^T \lambda e^{-(\lambda + \tilde{r} - r)t}\mathrm{d}t, \qquad \frac{\partial V}{\partial u} = \frac{\gamma e^{-rT}}{\gamma + r}.$$

The second part of (18) then becomes

$$g(T)\frac{\gamma e^{-\tilde{r}T}}{\gamma + \tilde{r}}(\pi^*)'(S(v)) - \frac{p_0 \gamma e^{-rT}}{\gamma + r}\int_0^T \lambda e^{-(\lambda + \tilde{r} - r)t}\mathrm{d}t = (\pi^*)'(v)\frac{\gamma e^{-rT}}{\gamma + r},$$

which we can rewrite as

$$(\pi^*)'(S(v)) = \frac{(\gamma + \tilde{r})e^{(\tilde{r} - r)T}}{(\gamma + r)g(T)}\left((\pi^*)'(v) + p_0 \int_0^T \lambda e^{-(\lambda + \tilde{r} - r)t}\mathrm{d}t\right). \tag{19}$$

The leading factor at least one, indicating that $(\pi^*)'(S(v)) > (\pi^*)'(v)$ as long as $S(v) > 0$. Since π^* is concave, this means $S(v) < v$, proving part (a).

To prove part (b), we first note there is an upper bound on $(\pi^*)'(u)$. Since the agent always earns at least the value of the flow payment $c\mathrm{d}t$, we necessarily have

$$\frac{\pi^*(u)}{u} \leq \frac{\lambda p_0 y - 2c}{c}$$

for all u, which implies $(\pi^*)'(u) \leq \frac{\lambda p_0 y - 2c}{c}$. In fact, by considering a contract with a single working phase of length ϵ, we can see that

$$\lim_{u \to 0} \frac{\pi^*(u)}{u} \geq \frac{\lambda p_0 y - 2c}{c},$$

which then implies $(\pi^*)'(0) = \frac{\lambda p_0 y - 2c}{c}$. If $\tilde{r} > r$, then the leading factor in (19) is strictly larger than $1 + \epsilon$ for some $\epsilon > 0$ and all $T \geq 0$. The recursion for $(\pi^*)'(u_i)$ will eventually exceed $\frac{\lambda p_0 y - 2c}{c}$, implying the principal optimally terminates the agent's employment.

If $r = \tilde{r}$, equation (19) simplifies to

$$(\pi^*)'(S(v)) = \frac{1}{1 - p_0 + p_0 e^{-\lambda T}}\left((\pi^*)'(v) + p_0\left(1 - e^{-\lambda T}\right)\right).$$

By similar reasoning, if $T_i > \epsilon$ for some $\epsilon > 0$ infinitely often, then the left hand side increases without bound. Assuming that $\lim_{i \to \infty} T_i = 0$, we necessarily have $\lim_{i \to \infty} u_i = 0$. We now make use of the first part of (18). We compute the remaining partial derivatives, assuming $\tilde{r} = r$, as

$$\frac{\partial \pi}{\partial T} = \left(\lambda p_0 y e^{-\lambda T} - 2cg(T)\right) e^{-rT} - \frac{\lambda p_0 \gamma u}{\gamma + r} e^{-(\lambda + r)T}$$
$$- \left(\lambda p_0 e^{-\lambda T} + rg(T)\right) \frac{\gamma e^{-rT}}{\gamma + r} \pi^*(u)$$
$$+ \left(\frac{r p_0 u \gamma e^{-rT}}{\gamma + r} - cg(T) e^{(\lambda - r)T}\right) \int_0^T \lambda e^{-\lambda t} \mathrm{d}t, \quad \text{and}$$

$$\frac{\partial V}{\partial T} = cg(T) e^{(\lambda - r)T} - \frac{r u \gamma e^{-rT}}{\gamma + r}.$$

Substituting and simplifying, equation (18) becomes

$$g(T)(\pi^*)'(u) \left(cg(T) e^{\lambda T} - \frac{r u \gamma}{\gamma + r}\right) + \left(\lambda p_0 e^{-\lambda T} + rg(T)\right) \frac{\gamma}{\gamma + r} \pi^*(u) + \frac{\lambda p_0 \gamma u}{\gamma + r} e^{-\lambda T}$$
$$= \left(\lambda p_0 y e^{-\lambda T} - 2cg(T)\right) - c(1 - p_0) g(T) e^{\lambda T} \int_0^T \lambda e^{-\lambda t} \mathrm{d}t.$$

Defining $\tilde{g}(t) = g(t) e^{\lambda t}$, we can further simplify this to

$$(\lambda p_0 y - c)\, \tilde{g}(T)^2 + \left(c + \frac{r\gamma}{\gamma + r} \left(\pi^*(u) - u(\pi^*)'(u)\right)\right) \tilde{g}(T)$$
$$+ \frac{\lambda p_0 \gamma}{\gamma + r} \left(\pi^*(u) + u\right) - \lambda p_0 y = 0. \tag{20}$$

If the principal never fires the agent, this always holds with equality. In the limit as u approaches zero, this approaches

$$(\lambda p_0 y - c) \tilde{g}(T)^2 + c\tilde{g}(T) - \lambda p_0 y = 0,$$

with solution

$$\tilde{g}(T) = \frac{\lambda p_0 y}{\lambda p_0 y - c}.$$

This implies that

$$\lim_{i \to \infty} T_i = \ln\left[\frac{1}{1 - p_0} \left(\frac{\lambda p_0 y}{\lambda p_0 y - c} - p_0\right)\right] = \ln\left[\frac{\lambda p_0 y}{\lambda p_0 y - c} + \frac{c p_0}{(1 - p_0)(\lambda p_0 y - c)}\right] > 0,$$

contradicting $\lim_{i \to \infty} T_i = 0$. We conclude that the principal fires the agent.

□

Proposition 7 shows that the continuation values associated with each new working phase are monotonically decreasing, and the principal eventually terminates the agent's employment if no success arrives. An optimal contract thus front-loads the agent's effort and becomes less productive over time. Since the principal eventually fires the agent, with positive probability we never reach success. Hence, with an impatent principal, agency costs appear in three ways: some worthwhile projects are not pursued, effort is inefficiently front-loaded, and sometimes we quit without a success.

