Discussion Paper No. 919

A TENURE-CLOCK PROBLEM

Chia-Hui Chen
Junichiro Ishida

January 2015

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

A Tenure-Clock Problem

Chia-Hui Chen* and Junichiro Ishida†

January 6, 2015

Abstract

We consider a "tenure-clock problem" in which a principal may set a deadline by which she needs to evaluate an agent's ability and decides whether to promote him or not. We embed this problem in a continuous-time model with both hidden action and hidden information, where the principal must induce the agent to exert effort to facilitate her learning process. The value of committing to a deadline is examined in this environment, and factors that make the deadline more profitable are identified. Our simple framework allows us to obtain a complete characterization of the equilibrium, both with and without commitment, and provides insight into why up-or-out contracts are prevalent in some industries while they are almost non-existent in others.

JEL Classification Number: D82, D83
Keywords: dynamic agency, deadlines, experimentation, standard promotion practices, up-or-out contracts.

*Institute of Economic Reserach, Kyoto University. Email: chchen@kier.kyoto-u.ac.jp
†Institute of Social and Economic Research, Osaka University. Email: jishida@iser.osaka-u.ac.jp

1 Introduction

Suppose that an employer needs to hire a worker to carry out a project over time. The project is ability-intensive in that the worker can successfully complete the project only if he is sufficiently productive. As is often the case, however, the worker's productivity is not directly observable to the employer, who must instead make an inference about his productivity from a sequence of observed outputs. A number of issues arise in this dynamic environment. How much time should the employer give the worker before she terminates the project? Should the employer commit to a deadline at the outset? If so, under what conditions?

In this paper, we attempt to address these issues by constructing a stylized model of a "tenure-clock problem," in which a principal hires an agent and evaluates his productivity over time in an ongoing relationship. The problem is embedded in a continuous-time framework with both hidden action and hidden information. At each instance, the agent privately chooses how much effort to supply. The outcome is either a success or a failure, depending on his effort choice as well as his productivity type. The game ends immediately when the agent achieves a success (or a "breakthrough"). The principal's task in this environment is to determine when to terminate the project, conditional on no success having occurred.

Two distinct cases are analyzed within this environment to evaluate the value of setting an evaluation deadline. In one case, the principal sets a deadline and commits to it at the outset, and the project is terminated automatically when the deadline is reached without attaining a success (but never before). In the other, the principal makes no such commitment, thereby retaining discretion to terminate the project at any instance with no specific deadline. The principal simply terminates the project when the principal judges that the agent is not up to the task. By directly comparing these two solutions, we evaluate the extent to which the principal benefits from committing to an evaluation deadline at the outset of the relationship in this dynamic environment.

The driving force of our analysis is a dynamic strategic interaction between the agent's effort choices and the principal's termination strategy. On one hand, the agent's effort choices depend clearly on how much time is left until the project is terminated: since the net value of achieving a success is low when the project is still far from termination, the agent tends to start off with low effort and gradually shift to higher effort as the expected termination date approaches. On the other hand, the principal's willingness to terminate the project depends also on the agent's effort choices: when the agent is less motivated and exerts low effort, less information is revealed about his type, which makes the principal more reluctant to terminate the project. This strategic interaction can generate a vicious cycle where the principal's reluctance to terminate the project diminishes the agent's motivation, which in turn makes the principal even more reluctant. It is in general profitable to commit to a deadline when each player's incentive to "procrastinate" is sufficiently strong.

We provide an analytical framework that can capture this dynamic interaction, arguably in its simplest form. Our simple framework allows us to obtain a complete characterization of the equilibrium, both with and without commitment. The case that is technically more challenging is the no-commitment one where the principal's termination strategy must be sequentially rational along the way, and the belief off the equilibrium path plays a crucial role. Even in this case, we show that a (pure-strategy) perfect Bayesian equilibrium always exists and is generically unique for any given set of parameters. We then build on this result to derive a necessary and sufficient condition under which the principal can strictly benefit from committing to an evaluation deadline.

Among other things, an interesting, somewhat counterintuitive, property of our model is that the average duration of the project can be either shorter or longer with commitment than without. In other words, there exists an equilibrium in which the principal prematurely terminates the project when she does not commit to a deadline. This result is somewhat surprising, provided that the inefficiency of the problem stems partly from the principal's reluctance to terminate the project. The principal tends to terminate the project too early when the agent's productivity under low effort is sufficiently small. To see this possibility most clearly, suppose that the agent's productivity under low effort is arbitrarily close to zero, in which case the principal can learn very little from failures (during the phase where the agent is supposed to exert low effort). The principal's belief thus declines very slowly over time, forcing her to incur a large amount of loss if she is to wait until the instantaneous payoff equals zero. We show that if the expected loss is prohibitively large, the principal chooses to terminate the project early even when the belief is still relatively high.

The most prominent example of evaluation schemes with deadlines is perhaps the "up-or-out system," which is widely observed in academia and professional service industries such as law, accounting, and consulting. From a practical point of view, our analysis offers some implications for when up-or-out contracts are more valuable by identifying several key factors that favor the use of an evaluation deadline: high ability intensity, stable job descriptions, similar jobs across ranks, and initial screening. Each of these factors intensifies either the agent's incentive to delay exerting high effort for a given deadline or the principal's incentive to delay terminating the project for a given effort sequence (or both). We later argue that these factors are staple features of industries that are traditionally characterized by the up-or-out system.

The paper is organized as follows. The literature review is provided in the remainder of this section. The model is presented in section 2 and analyzed in section 3, where we characterize both the commitment and no-commitment solutions. These solutions are then compared in section 4, in order to derive a necessary and sufficient condition under which it is optimal to commit to a deadline and conduct comparative statics exercises. Some extensions of the baseline model are discussed in section 5, and concluding remarks are offered in section 6.

Related Literature: The current analysis is most closely related to the experimentation literature in that the principal here attempts to uncover the agent's type though a sequence of experiments.[1] From the principal's point of view, our model can be seen as a variant of the canonical two-armed bandit problem with one safe arm (continuing the project) and one risky arm (terminating the project). A crucial difference from the existing models is that experiments in our context are "intermediated," i.e., experiments are conducted not by the principal herself but by an informed intermediary who may or may not behave as the principal wishes.[2]

Recently, there have been increasingly many works that explore the optimal provision of incentives in bandit problems. Manso (2011) considers the classic two-armed bandit problem and shows that the optimal contract in this context must tolerate, or even reward, early failures in order to encourage exploratory activities. Bergemann and Hege (1997, 2005) and Hörner and Samuelson (2013) analyze a financing problem of a venture capitalist where the principal provides funding to the agent who conducts experiments on a project of unknown quality. Gerardi and Maestri (2012) consider a similar environment where an agent conducts experiments but assume that the outcome of each experiment can only be observed by the agent. The principal must hence devise a contract not only to induce costly effort but also to truthfully reveal the information. Halac et al. (2013) analyze a model of long-term contracting for experimentation with hidden information about the agent's ability and dynamic moral hazard and obtain an explicit characterization of optimal contracts. Aside from some technical differences,[3] these previous works are primarily concerned with characterizing op-

[1]An early economic application of the bandit problem can be found in Rothschild (1974). See Bergemann and Välimäki (2008) for a survey.