The proof of the proposition also offers a way to explicitly compute the number of working phases in an optimal contract, along with other properties of $\pi^*(v)$. Working backwards from the point at which the agent is fired, we can define a sequence of agent values $v_1 < v_2 < \ldots$ such that an optimal contract delivering $v \in (v_{k-1}, v_k]$ to the agent uses k working phases. We can recursively compute these thresholds via

$$v_k = V(t_{k-1}, v_{k-1}), \quad \pi^*(v_k) = \pi(t_{k-1}, v_{k-1}), \quad (\pi^*)'(v_k) = \frac{\frac{\partial \pi}{\partial T}}{\frac{\partial V}{\partial T}}(t_{k-1}, v_{k-1}).$$

The working phase length t_{k-1} can be computed from v_{k-1} using the first part of (18). In the special case in which $\tilde{r} = r$, we can solve (20) to obtain

$$\tilde{g}(t_k) = \frac{-\left(c + \frac{r\gamma}{\gamma+r}\left(\pi^*(v_k) - v_k(\pi^*)'(v_k)\right)\right)}{2(\lambda p_0 y - c)} + \frac{\sqrt{\left(c + \frac{r\gamma}{\gamma+r}\left(\pi^*(v_k) - v_k(\pi^*)'(v_k)\right)\right)^2 + 4\lambda p_0\left(\lambda p_0 y - c\right)\left(y - \frac{\gamma}{\gamma+r}(\pi^*(v_k) + v_k)\right)}}{2(\lambda p_0 y - c)}.$$

Following this recursion, when we reach k such that $(\pi^*)'(v_k) < 0$, then k is the number of working phases in an optimal contract.

4.3 Patience, Delay, and Tenure

This subsection addresses the case in which $\tilde{r} < r$. With a patient principal, analysis becomes more difficult because delaying the agent's effort is a potentially useful way to reduce incentive costs.

Proposition 8. *For v sufficiently close to zero, the contract $(T, x(t), u)$ that achieves $\pi^*(v)$ sets $x(t) = 0$ for some t.*

Proof. See Appendix. □

Delay reduces the agent's continuation value, thereby reducing earlier incentive costs. Of course, delay also reduces the principal's continuation value, but at a lower rate. For some promised continuation values, it is optimal for the principal to request a delay, so we see a new form of inefficiency that fails to appear when $\tilde{r} \geq r$.

On the other hand, with a patient principal, there is no reason to fire the agent. Given an implementable value pair (π, v), we can also implement any pair $(\pi e^{-\tilde{r}t}, ve^{-rt})$ through delay. Taking t sufficiently large, this implies that the ratio

$$\frac{\pi^*(v)}{v}$$

is unbounded near $v = 0$. This means the principal can always find a continuation contract that is better than firing the agent.

Proposition 9. *Suppose $\tilde{r} < r$. In an optimal contract $\{(T_i, u_i)\}$:*

(a) The continuation values u_i are bounded away from zero: the principal never fires the agent;

(b) Success is reached with probability one.

Proof. See Appendix. □

Proposition 9 highlights another important distinction in the contract structure depending on who is more patient. The contract still exhibits front-loading, but the agent has a guaranteed lower bound on his continutation utility, no matter how many working phases pass without a success. With a patient principal, optimal contracts resemble a tenure arrangement. In contrast with the last section, we are sure to obtain a success eventually.

4.4 Discussion

A simple extension allows us to study how adverse selection affects optimal contracts. Suppose there are two types of agents, a high type as above and a low type that can never succeed. The agent knows his type, but the principal only has some prior. As work progresses without a success, the principal becomes more pessimistic about the agent's type. Since the low type never succeeds, the pricipal writes a contract to incentivize the high type as before, but chooses a time to fire the agent regardless of the relative discount rates. This stopping time is naturally earlier than it would be without the low type, resulting in more inefficient front-loading and a decrease in the high type's reward relative to the model without adverse selection.

In many cases, it seems natural that a principal should observe on *what* the agent is working but not the level of effort or the diversion of resources—entrepreneurs pitch their ideas to their supporters, managers talk to researchers and ask for updates. Still, we might ask what happens if we relax this assumption. Doing so affords the agent another potentially profitable deviation. Instead of working until the end of a planned working phase, the agent could quit early and get a head start on coming up with a new idea. This brings the next working phase forward in time and increases the continuation payoff. The principal would need to compensate the agent for the value of this option, and it seems likely that this would exacerbate front-loading in an optimal contract. However, there are two complications that make analysis of such a model less than straightforward. First, since the agent cannot actually start working on the new idea until the previous planned working phase concludes, valuing this option is a difficult exercise. Second, if the agent cheats multiple times, the principal may infer this from observing creative phases that are too short.

5 Final Remarks

Dead ends are a pervasive and important feature of the innovative process. The same problem often has multiple solutions, and not every idea a researcher tries will prove successful. The implicit assumption that there is only one path to success is a key limitation in the existing experimentation literature. Including this feature helps us understand how incentives affect the rate of innovation, and it reveals a novel agency problem that is unique to creative workers.

One contribution is a new rationale for why low-powered incentives encourage creativity. The choice between continuing with the current idea and incurring delay to think up a new one is fundamentally distinct from the more classical exploration-exploitation tradeoff. In the latter case, strong incentives discourage risk taking for fear of failure. Here, strong incentives lead the agent to spend more time on ideas that are more likely to fail: the agent is reluctant to change course due to the cost of delay. Lower rewards can lead to faster innovation because the agent is more willing to give up and try something new. This effect carries implications for the design of intellectual property protections. We can explain the inverse U relationship between intellectual property rights and innovative output and why increased R&D spending may not lead to increased innovation.

In a principal-agent model, the ability of the agent to wait and come up with a new idea leads to higher incentive costs. When the agent can earn a high continuation value starting from his next idea, it becomes more expensive to incentivize effort on the current idea. This gives the principal a reason to reduce the productivity of continuation contracts: the agent's effort is front-loaded, and the relationship becomes progressively less productive over time. The structure of optimal contracts depends crucially on who is more patient. An impatient principal gives the agent a deadline and fires him if success takes too long. A patient principal grants the agent tenure, but may ask him to take breaks in order to reduce incentive costs.