[2]There are now increasingly many works, often called strategic experimentation, where a group of individuals, rather than a single individual, face bandit problems (Bolton and Harris, 1999; Keller et al., 2005; Klein and Rady, 2011; Bonatti and Hörner, 2011). There is some connection to this strand of literature in that the experimentation process in the current setup involves decisions of multiple individuals. In a different vein, Guo (2014) analyzes a model of "delegated experimentation" in which a principal delegates experimentation to an agent and solves for the optimal delegation rule.

[3]As a key technical difference, we consider a case where the agent knows his own productivity (or the project quality), so that our model belongs to the class of dynamic signaling, rather than of experimentation,

timal contracts. In contrast, the aim of this paper is to compare the allocations under full commitment and no commitment, while focusing on a narrower class of contracts, in order to evaluate the extent to which the principal can gain from committing to a deadline at the outset.[4]

It is well known that players often wait until the deadline to reach an agreement in finite-horizon models. This behavior, which is referred to as the deadline effect, is a topic of utmost concern in many bargaining and war of attrition models (Hendricks et al., 1988; Spier, 1992; Fershtman and Seidman, 1993; Hörner and Samuelson, 2011; Chen, 2012; Damiano et al., 2012; Fuchs and Skrzypacz, 2013). Some recent works also explore the role of deadlines in dynamic problems with multiple agents. Bonatti and Hörner (2011) analyze a dynamic moral-hazard problem with a team in which two agents work on a project of unknown quality and briefly discuss the optimal deadline in this context. Campbell et al. (2013) consider a similar environment where two agents work jointly on a project. They assume that there is only one project type but assume that one's own outcomes are his private information. Each agent can exert effort to produce a breakthrough individually, and a successful agent can reveal that he has been successful. The focus of these works is placed on the interaction between the agents, especially the freerider problem, whereas ours is on the dynamic interaction between the principal's termination decisions and the agent's effort choices.

The use of up-or-out is clearly intriguing because of its unique features, and has naturally spurred a lot of interest among economists. Two features are particularly crucial as defining characteristics of up-or-out: (i) there is an evaluation deadline by which a promotion must occur; (ii) a worker must leave the firm if he is not promoted by the deadline.[5] Most previous studies focus on the second feature (that unpromoted workers must leave their respective firms) and rationalize the use of up-or-out from that perspective.[6] In contrast, our analysis

from the agent's point of view.

[4]Mason and Välimäki (2014) consider a dynamic moral hazard problem and derives optimal wage contracts, both with and without commitment on wage payments. However, their work assumes only one agent type with an exclusive focus on moral hazard and thus does not consider hidden information about the agent's productivity.

[5]In addition, Kahn and Huberman (1988) also argue that an up-or-out contract must specify a retention wage, which is taken implicitly here.

[6]Some of the influential works on this topic focus on the second feature and emphasize its role in inducing human capital investments. Kahn and Huberman (1988) consider an environment where a worker invests in firm-specific human capital but the investment level is not verifiable. The unverifiability invites a problem, often referred to as double moral hazard, where the firm cannot credibly compensate the worker for making the investment. They show that the firm can effectively solve this problem by the use of an up-or-out contract, because the contract gives the firm an incentive to retain the worker if and only if his productivity is high enough. Similarly, building on the model of asymmetric learning, Waldman (1990) emphasizes the role of signaling in inducing general human capital investments. From a different perspective, O'Flaherty and Siow (1992) consider an environment where junior-level jobs are used to predict a worker's productivity at senior-level jobs and argue that up-or-out contracts prevail when junior-level job slots are scarce.

focuses exclusively on the first, i.e., the presence of an evaluation deadline, as the essential feature of up-or-out.[7] We focus on this feature primarily because of the recent trend observed in professional service industries: the way up-or-out is implemented in practice has become diverse even within a specific industry, and there are now increasingly many instances of up-or-out that do not necessarily force unpromoted workers to leave their respective firms.[8] This recent trend away from the traditional sense of up-or-out suggests that the second feature is not necessarily an essential aspect of up-or-out whereas the presence of a deadline is clearly so, almost by definition. Moreover, by focusing on the optimal deadline problem, the current framework can be applied to a wide range of circumstances in which a principal (an evaluator) must assess an agent's upside potential that is only gradually revealed in an ongoing relationship, e.g., a manager who must evaluate subordinates, a professor who must evaluate graduate students, a head coach in professional sports who must evaluate players, and so on.

As another departure from the literature, it is also worth noting that our analysis directly evaluates the value of up-or-out by comparing the commitment and no-commitment solutions. This aspect of our analysis stands in sharp contrast to the existing literature: as pointed out by Ghosh and Waldman (2010), most previous studies analyze up-or-out contracts independently, with no specific reference to "standard promotion practices," thereby leaving unresolved why up-or-out contracts are commonly instituted in some industries whereas they are virtually nonexistent in others.[9] Our analysis provides a new take on when up-or-out contracts can be more valuable vis-a-vis other more standard promotion practices, from a perspective previously unexplored in the literature.

2 A tenure-clock problem

Environment: We employ a continuous-time model because of its greater tractability; a discrete-time counterpart of the model is briefly discussed in Appendix B. Consider a situation where a principal (female) hires an agent (male) to complete a project. The game ends either when the agent attains a success (a "breakthrough") or when the principal terminates the project. The agent is either good with prior probability $p_0 \in (0, 1)$ or mediocre with prob-

[7]As such, we are not necessarily concerned about what happens after the deadline: the agent in our context may or may not be fired after the deadline; all we need is that the agent must incur some cost, in whatever form, for not being promoted before the deadline.

[8]For instance, many law firms now shift away from the traditional up-or-out model, known as the Cravath system, and increasingly adopt a two-tier partnership model where they retain those who fail to make partner as permanent associates or non-equity partners.