The tractability of this framework makes it suitable for a variety of applications. For instance, we can ask how dead ends might impact contest design. Innovation contests may specify a particular goal (e.g. a clock that keeps time within a given accuracy at sea), but there are usually many possible ways to approach that goal. A contestant must continually decide whether to keep going on the current approach or to try a different idea. We could also ask how dead ends impact screening contracts in a richer setting than what I discussed. Another promising avenue is to explore a market for creative talent in which agents can move between different employers.

References

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William Kerr (2013), “Innovation, Reallocation, and Growth.” NBER Working Paper No. 18993.

Acemoglu, Daron, Ufuk Akcigit, and Murat Alp Celik (2016), “Young, Restless, and Creative: Openness to Disruption and Creative Innovations.” Working Paper.

Akcigit, Ufuk and Qingmin Liu (2016), “The Role of Information in Innovation and Competition.” *Journal of the European Economic Association*, 14, 828–870.

Azoulay, Pierre, Joshua Graff Zivin, and Gustavo Manso (2011), “Incentives and Creativity: Evidence from the Academic Life Sciences.” *RAND Journal of Economics*, 42, 527–554.

Bénabou, Roland and Jean Tirole (2003), “Intrinsic and Extrinsic Motivation.” *Review of Economic Studies*, 70, 489–520.

Bergemann, Dirk and Ulrich Hege (2005), “The Financing of Innovation: Learning and Stopping.” *RAND Journal of Economics*, 36, 719–752.

Bishai, David (2004), “Does Time Preference Change with Age?” *Journal of Population Economics*, 17, 583–602.

Byun, Seong (2015), “Incentivizing Innovation Under Ambiguity.” Working Paper.

Chang, Xin, Kangkang Fu, Angie Low, and Wenrui Zhang (2015), “Non-Executive Employee Stock Options and Corporate Innovation.” *Journal of Financial Economics*, 115, 168–188.

Deci, Edward, Richard Koestner, and Richard Ryan (1999), “A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation.” *Psychological Bulletin*, 125, 627–668.

Demarzo, Peter and Yuliy Sannikov (2006), “Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model.” *Journal of Finance*, 61, 2681–2724.

Di Tella, Sebastian and Yuliy Sannikov (2016), “Optimal Asset Management Contracts with Hidden Savings.” Working Paper.

Ederer, Florian and Gustavo Manso (2013), “Is Pay-for-Performance Detrimental to Innovation?” *Management Science*, 59, 1496–1513.

Francis, Bill, Iftekhar Hasan, and Zenu Sharma (2011), “Incentives and Innovation: Evidence from CEO Compensation Contracts.” Working Paper.

Furukawa, Yuichi (2010), “Intellectual Property Protection and Innovation: an Inverted-U Relationship.” *Economics Letters*, 109, 99–101.

Garrett, Robert (2010), "Does Employee Ownership Increase Innovation?" *New England Journal of Entrepreneurship*, 13, 37–46.

Gilhooly, Kenneth (2016), "Incubation and Intuition in Creative Problem Solving." *Frontiers in Psychology*, 7.

Gould, David and William Gruben (1996), "The Role of Intellectual Property Rights in Economic Growth." *Journal of Development Economics*, 48, 323–350.

Guo, Yingni (2016), "Dynamic Delegation of Experimentation." *American Economic Review*, 106, 1969–2008.

Halac, Marina, Navin Kartik, and Qingmin Liu (2016a), "Contests for Experimentation." Forthcoming in Journal of Political Economy.

Halac, Marina, Navin Kartik, and Qingmin Liu (2016b), "Optimal Contracts for Experimentation." Forthcoming in Review of Economic Studies.

Harrison, Glenn, Morten Lau, and Melonie Williams (2002), "Estimating Individual Discount Rates in Denmark: A Field Experiment." *American Economic Review*, 92, 1606–1617.

Hélie, Sébastien and Ron Sun (2010), "Incubation, Insight, and Creative Problem Solving: A Unified Theory and a Connectionist Model." *Psychological Review*, 117, 994–1024.

Hörner, Johannes and Larry Samuelson (2013), "Incentives for Experimenting Agents." *RAND Journal of Economics*, 44, 632–663.

Horowitz, Andrew and Edwin Lai (1996), "Patent Length and the Rate of Innovation." *International Economic Review*, 37, 785–801.

Jones, Charles (1995), "Time Series Tests of Endogenous Growth Models." *Quarterly Journal of Economics*, 110, 495–525.

Kanwar, Sunil and Robert Evenson (2003), "Does Intellectual Property Protection Spur Technological Change?" *Oxford Economic Papers*, 55, 235–264.

Keller, Godfrey and Sven Rady (2015), "Breakdowns." *Theoretical Economics*, 10, 175–202.

Keller, Godfrey, Sven Rady, and Martin Cripps (2005), "Strategic Experimentation with Exponential Bandits." *Econometrica*, 73, 39–68.

Klette, Tor and Samuel Kortum (2004), "Innovating Firms and Aggregate Innovation." *Journal of Political Economy*, 112, 986–1018.

Lentz, Rasmus and Dale Mortensen (2008), "An Empirical Model of Growth through Product Innovation." *Econometrica*, 76, 1317–1373.

Lerner, Josh (2009), "The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues." *American Economic Review: Papers and Proceedings*, 99, 343–348.

Lerner, Josh and Bronwyn Hall (2010), "The Financing of R&D and Innovation." In *Handbook of the Economics of Innovation* (Bronwyn Hall and Nathan Rosenberg, eds.), volume 1, Elsevier.

Lerner, Josh and Julie Wulf (2007), "Innovation and Incentives: Evidence from Corporate R&D." *Review of Economics and Statistics*, 89, 634–644.

Lipnowski, Elliot and Joao Ramos (2015), "Repeated Delegation." Working Paper.

Manso, Gustavo (2011), "Motivating Innovation." *Journal of Finance*, 66, 1823–1860.

Moroni, Sofia (2015), "Experimentation in Organizations." Working Paper.

Opp, Marcus and John Zhu (2015), "Impatience versus Incentives." *Econometrica*, 83, 1601–1617.

Prat, Julien and Boyan Jovanovic (2014), "Dynamic Contracts when the Agent's Quality is Unknown." *Theoretical Economics*, 9, 865–914.

Qian, Yi (2007), "Do National Patent Laws Stimulate Domestic Innovation in a Global Patenting Environment? A Cross-Country Analysis of Pharmaceutical Patent Protection." *Review of Economics and Statistics*, 89, 436–453.

Sannikov, Yuliy (2008), "A Continuous-Time Version of the Principal-Agent Problem." *Review of Economic Studies*, 75, 957–984.

Sio, Ut Na and Elisabeth Rudowicz (2007), "The Role of an Incubation Period in Creative Problem Solving." *Creativity Research Journal*, 19, 307–318.

Tian, Xuan and Tracy Wang (2014), "Tolerance for Failure and Corporate Innovation." *Review of Financial Studies*, 27, 211–255.

Yanadori, Yoshio and Janet Marler (2006), "Compensation Strategy: Does Business Strategy Influence Compensation in High-Technology Firms?" *Strategic Management Journal*, 27, 559–570.

Zhu, John (2013), "Optimal Contracts with Shirking." *Review of Economic Studies*, 80, 812–839.

A Appendix

Proof of Proposition 1

We derive a Bellman equation to characterize the agent's behvaior. Write $u(p)$ for the agent's continuation payoff when currently working on an idea that is good with probability p, and write τ for the exponentially distributed time spent in the creative phase. Since the agent always has the option to incur delay and come up with a new idea, we have for all p that

$$u(p) \geq \mathbb{E}\left[e^{-r\tau}\right] u(p_0) = \frac{\gamma}{\gamma + r} u(p_0).$$

This will hold with equality at the abandonment belief $\underline{p}$.

During a "period" dt, the agent succeeds with probability $\lambda \mathrm{d}t$ if the idea is good, and with probability zero otherwise. In the absence of success, the agent updates her belief to

$$p + \mathrm{d}p = \frac{p(1 - \lambda \mathrm{d}t)}{1 - p + p(1 - \lambda \mathrm{d}t)}.$$

Equivalently, the belief change is $\mathrm{d}p = -\lambda p(1-p)\mathrm{d}t$. At the same time, the agent incurs a flow cost $c\mathrm{d}t$ of effort. The value function therefore satisfies

$$u(p) = \max\left\{\frac{\gamma}{\gamma + r} u(p_0),\ -c\mathrm{d}t + e^{-r\mathrm{d}t}\mathbb{E}[u(p + \mathrm{d}p)\,|\,p]\right\}.$$

We split the continuation value $\mathbb{E}[u(p + \mathrm{d}p)\,|\,p]$ into two parts. The agent succeeds with probability $\lambda p \mathrm{d}t$, yielding a continuation payoff of y. Otherwise, the continuation payoff is $u(p) + u'(p)\mathrm{d}p$. Taking $e^{-r\mathrm{d}t} \approx 1 - r\mathrm{d}t$, when the agent continues working we have

$$\begin{aligned} u(p) &= -c\mathrm{d}t + (1 - r\mathrm{d}t)\left(\lambda p y \mathrm{d}t + (1 - \lambda p \mathrm{d}t)(u(p) + u'(p)\mathrm{d}p)\right) \\ &\approx (\lambda p y - c)\mathrm{d}t + \left(1 - (\lambda p + r)\mathrm{d}t\right) u(p) - \lambda p(1-p)u'(p)\mathrm{d}t, \end{aligned}$$

where we have removed higher order terms. Rearranging, we arrive at the Bellman equation

$$u(p) = \max\left\{\frac{\gamma}{\gamma + r} u(p_0),\ \frac{1}{\lambda p + r}\left(\lambda p\left[y - (1-p)u'(p)\right] - c\right)\right\}.$$

Observe that the flow payoff can be positive if and only if $\lambda p_0 y > c$, implying the first claim.

As long as the agent continues with the current idea, the value function satisfies the differential equation

$$u'(p) + \frac{r + \lambda p}{\lambda p(1-p)} u(p) = \frac{y}{1-p} - \frac{c}{\lambda p(1-p)}.$$

We can solve this differential equation using the integrating factor

$$g(p) = e^{\int \frac{r + \lambda p}{\lambda p(1-p)}\mathrm{d}p} = \left(\frac{p}{1-p}\right)^{\frac{r}{\lambda}} \frac{1}{1-p}.$$

Solutions of this differential equation take the form

$$\begin{aligned} u(p) &= \frac{1}{g(p)} \int g(p) \left(\frac{y}{1-p} - \frac{c}{\lambda p(1-p)} \right) \mathrm{d}p + \frac{C}{g(p)} \\ &= \frac{\lambda p}{\lambda + r} \left(y + \frac{c}{r} \right) - \frac{c}{r} + C(1-p) \left(\frac{1-p}{p} \right)^{\frac{r}{\lambda}}, \end{aligned}$$

where C is an arbitrary constant.

To find C and the abandonment belief $\underline{p}$, we impose the boundary conditions $u(\underline{p}) = \frac{\gamma}{\gamma+r} u(p_0)$ and $u'(\underline{p}) = 0$. We compute the derivative

$$u'(p) = \frac{\lambda}{\lambda + r} \left(y + \frac{c}{r} \right) - C \frac{r + \lambda p}{\lambda p} \left(\frac{1-p}{p} \right)^{\frac{r}{\lambda}}.$$

The condition $u'(\underline{p}) = 0$ gives

$$C = \left(\frac{\lambda}{\lambda + r} \right) \left(\frac{\lambda \underline{p}}{\lambda \underline{p} + r} \right) \left(y + \frac{c}{r} \right) \left(\frac{\underline{p}}{1 - \underline{p}} \right)^{\frac{r}{\lambda}}.$$

Substituting this into $u(\underline{p}) = \frac{\gamma}{\gamma+r} u(p_0)$ gives

$$\begin{aligned} \frac{\gamma}{\gamma + r} u(p_0) = &\frac{\lambda \underline{p}}{\lambda + r} \left(y + \frac{c}{r} \right) - \frac{c}{r} \\ &+ \left(\frac{\lambda}{\lambda + r} \right) \left(\frac{\lambda \underline{p}(1 - \underline{p})}{\lambda \underline{p} + r} \right) \left(y + \frac{c}{r} \right). \end{aligned}$$

Multiplying through by $\lambda \underline{p} + r$, we find the quadratic term drops out, and we can solve for

$$\underline{p} = \frac{\gamma r u(p_0) + (\gamma + r)c}{\lambda \left((\gamma + r)y - \gamma u(p_0) \right)},$$

which then implies

$$\frac{\underline{p}}{1 - \underline{p}} = \frac{p_0}{1 - p_0} \rho = \frac{\gamma r u(p_0) + (\gamma + r)c}{(\gamma + r)(\lambda y - c) - \gamma(\lambda + r)u(p_0)}. \tag{21}$$

To complete the proof, we must compute $u(p_0)$. The expected payoff is the probability of finding a breakthrough in the current project times the discounted payoff from a breakthrough, plus the probability of not finding a breakthrough times the continuation payoff from starting over, less the accumulated flow cost of effort. Let T denote the amount of time the agent spends on any research project. From Bayes' rule, we have $e^{-\lambda T} = \frac{\underline{p}(1-p_0)}{p_0(1-\underline{p})} = \rho$.