[9]As an exception, Ghosh and Waldman (2010) build on an asymmetric-learning model and compare the performance of up-or-out contracts with that of standard promotion practices.

ability $1-p_0$. The ability type is the agent's private information and is not directly observable to the principal who must instead evaluate it from a sequence of observed outcomes.

Production: The agent makes unobservable effort $a_t \in \{l, h\}$ at each instance t. We interpret that low effort ($a_t = l$) refers to the minimum level of effort that can be induced via input monitoring while high effort ($a_t = h$) refers to any part of effort that cannot be directly monitored by any means. The instantaneous cost of effort a is denoted by d_a where $d_h = d > 0$ and $d_l = 0$. A success arrives stochastically, depending on the effort choice as well as the agent's type. If the good type chooses $a_t = a$ over time $[t, t + dt)$, he attains a success with probability $\lambda_a dt$ where $\lambda_h > \lambda_l > 0$. In contrast, the mediocre type can never succeed with any effort level.[10] Define $\Delta_\lambda := \lambda_h - \lambda_l$.

Payoffs: A success yields a net present value of $y > 0$ to the principal; otherwise, she receives zero. If the agent attains a success, the principal pays a bonus $b \geq 0$ to him. Aside from this bonus, the principal also needs to pay a fixed wage $w \geq 0$ at each instance as long as she hires the agent. We call $(b, w) \in (0, \infty)^2$ the wage scheme for clarity and assume for most part that it is given exogenously: for a practical interpretation, w reflects the ongoing market wage of entry-level workers whereas b reflects that of senior-level workers.[11] The reservation payoff is assumed to be zero for both players. The common discount rate is denoted by $r \in (0, \infty)$.

Remark on exogenous wages: For most part of the analysis, we restrict our attention to the case where the principal can only set a deadline to motivate the agent while taking the wage scheme as given; the case with endogenous wages is briefly discussed in section 5.2. Although we make this assumption to clarify the role of deadlines, we also argue that this constitutes a nice starting point for several reasons. First, in many instances, equilibrium wages are dictated by market forces. Second, there are also many cases in reality where contingent monetary transfers are neither feasible due to contractual incompleteness nor desirable as in a professor-student relationship. Finally, the cost and benefit of achieving a success, or simply staying in a particular job, may involve non-transferrable gains such as prestige, authority, and the sense of achievement. Monetary transfers are of secondary importance when this non-pecuniary motive is sufficiently large. In any of these cases, the principal has little control over what we loosely refer to as the wage scheme, whereas the

[10] Our model specification is thus the "breakthrough" type in which one success can resolve all the uncertainty regarding the agent's type – an assumption that is predominant in the experimentation literature. See, for instance, Keller et al. (2005) and Bonatti and Hörner (2011, 2013). The analysis remains essentially the same even when the mediocre type can succeed with some probability, as long as the success probabilities are sufficiently small. See section 5.3 for details.

[11] In a broader sense, while we call b a bonus for expositional clarity, it is meant to capture all the benefits which accrue from a promotion, including non-transferrable gains such as prestige, authority, and job security; similarly, w is also meant to capture all the flow benefit from remaining in the current job.

option of setting a deadline is almost always viable.

Contract: The only contractible decision for the principal in this environment is whether to set a deadline, and if so, at what point; we thus do not consider the possibility of any screening contract (no direct mechanism).[12] An incentive scheme offered by the principal is thoroughly summarized by $T \in \mathbb{R}_+ \cup \{\varnothing\}$, while the wage scheme (b, w) is taken as given. If $T \in \mathbb{R}_+$, the principal terminates the project at time T (but never before) if the agent has not attained a success up to that point. If $T = \varnothing$, the principal chooses not to set any specific deadline, in which case she can terminate the project at any instance. In either case, both parties receive zero payoff once the project is terminated.

3 Analysis

3.1 Agent's effort decision

The agent decides whether or not to exert high effort at each instance. To analyze this problem, it is important to note that the principal's belief affects the agent's payoff only through her termination decision. The agent's optimal effort choice thus depends only on the remaining time to the termination date (hereafter, simply the remaining time), i.e., the maximum length of time for which the principal continues the project without attaining a success. The remaining time is obvious when the principal sets a deadline at T, in which case the remaining time at time t is simply given by $T - t$. Even without such an explicit commitment, however, the remaining time can be computed from the principal's equilibrium strategy in essentially the same manner as we shall discuss below. For now, we proceed with the presumption that the remaining time exists and is well-defined at any given point in time.

The agent's problem is rather straightforward since the mediocre type, knowing that his marginal value of effort is zero, never exerts high effort. As such, we can focus on the good type whom we refer to simply as the agent in what follows. Denote by $U(k)$ the agent's value function when the remaining time is k. Taking $k > 0$ as given, the value function can be written as

$$U(k) = \max_{a \in \{l,h\}} \Big((\lambda_a b - d_a + w)dt + e^{-rdt}(1 - \lambda_a dt)U(k - dt) \Big).$$

Taking limit $dt \to 0$, we obtain the Bellman equation:

$$rU(k) = \max_{a \in \{l,h\}} \Big(\lambda_a b - d_a + w - \lambda_a U(k) - \dot{U}(k) \Big), \tag{1}$$

[12] As we will discuss below, we only consider deterministic deadlines, in which case there is no feasible way to screen out the mediocre type. If we allow for stochastic deadlines, on the other hand, it may be possible for the principal to screen out the mediocre type by a menu of contracts. In this paper, however, we do not consider this possibility for the reasons mentioned later in footnote 13.

with $\lim_{k \to 0} U(0) = 0$. It is clear from this that the agent chooses high effort if and only if

$$\Delta_\lambda \big(b - U(k)\big) \geq d.$$

As usual in this type of setup (see, e.g., Bergemann and Hege, 2005), a success today implies the loss of future opportunities to earn rent, which inevitably diminishes the agent's incentive to work hard today. This rent-dissipation effect is naturally more accentuated when the agent has more time to prove himself. As such, the agent exerts high effort only when the project is sufficiently close to termination. This is a manifestation of the deadline effect that lies at the core of our entire analysis.

Proposition 1 *If*

$$\frac{\lambda_h b - d + w}{\lambda_h + r} > b - \frac{d}{\Delta_\lambda}, \tag{2}$$

there exists k^* such that the good type exerts high effort if the remaining time is less than or equal to k^* and low effort otherwise. The threshold k^* is given by

$$k^* = \begin{cases} -\frac{1}{\lambda_h + r} \ln\big(1 - \frac{(\lambda_h + r)(b - \frac{d}{\Delta_\lambda})}{\lambda_h b - d + w}\big) & \textit{if } b > \frac{d}{\Delta_\lambda}, \\ 0 & \textit{if } \frac{d}{\Delta_\lambda} \geq b. \end{cases} \tag{3}$$

If (2) does not hold, the agent always exerts high effort.

Proof: See Appendix A.

When $\Delta_\lambda b \leq d$, the static incentive is too weak for the agent to exert high effort for any remaining time. In contrast, when (2) fails to hold, the static incentive is strong enough to overcome the rent-dissipation effect, and the agent is willing to supply effort under any circumstance. As these cases only result in trivial solutions and are clearly of less interest for the question we pose here, we restrict our attention to the case where the strength of the static incentive lies in some intermediate range by making the following assumption.

Assumption 1 $\frac{\lambda_h b - d + w}{\lambda_h + r} > b - \frac{d}{\Delta_\lambda} > 0.$

Among other things, the assumption implies that the optimal threshold k^* is bounded from above and away from zero, i.e., $k^* \in (0, \infty)$.

3.2 Principal's termination decision

Let p_t denote the principal's belief that the agent is good at time t. Conditional on no success having occurred by time t, the updated belief evolves according to

$$\dot{p}_t = -\lambda_{a_t} p_t(1 - p_t), \tag{4}$$

taking the agent's effort choice as given. It is clear that the belief is strictly decreasing over time for any effort choice (due to the fact that $\lambda_l > 0$) until the agent attains a success, in which case the belief immediately jumps up to one.

In the current setup, because the principal's belief is strictly decreasing over time, there is a one-to-one relationship between the belief and the time for any given effort sequence. Since time is only an intermediate concept that by itself does not admit any economically meaningful interpretation, it is often more convenient to represent strategies in terms of the belief rather than time. The principal's problem can thus be defined as one wherein she chooses a cutoff belief (in the commitment case) or a set of cutoff beliefs (in the no-commitment case), at which she terminates the project. Accordingly, we represent the agent's strategy as $A : P \to \{l, h\}$ where $P \subset [0, 1]$ refers to the relevant domain of the belief (to be defined more precisely in later subsections).

Let $\tau(p; A)$ denote the time at which the belief reaches p under a given strategy A, with $\tau(p_0; A) = 0$ and $p_{\tau(p;A)} = p$ for any p. The principal's continuation payoff at time t for some termination belief q is obtained as

$$\int_{\tau(p_t;A)}^{\tau(q;A)} v(p_s; A) e^{\int_{\tau(p_t;A)}^{s} -(\lambda_{a_u} p_u + r) du} ds,$$

subject to (4), where $v(p; A)$ is the instantaneous payoff given by

$$v(p; A) = \lambda_{A(p)} p y - w.$$

It is immediate to see that for a given effort level, there exists a threshold belief below which the instantaneous payoff is negative for the principal. We denote by $q_a := \frac{w}{\lambda_a y}$ the "break-even" belief at which the principal's instantaneous payoff equals zero under effort a (whenever it is well defined). Let $q_a = 1$ if $\frac{w}{\lambda_a y} > 1$. Combined with the fact that the belief is strictly decreasing over time for any effort sequence, q_h then represents the absolute lowerbound of the belief, as the principal clearly has no incentive to continue the project once her belief dips below this level. Since the belief must reach this level sooner or later (due to the fact that $\lambda_l > 0$ and $w > 0$), the presence of such a lowerbound suggests that the game must end in some finite time. This allows us to solve the game by backward induction, analogously to the "gap case" of the durable-good monopoly problem.

Finally, if the value of a success is too small, the model only admits a trivial solution where the principal chooses to stop immediately (or not to hire the agent in the first place). In what follows, therefore, we assume that the value of a success is large enough for the principal to hire the agent at least for some positive duration.

Assumption 2 $p_0 > q_h \;\Leftrightarrow\; \lambda_h p_0 y > w$.

3.3 The commitment solution

The principal's problem under commitment is to choose a termination belief q which maximizes her expected payoff at time 0. Throughout the analysis, we restrict attention to deterministic deadlines where the principal stops when the belief reaches q with probability one.[13] As we have already seen, the high-ability agent exerts high effort only in $[\max\{T-k^*, 0\}, T]$ for any given deadline T. Let A^*_q denote the agent's best response, given the principal's strategy q. To represent the agent's best response, it is convenient to define a backward operator:

$$\phi(p) := \frac{p}{p+(1-p)e^{-\lambda_h k^*}}.$$

That is, $\phi(p)$ indicates the belief wherein starting with $p_s = \phi(p)$, if the agent exerts high effort for $t \in [s, s+k^*]$, we end up with $p_{s+k^*} = p$. Given this definition, the agent's best response is given by

$$A^*_q(p) = \begin{cases} l & \text{for } p_0 \geq p > \phi(q), \\ h & \text{for } \phi(q) \geq p \geq q, \end{cases}$$

if $p_0 > \phi(q)$; otherwise, $A^*_q(p) = h$ for all $p \in [q, p_0]$. In what follows, we abbreviate the subscript q and denote the best response simply by A^* to ease notation.

Let $\tau^*(p) := \tau(p; A^*)$, $\sigma^*(q) := \max\{\tau^*(q) - k^*, 0\}$ and $V^*(p; q) := V(p; q, A^*)$. The principal's continuation payoff at time 0 can then be written as

$$\begin{aligned} V^*(p_0; q) = &\int_0^{\sigma^*(q)} (\lambda_l p_s y - w) e^{-\int_0^s (\lambda_l p_u + r) du} ds \\ &+ e^{-\int_0^{\sigma^*(q)} (\lambda_l p_u + r) du} \int_{\sigma^*(q)}^{\tau^*(q)} (\lambda_h p_s y - w) e^{-\int_{\sigma^*(q)}^s (\lambda_h p_u + r) du} ds. \end{aligned}$$

[13] Our exclusive focus on deterministic deadlines is motivated by the fact that we almost never observe a contract in which the deadline is stochastic in nature. One possible reason for this is that a contract with stochastic deadlines inherently requires a credible randomization device and is thus hard to implement in practice. For instance, consider a contract where the principal terminates the project at t^1 or t^2, $t^1 < t^2$, with equal probability. At time t^1, however, the principal's preference is generically strict, i.e., it is strictly better either to stop or to continue, so that there is no incentive to randomize at that point. If it is strictly better to stop at time t^1, for instance, the principal stops at time t^1 with probability one but knowing that, the contract effectively has a deterministic deadline.