We calculate $u(p_0)$ as

$$\begin{aligned}
u(p_0) = {} & p_0 y \int_0^T \lambda e^{-(r+\lambda)t} \mathrm{d}t + \left(1 - p_0 + p_0 e^{-\lambda T}\right) e^{-rT} \frac{\gamma}{\gamma + r} u(p_0) \\
& - p_0 \int_0^T \lambda e^{-\lambda t} \left(\frac{c}{r}(1 - e^{-rt})\right) dt - \left(1 - p_0 + p_0 e^{-\lambda T}\right) \frac{c}{r}(1 - e^{-rT}) \\
= {} & \left(p_0 e^{-(r+\lambda)T} + (1 - p_0) e^{-rT}\right) \frac{\gamma}{\gamma + r} u(p_0) + p_0 \frac{\lambda y}{\lambda + r} \left(1 - e^{-(r+\lambda)T}\right) \\
& - \frac{c}{r} \left(\frac{p_0 r}{\lambda + r}(1 - e^{-(r+\lambda)T}) + (1 - p_0)(1 - e^{-rT})\right)
\end{aligned}$$

Solving and substituting ρ for $e^{-\lambda T}$ gives

$$u(p_0) = \frac{(\gamma + r)\left(p_0 r(1 - \rho^{1+\frac{r}{\lambda}})(\lambda y - c) - c(\lambda + r)(1 - p_0)(1 - \rho^{\frac{r}{\lambda}})\right)}{r(\lambda + r)\left(r + \gamma\left((1 - p_0)(1 - \rho^{\frac{r}{\lambda}}) + p_0(1 - \rho^{1+\frac{r}{\lambda}})\right)\right)}.$$

From (21) we now have

$$\frac{p_0}{1 - p_0}\rho = \frac{r\left(\lambda\gamma p_0(ry + c)(1 - \rho^{1+\frac{r}{\lambda}}) + cr(\lambda + r)\right)}{(\lambda + r)\left(\lambda\gamma(ry + c)(1 - p_0)(1 - \rho^{\frac{r}{\lambda}}) + r^2(\lambda y - c)\right)}.$$

Multiplying through by the denominator and simplifying gives our result:

$$\frac{\gamma}{r}\rho\left(1 - \frac{\lambda\rho^{\frac{r}{\lambda}}}{\lambda + r}\right) + r\rho\frac{\lambda y - c}{\lambda(ry + c)(1 - p_0)} = \frac{\gamma}{\lambda + r} + \frac{cr}{p_0\lambda(ry + c)}.$$

If we differentiate the left hand side with respect to ρ, we obtain

$$\frac{\gamma}{r}\left(1 - \rho^{\frac{r}{\lambda}}\right) + r\frac{\lambda y - c}{\lambda(ry + c)(1 - p_0)} > 0.$$

Therefore, the left hand side is strictly increasing in ρ. For $\rho \in (0, 1)$, the left hand side ranges from zero to

$$\frac{\gamma}{\lambda + r} + r\frac{\lambda y - c}{\lambda(ry + c)(1 - p_0)} > \frac{\gamma}{\lambda + r} + r\frac{\frac{c}{p_0} - c}{\lambda(ry + c)(1 - p_0)} = \frac{\gamma}{\lambda + r} + \frac{cr}{p_0\lambda(ry + c)},$$

implying there is a unique solution. □

Proof of Proposition 2

We derive each of these results by implicitly differentiating (2) with respect to the variable in question. Starting with the cost c, we multiply both sides by $ry + c$ and differentiate to obtain

$$\frac{\partial\rho}{\partial c}\left(\frac{\gamma(ry + c)}{r}(1 - \rho^{\frac{r}{\lambda}}) + r\frac{\lambda y - c}{\lambda(1 - p_0)}\right) + \frac{\gamma}{r}\rho\left(1 - \frac{\lambda\rho^{\frac{r}{\lambda}}}{\lambda + r}\right) - \frac{\rho}{\lambda(1 - p_0)} = \frac{\gamma}{\lambda + r} + \frac{r}{p_0\lambda}.$$

Solving gives

$$\frac{\partial \rho}{\partial c} = \frac{\frac{\rho}{\lambda(1-p_0)} + \frac{r}{\lambda p_0} + \frac{\gamma}{r(\lambda+r)}\left(r - (\lambda+r)\rho + \lambda\rho^{1+\frac{r}{\lambda}}\right)}{\frac{r(\lambda y - c)}{\lambda(1-p_0)} + \frac{\gamma}{r}(ry+c)(1-\rho^{\frac{r}{\lambda}})}.$$

The denominator is clearly positive, and the numerator is positive if

$$r - (\lambda+r)\rho + \lambda\rho^{1+\frac{r}{\lambda}} > 0. \tag{22}$$

Taking a first derivative shows the left hand side in (22) is strictly decreasing in ρ, taking the value 0 at $\rho = 1$, so (22) holds, and $\frac{\partial \rho}{\partial c} > 0$ as desired. The calculation to show $\frac{\partial \rho}{\partial y} < 0$ is essentially identical, and I leave it as an exercise.

Proceeding similary, we differentiate (2) with respect to p_0 to obtain

$$\frac{\partial \rho}{\partial p_0}\left(\frac{\gamma}{r}(1-\rho^{\frac{r}{\lambda}}) + r\frac{\lambda y - c}{\lambda(ry+c)(1-p_0)}\right) = -\left(r\rho\frac{\lambda y - c}{\lambda(ry+c)(1-p_0)^2} - \frac{cr}{p_0^2\lambda(ry+c)}\right),$$

which immediately implies $\frac{\partial \rho}{\partial p_0} < 0$. Differentiating with respect to γ gives

$$\frac{\partial \rho}{\partial \gamma}\left(\frac{\gamma}{r}(1-\rho^{\frac{r}{\lambda}}) + r\frac{\lambda y - c}{\lambda(ry+c)(1-p_0)}\right) = -\frac{\rho}{r}\left(1 - \frac{\lambda\rho^{\frac{r}{\lambda}}}{\lambda+r}\right) + \frac{1}{\lambda+r}.$$

The inequality (22) implies the right hand side is positive, so $\frac{\partial \rho}{\partial \gamma} > 0$. Differetiating with respect to λ gives

$$\begin{aligned}\frac{\partial \rho}{\partial \lambda}\left(\frac{\gamma}{r}(1-\rho^{\frac{r}{\lambda}}) + r\frac{\lambda y - c}{\lambda(ry+c)(1-p_0)}\right) &= -\frac{cr}{\lambda^2(ry+c)}\left(\frac{1}{p_0} + \frac{\rho}{1-p_0}\right) \\ &\quad + \frac{\gamma}{\lambda(\lambda+r)^2}\left(\rho^{1+\frac{r}{\lambda}}\left(\lambda - (\lambda+r)\ln\rho\right) - \lambda\right).\end{aligned}$$

Using that $\ln\rho \geq \frac{\rho-1}{\rho}$, the term on the second line is bounded above by

$$\frac{\gamma}{\lambda(\lambda+r)^2}\left((\lambda+r)\rho^{\frac{r}{\lambda}} - r\rho^{1+\frac{r}{\lambda}} - \lambda\right).$$

If we flip λ and r in (22), the inequality still holds by the same argument, implying this term is negative, and hence $\frac{\partial \rho}{\partial \lambda} < 0$.

Finally, multiply both sides of (2) by $y + \frac{c}{r}$ and differentiate with respect to r to obtain

$$\begin{aligned}&\frac{\partial \rho}{\partial r}\left(\frac{\gamma(ry+c)}{r^2}(1-\rho^{\frac{r}{\lambda}}) + \frac{\lambda y - c}{\lambda(1-p_0)}\right) \\ &\quad = -\frac{\gamma\left(y+\frac{c}{r}\right)}{r^2(\lambda+r)^2}\left(r^2 - (\lambda+r)^2\rho + \lambda(\lambda+2r)\rho^{1+\frac{r}{\lambda}} + r(\lambda+r)\rho^{1+\frac{r}{\lambda}}\ln\rho\right) \\ &\qquad - \frac{\gamma c}{r^3(\lambda+r)}\left(r - (\lambda+r)\rho + \lambda\rho^{1+\frac{r}{\lambda}}\right)\end{aligned}$$

The second term on the right is negative by (22). To complete the proof, we show that

$$r^2 - \rho(\lambda + r)^2 + \lambda\rho^{1+\frac{r}{\lambda}}(\lambda + 2r) - r(\lambda + r)\rho^{1+\frac{r}{\lambda}} \ln \rho$$

is positive for $\rho \in (0, 1)$. Differentiating this expression with respect to ρ gives

$$(\lambda + r)^2 \left(\rho^{\frac{r}{\lambda}} \left(1 - \frac{r}{\lambda} \ln \rho\right) - 1\right),$$

and differentiating a second time yields

$$-\frac{r^2}{\lambda^2}(\lambda + r)^2 \rho^{\frac{r}{\lambda}-1} \ln \rho > 0.$$

The expression and its first derivative are both zero at $\rho = 1$. Since the second derivative is positive, the first derivative is negative on $(0, 1)$, and hence the function is positive on $(0, 1)$ as desired. □

Proof of Lemma 1

I first derive the reward function $w(t)$ using a first order approach, and I subsequently check that the agent cannot improve through multiple deviations. The net benefit from effort at the instant t is

$$\left(\lambda p_0 e^{-\lambda X(t)} w(t) - cg(t)\right) e^{-rt} \mathrm{d}t, \tag{23}$$

the expected value of success less the funds the agent could divert. Besides the immediate gain from diverting funds, shirking delivers benefits that arrive later. In the current working phase, there is an increased chance of success at times $s > t$, since $p(s)$ is higher than it would have been. At the end of the working phase, there is an increased continuation payoff, since the chance of success in the current working phase goes down. Differentiating the expected reward with respect to cumulative effort X at time $s > t$ gives

$$-\lambda^2 p_0 w(s) x(s) e^{-\lambda X(s) - rs} \mathrm{d}s \mathrm{d}t,$$

and differentiating the continuation payoff with respect to cumulative effort gives

$$-\lambda \frac{p_0 \gamma}{\gamma + r} e^{-\lambda X(T) - rT} u \mathrm{d}t.$$

The delayed benefits of shirking at time t are then

$$\lambda \mathrm{d}t \left(\int_t^T \lambda p_0 w(s) x(s) e^{-\lambda X(s) - rs} \mathrm{d}s + \frac{p_0 \gamma}{\gamma + r} e^{-\lambda X(T) - rT} u\right). \tag{24}$$

The agent is willing to exert effort, and refrain from diverting funds, whenever (23) is at least as large as (24). If (23) is less than (24), the agent can gain from a one-shot deviation. The principal minimizes the agent's reward by setting the two equal.

We can evaluate the integral in (24) using integration by parts. We have

$$\int_t^T \lambda p_0 w(s)x(s)e^{-\lambda X(s)-rs}\mathrm{d}s = p_0\left(w(t)e^{-\lambda X(t)-rt} - w(T)e^{-\lambda X(T)-rT}\right) + \int_t^T p_0 e^{-\lambda X(s)-rs}\left(w'(s)-rw(s)\right)\mathrm{d}s.$$

Setting (23) equal to (24) and simplifying gives

$$\begin{aligned}
&\lambda p_0 w(T)e^{-\lambda X(T)-rT} - \frac{\lambda p_0\gamma}{\gamma+r}e^{-\lambda X(T)-rT}u \\
&= \int_t^T \lambda p_0 e^{-\lambda X(s)-rs}\left(w'(s)-rw(s)\right)\mathrm{d}s + cg(t)e^{-rt} \\
&= \int_t^T \lambda p_0 e^{-\lambda X(s)-rs}\left(w'(s)-rw(s)\right) + crg(s)e^{-rs} + \lambda p_0 cx(s)e^{-\lambda X(s)-rs}\mathrm{d}s \\
&\quad + cg(T)e^{-rT}.
\end{aligned}$$