With some computation, we obtain

$$\begin{aligned}V^*(p_0;q) = &\frac{p_0\pi_l}{\lambda_l+r}(1-e^{-(\lambda_l+r)\sigma^*(q)}) + \frac{p_0 e^{-(\lambda_l+r)\sigma^*(q)}\pi_h}{\lambda_h+r}(1-e^{-(\lambda_h+r)(\tau^*(q)-\sigma^*(q))})\\ &-\frac{(1-p_0)w}{r}(1-e^{-r\tau^*(q)}),\end{aligned}$$

where $\pi_a := \lambda_a y - w$ is the expected instantaneous payoff from the good type under effort a. The first two terms represent the expected gain from the high-ability type whereas the last term represents the expected loss from the mediocre type.

The commitment solution, denoted by q^C, is defined as

$$q^C := \text{argmax}_{q\in[0,p_0]}\ V^*(p_0;q).$$

The continuation payoff is continuous but kinked at $q = \psi_0 := \phi^{-1}(p_0)$. For $p_0 > q \geq \psi_0$, the first-order condition is given by

$$p_0\pi_h e^{-\lambda_h\tau^*(q)} - (1-p_0)w = 0. \tag{5}$$

For $\psi_0 > q$, we have

$$p_0 e^{-(\lambda_l+r)\sigma^*(q)}\Big(\pi_l - \frac{\lambda_l+r}{\lambda_h+r}\pi_h(1-e^{-(\lambda_h+r)k^*})\Big) - (1-p_0)we^{-r\tau^*(q)} = 0,$$

which can also be written as

$$\big(p_0\pi_h e^{-(\lambda_l\sigma^*(q)+\lambda_h k^*)} - (1-p_0)w\big)e^{-r\tau^*(q)} = \frac{p_0\Delta_\lambda e^{-(\lambda_l+r)\sigma^*(q)}}{\lambda_h+r}\big(w+ry+\pi_h e^{-(\lambda_h+r)k^*}\big). \tag{6}$$

Note that the left-hand side of (6) is the expected instantaneous payoff for the principal at the time of termination, whereas the right-hand side is the marginal cost of extending the deadline. Since the right-hand side is strictly positive, the principal must set the deadline at a point where the expected instantaneous payoff is positive if $\psi_0 > q$.

Proposition 2 *For any given set of parameters $\Theta := (b, y, w, d, \lambda_l, \lambda_h, r, p_0)$ satisfying Assumptions 1 and 2, there exists a unique commitment solution $q^C \in [q_h, p_0)$. If*

$$\frac{\lambda_l+r}{\lambda_h+r}\pi_h(1-e^{-(\lambda_h+r)k^*}) \geq \pi_l, \tag{7}$$

the commitment solution is given by

$$q^C = \begin{cases} q_h & \textit{for } \phi(q_h) \geq p_0 > q_h,\\ \frac{p_0 e^{-\lambda_h k^*}}{1-p_0+p_0 e^{-\lambda_h k^*}} & \textit{for } 1 > p_0 > \phi(q_h).\end{cases}$$

If (7) fails to hold, the commitment solution is given by

$$q^C = \begin{cases} q_h & \text{for } \phi(q_h) \geq p_0 > q_h, \\ \frac{p_0 e^{-\lambda_h k^*}}{1-p_0+p_0 e^{-\lambda_h k^*}} & \text{for } \Pi \geq p_0 > \phi(q_h), \\ \hat{q} & \text{for } 1 > p_0 > \Pi, \end{cases}$$

where $\Pi := \frac{we^{-rk^*}}{\pi_l - \frac{\lambda_l+r}{\lambda_h+r}\pi_h(1-e^{-(\lambda_h+r)k^*})+we^{-rk^*}} \in (\phi(q_h), 1)$*, and* $\hat{q} \in (q_h, 1)$ *is the unique solution to (6).*

Proof: See Appendix A.

3.4 The no-commitment solution

The situation becomes more complicated, and perhaps more intriguing, when the principal makes no commitment at the outset. Formally, for this no-commitment case, we solve for a perfect Bayesian equilibrium in Markov strategies (hereafter, simply an equilibrium) with the belief as the state variable. The restriction to Markov strategies is without loss of generality in this context because the only observable history to the principal is the sequence of outcomes that consists only of failures at any continuation game (conditional on no breakthrough having occurred). The principal's termination decision can thus depend only on her belief. Given this and the fact that effort is unobservable, the agent's strategy must also take a Markovian form that depends only on the principal's belief.

We first note that the instantaneous payoff must be negative for any effort choice once the belief dips below q_h. It is therefore a dominant strategy for the principal to stop for any $p \leq q_h$. This implies that the strategies beyond that point are totally irrelevant, and we specify the strategies only over $[q_h, p_0]$.[14] The agent's strategy is thus specified over this range as $A : [q_h, p_0] \to \{l, h\}$. The principal's termination strategy is now specified in a different way, as her strategy off the equilibrium path may matter. More precisely, the principal's strategy is given by a set of termination beliefs $Q \subset [q_h, p_0]$ that may not be a singleton. The agent's best response is now denoted by A^*_Q (or simply A^* when it does not cause any confusion). With slight abuse of notation, let $V^*(p; Q)$ denote the continuation payoff under the best response A^*.

We denote the equilibrium strategy by Q^*. One can then easily show that although Q^* may or may not be a singleton, it is always finite.

Lemma 1 Q^* *is always finite.*

[14]To be more precise, the principal always terminates the project immediately for any $p \leq q_h$. Given this, the best response for the agent is to exert high effort for any $p \leq q_h$, but this has no consequence, even off the equilibrium path, because it is a dominant strategy to stop when $p \leq q_h$.

Proof:

Since Q is defined on a compact set $[q_h, p_0]$, it suffices to show that there is no continuous interval $L \subset [q_h, p_0]$ such that the principal stops for $p \in L$. To show this, suppose on the contrary that there is such an interval. Since the remaining time is zero during this whole interval, we must have $A^*(p) = 1$ for $p \in L$. However, this implies that the principal's expected payoff must be positive since $p > q_h$. As this gives her an incentive to deviate, there cannot be such an interval in equilibrium. ∎

Given this result, we denote each element of Q by q^i where $q^1 < q^2 < \cdots < q^n$. Accordingly, the game is divided into distinct segments P^i, $i = 1, 2, ..., n$ by termination beliefs, where $P^i = [q^i, q^{i+1})$ for $i = 1, 2, ..., n-1$ and $P^n = [q^n, p_0]$. Only the last segment P^n is actually played on the equilibrium path, as the game ends with probability one by the time the belief reaches q^n. Still, when the principal makes no commitment to a deadline, the agent's strategy off the equilibrium path may play a role in constructing an equilibrium. For expositional clarity, we let $q^{NC} = q^n$ and call it the no-commitment solution.