This implies that the reward $w(t)$ satisfies the differential equation

$$w'(t) - rw(t) = -\frac{crg(t)}{\lambda p_0}e^{\lambda X(t)} - cx(t) = -\frac{cr(1-p_0)}{\lambda p_0}e^{\lambda X(t)} - c\left(\frac{r}{\lambda} + x(t)\right).$$

Using the integrating factor e^{-rt}, we have

$$\frac{\partial}{\partial t}\left(w(t)e^{-rt}\right) = -ce^{-rt}\left(\frac{r(1-p_0)}{\lambda p_0}e^{\lambda X(t)} + \frac{r}{\lambda} + x(t)\right),$$

implying

$$w(t) = e^{rt}\left(C + \int_t^T ce^{-rs}\left(\frac{r(1-p_0)}{\lambda p_0}e^{\lambda X(s)} + \frac{r}{\lambda} + x(s)\right)\mathrm{d}s\right),$$

for some constant C. Taking $t = T$, we have the boundary condition

$$\lambda p_0 w(T) - \frac{\lambda p_0 \gamma}{\gamma + r}u = c(1-p_0)e^{\lambda X(T)} + cp_0,$$

which gives us

$$w(T) = \frac{\gamma}{\gamma+r}u + \frac{c(1-p_0)e^{\lambda X(T)}}{\lambda p_0} + \frac{c}{\lambda}, \quad C = \frac{\gamma e^{-rT}}{\gamma+r}u + \frac{c(1-p_0)e^{\lambda X(T)-rT}}{\lambda p_0} + \frac{ce^{-rT}}{\lambda}.$$

Substituting and simplifying we have

$$\begin{aligned}
w(t) &= \frac{\gamma e^{-r(T-t)}}{\gamma+r}u + \frac{c(1-p_0)e^{\lambda X(T)-r(T-t)}}{\lambda p_0} + \frac{ce^{-r(T-t)}}{\lambda} \\
&\quad + \frac{c}{\lambda}\left(1-e^{-r(T-t)}\right) + ce^{rt}\int_t^T \frac{r(1-p_0)}{\lambda p_0}e^{\lambda X(s)-rs} + x(s)e^{-rs}\mathrm{d}s \\
&= \frac{\gamma e^{-r(T-t)}}{\gamma+r}u + \frac{c(1-p_0)e^{\lambda X(T)-r(T-t)}}{\lambda p_0} + \frac{c}{\lambda} \\
&\quad + \frac{c(1-p_0)}{\lambda p_0}\int_t^T re^{\lambda X(s)-r(s-t)}\mathrm{d}s + c\int_t^T x(s)e^{-r(s-t)}\mathrm{d}s \\
&= \frac{\gamma e^{-r(T-t)}}{\gamma+r}u + \frac{c(1-p_0)e^{\lambda X(t)} + cp_0}{\lambda p_0} + \int_t^T \frac{cg(s)}{p_0}x(s)e^{\lambda X(s)-r(s-t)}\mathrm{d}s,
\end{aligned}$$

where the last line follows using integration by parts.

I now check that it is in fact an optimal strategy for the agent to follow the prescribed experimentation policy $x(t)$. Suppose instead the agent adopts the policy $\hat{x}(t)$, and we write $\hat{X}$ and $\hat{g}$ for the corresponding functions of the agent's actual policy as opposed to the prescribed policy. Using Lemma **??**, we compute the agent's value as

$$\begin{aligned}
v(\zeta) &= \int_0^T \lambda p_0 w_{x,u}(t)\hat{x}(t)e^{-\lambda\hat{X}(t)-rt} + c\hat{g}(t)\left(x(t)-\hat{x}(t)\right)e^{-rt}\mathrm{d}t + \hat{g}(T)\frac{\gamma e^{-rT}}{\gamma+r}u \\
&= c\int_0^T \left(g(t)\hat{x}(t)e^{\lambda(X(t)-\hat{X}(t))} + \hat{g}(t)\left(x(t)-\hat{x}(t)\right)\right)e^{-rt}\mathrm{d}t \\
&\quad + c\int_0^T \lambda\hat{x}(t)e^{-\lambda\hat{X}(t)}\int_t^T g(s)x(s)e^{\lambda X(s)}e^{-rs}\mathrm{d}s\mathrm{d}t + \frac{\gamma e^{-rT}}{\gamma+r}u \\
&= c\int_0^T (1-p_0)\left(e^{\lambda(X(t)-\hat{X}(t))}-1\right)\hat{x}(t)e^{-rt} + \hat{g}(t)x(t)e^{-rt}\mathrm{d}t \\
&\quad + c\int_0^T \left(1-e^{-\lambda\hat{X}(t)}\right)g(t)x(t)e^{\lambda X(t)}e^{-rt}\mathrm{d}t + \frac{\gamma e^{-rT}}{\gamma+r}u \\
&= c\int_0^T e^{-rt}(1-p_0)\left(e^{\lambda(X(t)-\hat{X}(t))}-1\right)\left(\hat{x}(t)-x(t)\right) + g(t)x(t)e^{\lambda X(t)-rt}\mathrm{d}t \\
&\quad + \frac{\gamma e^{-rT}}{\gamma+r}u,
\end{aligned}$$

where the second equality is obtained by changing the order of integration. Since $\hat{x}(t) \leq x(t)$, the integral is bounded above by

$$c\int_0^T g(t)x(t)e^{\lambda X(t)-rt}\mathrm{d}t,$$

obtaining equality if and only if $\hat{x}(t) = x(t)$. □

Proof of Propositions 8 and 9

I first make a change of variables in our expression for the principal's value (14). Let $t(x)$ denote the smallest t such that

$$\int_0^t x(s)\mathrm{d}s = x.$$

Taking $X = X(T)$, and $g(x) = 1 - p_0 + p_0 e^{-\lambda x}$, the principal's maximum profit is

$$\pi\left(t(x), X, T, u\right) = \int_0^X e^{-\tilde{r}t(x)}\left(\lambda p_0 y e^{-\lambda x} - 2cg(x)\right)\mathrm{d}x + g(X)\frac{\gamma e^{-\tilde{r}T}}{\gamma + \tilde{r}}\pi^*(u)$$
$$- \int_0^X \lambda e^{-\lambda x - (\tilde{r}-r)t(x)}\left(\frac{p_0\gamma e^{-rT}}{\gamma + r}u + c\int_x^X g(x)e^{\lambda y - rt(y)}\mathrm{d}y\right)\mathrm{d}x.$$

For notational convenience, define $\delta = \frac{\gamma}{\gamma+r}$ and $\tilde{\delta} = \frac{\gamma}{\gamma+\tilde{r}}$. The optimization problem defining $\pi^*(v)$ is then

$$\max_{t(x),X,T,u} \quad \int_0^X e^{-\tilde{r}t(x)}\left(\lambda p_0 y e^{-\lambda x} - 2cg(x)\right)\mathrm{d}x + g(X)\tilde{\delta}e^{-\tilde{r}T}\pi^*(u)$$
$$- \int_0^X \lambda e^{-\lambda x-(\tilde{r}-r)t(x)}\left(p_0\delta u e^{-rT} + c\int_x^X g(x)e^{\lambda y - rt(y)}\mathrm{d}y\right)\mathrm{d}x$$
$$s.t. \quad \int_0^X cg(x)e^{\lambda x - rt(x)}\mathrm{d}x + \delta u e^{-rT} = v$$
$$\int_0^X t'(x)\mathrm{d}x \leq T$$
$$t'(x) \geq 1, \quad \forall x,$$

where $t'(x)$ is a distribution—in particular it can contain (positive) Dirac masses.

The Lagrangian for this problem is

$$\mathcal{L}\left(t(x), X, T, u\right) = \int_0^X e^{-\tilde{r}t(x)}\left(\lambda p_0 y e^{-\lambda x} - 2cg(x)\right)\mathrm{d}x + g(X)\delta e^{-\tilde{r}T}\pi^*(u)$$
$$- \int_0^X \lambda e^{-\lambda x-(\tilde{r}-r)t(x)}\left(p_0\delta u e^{-rT} + c\int_x^X g(y)e^{\lambda y - rt(y)}\mathrm{d}y\right)\mathrm{d}x$$
$$- \rho\left(\int_0^X cg(x)e^{\lambda x - rt(x)}\mathrm{d}x + \delta u e^{-rT} - v\right)$$
$$- \mu\left(\int_0^X t'(x)\mathrm{d}x - T\right) + \int_0^X \nu(x)(t'(x)-1)\mathrm{d}x.$$

There exist multipliers $\mu, \nu(x) \geq 0$ and $\rho \in \mathbb{R}$ such that $\mathcal{L}$ is maximized at an optimal solution $t(x), X, T, u$. The envelope theorem implies that $\rho = (\pi^*)'(v)$.

To prove Proposition 8, we compute the partial derivative of the Lagrangian with respect to X:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial X} &= e^{-\tilde{r}t(X)}\left(\lambda p_0 y e^{-\lambda X} - 2cg(X)\right) - \lambda p_0 e^{-\lambda X}\tilde{\delta} e^{-\tilde{r}T}\pi^*(u) \\ &\quad - \lambda p_0 u\delta e^{-\lambda X-(\tilde{r}-r)t(X)}e^{-rT} - \int_0^X \lambda e^{-\lambda x-(\tilde{r}-r)t(x)} cg(X) e^{\lambda X - rt(X)}\mathrm{d}x \\ &\quad - \rho cg(X)e^{\lambda X - rt(X)} - \mu t'(X)\end{aligned}$$

This implies that

$$\begin{aligned}(\pi^*)'(v) = \rho &\leq \frac{e^{-\lambda X+(r-\tilde{r})t(X)}\left(\lambda p_0 y e^{-\lambda X} - 2cg(X)\right)}{cg(X)} \\ &\leq e^{(r-\tilde{r})t(X)}\left(\frac{\lambda p_0 y e^{-\lambda X}}{cg(X)} - 2\right).\end{aligned}$$

For v close to zero, we know $(\pi^*)'(v)$ is unbounded, implying that the value of $t(X)$ must grow without bound as v gets close to zero. If there is no delay in the contract, we have $X = t(X)$, implying that

$$v \geq \int_0^{t(X)} cg(x)e^{(\lambda-r)x}\mathrm{d}x$$

This means v is bounded away from zero, a contradiction. We conclude that for v close to zero, the contract supporting $\pi^*(v)$ uses delay.

To prove Proposition 9, we take the first order condition for u:

$$0 = \frac{\partial \mathcal{L}}{\partial u} = g(X)\tilde{\delta}e^{-\tilde{r}T}(\pi^*)'(u) - \delta e^{-rT}\left(p_0\int_0^X \lambda e^{-\lambda x-(\tilde{r}-r)t(x)}\mathrm{d}x + \rho\right).$$

Since $(\pi^*)'(u)$ is unbounded near zero, the derivative of the Lagrangian with respect to u is necessarily positive for u sufficiently close to zero, so we never fire the agent. Rearranging the first order condition for u, and taking $\rho = (\pi^*)'(v)$, we have

$$(\pi^*)'\left(S(v)\right) = \frac{(\gamma+\tilde{r})e^{-(r-\tilde{r})T}}{(\gamma+r)g(X)}\left(p_0\int_0^X \lambda e^{-\lambda x-(\tilde{r}-r)t(x)}\mathrm{d}x + (\pi^*)'(v)\right).$$

If X is sufficiently close to zero, the leading term is strictly less than one. This implies that $(\pi^*)'\left(S(v)\right) < (\pi^*)'(v)$, and hence $S(v) > v$. If X_i converges to zero, then u_i does as well, but this is inconsistent with $S(v) > v$. Therefore, we must have a positive lower bound on X_i. This means that the sum of the X_i is unbounded, which establishes part (b).

Let $\underline{X} > 0$ denote a lower bound on the sequence $\{X_i\}$ in an optimal contract. We then have

$$u_{i-1} \geq e^{-r(T_i-\underline{X})}\int_0^{\underline{X}} cg(x)e^{(\lambda-r)x}\mathrm{d}x$$

for each i. If u_i converges to zero, then T_i must converge to infinity. However, for T sufficiently large, we can see again that $S(v) > v$, which is inconsistent with u_i converging to zero. We conclude that the sequence u_i is bounded away from zero, proving part (a). □