When the principal makes no commitment, her termination (or continuation) decision must be sequentially rational at each instance, given the current belief. To be more precise, for some given strategy A, the principal's continuation payoff at the time of termination must be zero, whereas it must be strictly positive elsewhere. Given that the agent adopts the best response A^*_Q, these restrictions amount to the following two equilibrium conditions which the principal's termination strategy Q must satisfy.

Condition T: (i) $q^1 = q_h$ and (ii) if $n > 1$, for each $i = 2, ..., n$, $\lim_{p \to q^i - 0} V^*(p; Q) = 0$.

Condition C: $V^*(p; Q) > 0$ for $p \in [q_h, p_0] \setminus Q$.

Condition T is the usual indifference condition that requires the principal to terminate the project when the continuation payoff (when she deviates) is zero.[15] This indifference condition alone is in general not sufficient because the instantaneous payoff v may not be monotonically decreasing over time: in any segment P^i, the agent may start off with low effort, during which the instantaneous payoff could be so small that the principal is tempted to stop prematurely. Condition C assures that the principal does not stop before the intended termination belief q^i is reached.

The next result establishes that there exists a perfect Bayesian equilibrium that is always unique even under no commitment.

[15] The right-hand limit of the continuation payoff must, by definition, be zero at the time of termination. The principal has an incentive to terminate the project only if the left-hand limit is also zero. This condition effectively implies that $V^*(p; Q^*)$ is continuous for all $p \in [q_h, p_0]$.

Proposition 3 *For any given set of parameters Θ satisfying Assumptions 1 and 2, there exists a unique equilibrium in which the principal stops at $q^{NC} \in [q_h, \min\{p_0, q_l\})$. The principal's equilibrium strategy is characterized by a set of $n \geq 1$ distinct termination beliefs where (i) $q^1 = q_h$ and (ii) if $n > 1$, $q^{i+1} > \phi(q^i) > q^i$ for all $i = 1, 2, ..., n-1$.*

Proof: See Appendix A.

When λ_l is sufficiently high, the principal's instantaneous payoff is positive even under low effort. In this case, the principal is not tempted to stop prematurely, and only Condition T is sufficient to pin down the equilibrium. Figure 1 depicts this situation: the agent starts with low effort and switches to high effort when the remaining time is k^*; the principal stops when the belief reaches the lowerbound q_h.

[Figure 1 about here]

In contrast, the instantaneous payoff may become negative under low effort when λ_l is relatively low, in which case we may have a situation where the principal's continuation payoff also becomes negative before the belief reaches q_h. Figures 2 and 3 show the evolution of the belief and the expected payoff while fixing the principal's strategy at $Q = \{q_h\}$. In the figures, q_l is so high that the instantaneous payoff is negative for the entire interval during which the agent exerts low effort, and there exists a point q^2 such that the continuation payoff is negative for $p \in (q^2, p_0]$. This implies that $Q = \{q_h\}$ does not constitute an equilibrium as it violates Condition C. In this case, the game is divided into two segments, P^1 and P^2, as illustrated in Figure 4. The pair of strategies now satisfies the equilibrium conditions since the instantaneous payoff is always positive in P^2.

[Figures 2-4 about here]

3.5 The cooperative solution: a benchmark

Before we move on to discuss when it is optimal to commit to a deadline, we briefly examine as a benchmark the case where both the principal and the agent attempt to maximize the sum of their individual payoffs to derive the first-best allocation. More precisely, the instantaneous payoff in the cooperative case, common to both players, is the difference between the expected benefit from a success minus the effort cost, i.e., $\lambda_a p(y + b) - d_a$ for a given effort a.

There is one important caveat for the analysis of the cooperative case. If the agent shared the same objective as we assume here, his private information could in principle be induced at no cost: given that the agent's reservation payoff is zero, the mediocre type is indifferent between participating and dropping out immediately (before time 0), as he would obtain zero payoff in either case for any given deadline. As it turns out, however, the optimal deadline is independent of the initial belief p_0 and is hence identical irrespective of whether or not the principal can immediately screen out the mediocre type. To make the situation comparable to the noncooperative case, we assume that for some reasons, no information can be exchanged between the principal and the agent.[16]

The agent's problem is essentially the same and only needs a slight modification. Letting $\tilde{U}(k)$ denote the agent's continuation payoff, the agent exerts high effort if and only if

$$\Delta_\lambda\big((y+b)-\tilde{U}(k)\big) \geq d.$$

As in the noncooperative case, there exists a threshold $\tilde{k}^*$ such that the good type exerts high effort if and only if the remaining time is less than or equal to $\tilde{k}^*$. Under the maintained assumptions, one can show that the agent is generally more motivated in the cooperative case than in the noncooperative case. Formally, we make the following statement.

Proposition 4 *Under Assumption 1, if*

$$\frac{\lambda_h(y+b)-d}{\lambda_h+r} > y+b-\frac{d}{\Delta_\lambda}, \tag{8}$$

the threshold is given by

$$\tilde{k}^* = -\frac{1}{\lambda_h+r}\ln\big(1-\frac{(\lambda_h+r)(y+b-\frac{d}{\Delta_\lambda})}{\lambda_h(y+b)-d}\big). \tag{9}$$

If (8) does not hold, the agent always exerts high effort. Moreover, $\tilde{k}^* > k^*$.

Proof: See Appendix A.

The principal's problem is, on the other hand, fundamentally different because in this setup, there is no social cost of hiring the mediocre type who simply yields zero surplus at every instance.[17] To make the argument parallel to the noncooperative case, we denote the principal's continuation payoff in the cooperative case by $\tilde{V}^*(p;q)$, which is simply a

[16] Since the mediocre type is indifferent, there is also an equilibrium where the mediocre type chooses to participate. Alternatively, one can think of this as a case where the mediocre type is selfish and cares about his own payoff, as in the noncooperative case. In either event, however, the optimal deadline is exactly the same; see the proof of Proposition 5 for details.

[17] Of course, the current setup is a special case in that the net value of the mediocre type is precisely zero. On the other hand, if there is a fixed cost of production, the net value of the mediocre type is strictly negative. In this case, however, the mediocre type has a strict incentive to drop out immediately. The qualitative nature of the solution with a fixed cost remains the same: the principal never terminates the project until the agent succeeds.

restatement of $\tilde{U}(k)$ in terms of beliefs. Further, define $\tilde{\tau}^*(p)$ and $\tilde{\sigma}^*(p) := \max\{\tilde{\tau}^*(p) - \tilde{k}^*, 0\}$ analogously. We then obtain

$$\tilde{V}^*(p_0; q) = \int_0^{\tilde{\sigma}^*(q)} \lambda_l p_s (y+b) e^{-\int_0^s (\lambda_l p_u + r) du} ds$$
$$+ e^{-\int_0^{\tilde{\sigma}^*(q)} (\lambda_l p_u + r) du} \int_{\tilde{\sigma}^*(q)}^{\tilde{\tau}^*(q)} p_s \big(\lambda_h (y+b) - d\big) e^{-\int_{\tilde{\sigma}^*(q)}^s (\lambda_h p_u + r) du} ds,$$

which can be written as

$$\tilde{V}^*(p_0; q) = \frac{p_0 \lambda_l (y+b)}{\lambda_l + r} (1 - e^{-(\lambda_l + r)\tilde{\sigma}^*(q)})$$
$$+ \frac{p_0 e^{-(\lambda_l + r)\tilde{\sigma}^*(q)} (\lambda_h (y+b) - d)}{\lambda_h + r} (1 - e^{-(\lambda_h + r)(\tilde{\tau}^*(q) - \tilde{\sigma}^*(q))}).$$

The cooperative solution, denoted by $\tilde{q}$, is defined as

$$\tilde{q} := \mathrm{argmax}_{q \in [0, p_0]} \tilde{V}^*(p_0; q).$$

Since there is no social cost of hiring the mediocre type, the social surplus from hiring the agent is strictly positive for any belief level, and as such, it could be optimal for the principal to hire the agent indefinitely. This is indeed what happens in the cooperative case: the principal never terminates the project even though the agent would never exert high effort for any finite $\tilde{k}^*$ if the game were to continue indefinitely. The reason for this is simple. The agent exerts low effort only when the continuation payoff, which coincides with the social surplus by definition, is larger than the static gain. On the other hand, if it is ever optimal to set a deadline to induce high effort, the deadline must be set at time $\tilde{k}^*$, at which point the continuation payoff must be lower than the static gain. As such, when the preferences are aligned (and there is no fixed cost of production), the first-best allocation is to continue the project indefinitely until the agent succeeds, although the agent may only exert low effort all the way.

Proposition 5 *In the cooperative case, the principal never terminates the project, and the game continues indefinitely. The agent always exerts low effort if (8) is satisfied and high effort otherwise.*

Proof: See Appendix A.

4 Up-or-out versus standard promotion practices

4.1 The value of up-or-out

The most prominent example of evaluation schemes with deadlines is arguably what is called the "up-or-out system" – a promotion scheme that is widely observed in academia and professional service industries such as law, accounting, and consulting. Under a typical up-or-out contract, the employer sets a deadline by which a promotion must occur, and the worker must leave the firm if he is not promoted by the deadline. Our simple framework offers some practical implications for up-or-out by precisely evaluating the value of committing to such an evaluation deadline at the outset.

To clarify what we attempt to investigate at this juncture, it is important to note that the principal can gain no relevant information along the equilibrium path because the only feasible history at any continuation game is the one consisting only of failures up to that point. This immediately implies that the no-commitment solution can yield no higher payoff than the commitment solution. Still, under some conditions, the no-commitment solution can replicate the commitment solution, in which case the no-commitment solution becomes optimal if the principal faces a (possibly very small) commitment cost.[18] The main purpose of this section is to clarify this condition, i.e., clarify when the absence of commitment entails no additional cost, under the presumption that the commitment solution is generally more costly to enforce than the no-commitment solution.

Given that the no-commitment solution can at best replicate the allocation of the commitment solution, one can easily see that setting a deadline strictly benefits the principal if and only if $q^C \neq q^{NC}$. On one hand, it is clear that the no-commitment solution gives the principal a different, and necessarily lower, payoff when $q^C \neq q^{NC}$. On the other hand, when $q^C = q^{NC}$, the equilibrium allocation must be exactly the same because Proposition 1 ensures the unique equilibrium given the same horizon. Using this fact, we can obtain an exact necessary and sufficient condition under which setting a rigid deadline yields a strictly positive value for the principal.

Proposition 6 *The commitment solution generically yields a higher expected payoff for the principal than the no-commitment solution if and only if* $p_0 > \phi(q_h)$,[19] *i.e.,*

$$p_0 > \frac{q_h}{q_h + (1-q_h)e^{-\lambda_h k^*}} \;\Leftrightarrow\; \frac{1-q_h}{q_h}\Big(1 - \frac{(\lambda_h + r)(b - \frac{d}{\Delta_\lambda})}{\lambda_h b - d + w}\Big)^{\frac{\lambda_h}{\lambda_h + r}} > \frac{1-p_0}{p_0}. \qquad (10)$$

Proof: The sufficiency is obvious from the argument we have seen thus far: if $p_0 > \phi(q_h)$, the optimal condition for the commitment solution is given by (6), which clearly differs from that for the no-commitment solution. To show the necessity, note that if $\phi(q_h) \geq p_0$, the

[18] An obvious cost is that the principal must give up flexibility to adjust *ex post* to any uncertainty that may resolve during the course of play although this aspect is assumed away in the current setup for clarity. The commitment cost may also arise from the cost of writing and enforcing a formal contract.

[19] To be more precise, the expected payoff of the no-commitment solution is always lower if $p_0 > \phi(q_h) \geq q_l$: in this case, $q^C > q^{NC} = q_h$ always holds. If $p_0, q_l > \phi(q_h)$, we may have $q^{NC} > q_h$, and there could be some non-generic cases where $q^C = q^{NC}$ and the expected payoffs are identical.

agent exerts high effort from the beginning and the principal stops when the belief reaches q_h, which results in the same payoff for the principal, either with or without commitment. ∎

The proposition states that the principal benefits from committing to a deadline if the initial prior belief p_0 exceeds a certain threshold $\phi(q_h)$. The value of up-or-out depends crucially on each player's propensity to "procrastinate." In short, the threshold is lower and hence favors the commitment solution for a given initial belief either when the agent delays exerting high effort for a given deadline or when the principal delays terminating the project for a given effort sequence. Below, we briefly summarize how each player's propensity to procrastinate is determined:

1. The agent's propensity to procrastinate is determined by the tradeoff between the current gain of attaining a success and the potential loss of future payoffs. For a given deadline, the agent tends to procrastinate more (a small k^*) when the potential loss of future payoffs is large relative to the current gain.

2. The principal's propensity to procrastinate depends on the break-even belief and the likelihood of a success under high effort. The break-even belief is evidently the major force determining how patient the principal can be, where the principal is more tempted to wait for an eventual success when q_h is low. Further, for a given q_h, the principal tends to wait longer when λ_h is small, because the information about the agent's type is revealed only slowly in that case.

Before we move on, it is important to note that our characterization result owes largely to the fact that we consider a binary effort choice or, more importantly, that there is an upperbound for effort. In contrast, if the effort domain is unbounded with a convex effort cost, the commitment solution in general yields a strictly higher payoff than the no-commitment solution.[20] Even in this case, however, the nature of the problem remains largely intact. The intuition provided above carries over to this more general case as well because the value of up-or-out is still determined in essentially the same way, although the analysis would be somewhat more complicated.

[20] On the other hand, we can obtain essentially the same result even if the effort level is continuous as long as the feasible effort level is bounded from above.

4.2 When is up-or-out optimal?

Up-or-out contracts are quite common in academia and professional service industries (hereafter, we refer to them inclusively as professional service industries).[21] In academia, it is often the case that incoming assistant professors must be promoted within a certain period of time (typically around six years). Outside of academia, a leading example of industries with up-or-out is the legal service industry. Many law firms traditionally adopt a set of managerial practices, known as the Cravath system, which include up-or-out (or "partnership track") where incoming associates must make partner within a certain period of time (roughly seven to ten years). This contrasts sharply with other typical firm organizations which almost never specify deadlines by which workers must be promoted.

To understand why this is the case, note that these professional service industries are indeed unique in several regards. For instance, high knowledge intensity, combined with low capital intensity, is an obvious hallmark of these industries (von Nordenflycht, 2010).[22] Another distinctive feature is a professionalized workforce that builds on a particular knowledge base (Torres, 1991; von Nordenflycht, 2010). This in turn creates well-defined job boundaries, from which we derive three additional implications: stable job descriptions, job homogeneity across different ranks, and initial screening. In what follows, we interpret our main result in light of these characteristics to provide further insight into why up-or-out is more prevalent in professional service industries.

Ability intensity: It is often argued that knowledge intensity is perhaps the most fundamental characteristic of professional service industries. Combined with the fact that they are also less capital intensive,[23] the productivity of an organization depends crucially on the extent of knowledge embodied in individuals. A likely consequence of this fact is that innate ability matters and creates value that cannot be easily substituted by sheer effort, at least in the short run – a feature which we call ability intensity for expositional clarity.

Since the mediocre type can never attain a success in the current setup, one way to measure the extent of ability intensity is by the (good type's) success probabilities. We say that the project is more ability-intensive when λ_l and λ_h are larger with a fixed Δ_λ. Note that an increase in λ_h, while fixing Δ_λ, yields three conflicting effects. First, it raises the agent's productivity at the margin, which in turn lowers the threshold q_h. Second, it gives

[21] The military is another example which is characterized by up-or-out. We do not consider this case because the underlying structure appears to be quite different.

[22] The discussion in von Nordenflycht (2010) is mainly about professional service industries, but these characteristics are clearly applicable to academia as well.

[23] High knowledge intensity and low capital intensity are not equivalent in the strict sense of the word because, as argued by von Nordenflycht (2010), an industry can be both knowledge- and capital-intensive at the same time. For the purpose of this study, however, we do not make any distinction between them.

the agent a stronger incentive to procrastinate (a lower k^*) as it raises his expected future payoff under high effort. Finally, it also facilitates the principal's learning as each failure reveals more information. The first two effects raises the value of up-or-out while the last lowers it, such that the overall impact is not immediately clear.

One can readily show, however, that there exists some threshold $\bar{\lambda}$ such that it is optimal to set a deadline, thereby favoring the use of up-or-out, for $\lambda_h > \bar{\lambda}$. To see this, we rewrite (10) as

$$\frac{p_0}{1-p_0} > \frac{q_h}{(1-q_h)e^{-\lambda_h k^*}} = \frac{w}{(\lambda_h y - w)e^{-\lambda_h k^*}}. \tag{11}$$

Under the maintained assumptions, $\lambda_h > \underline{\lambda} := \max\{\frac{w}{p_0 y}, \frac{\Delta_\lambda}{d}\big(r(b-\frac{d}{\Delta_\lambda})+d-w\big)\}$. If $\frac{\Delta_\lambda}{d}\big(r(b-\frac{d}{\Delta_\lambda})+d-w > \frac{w}{p_0 y}$, we have $\lim_{\lambda_h \to \underline{\lambda}} k^* = \infty$, and hence, (11) is never satisfied if λ_h is small. In contrast, we have

$$\lim_{\lambda_h \to \infty} \frac{q_h}{(1-q_h)e^{-\lambda_h k^*}} = 0,$$

implying that (11) must hold if λ_h is sufficiently large.

Stable job descriptions: Due to various professional requirements and accreditation processes, workers in professional service industries are typically responsible only for a narrow and clear set of tasks, compared to workers in other industries. This implies that the nature of tasks that they are expected to carry out is very stable over time, and the production environment is relatively immune to stochastic shocks. For instance, a demand or technology shock that entirely changes the job description of a lawyer or a college professor is highly unlikely.

We can interpret r as the rate at which the project is terminated for exogenous reasons, e.g., the arrival of a stochastic shock that makes the project completely worthless.[24] Define $\bar{r}$ such that

$$\frac{\lambda_h b - d + w}{\lambda_h + \bar{r}} = b - \frac{d}{\Delta_\lambda}.$$

Assumption 1 then implies that r must be bounded between $(0, \bar{r})$. It is clearly not profitable to commit to a deadline if r is sufficiently close to $\bar{r}$ because $\lim_{r \to \bar{r}} k^* = \infty$. In contrast, there exists some $\underline{r}$ such that it is profitable to commit to a deadline for $r < \underline{r}$ if

$$\frac{1-q_h}{q_h}\Big(1 - \frac{\lambda_1(b-\frac{d}{\Delta_\lambda})}{\lambda_h b - d + w}\Big) > \frac{1-p_0}{p_0}.$$

[24]Here, we allow the principal to terminate the project when the shock arrives; see section 5.1 for the case where the principal must abide by the deadline under any circumstance.

It can be seen from this argument that a decrease in r (a more stable relationship) in general raises the value of up-or-out because the agent can be more forward-looking and tends to procrastinate more, which in turn makes the principal unable to stop at the right time.

Job similarity across ranks: Another immediate consequence of stable job descriptions is that jobs necessarily become similar across different ranks.[25] In the current setup, this aspect may be captured by y, which can be interpreted as the expected payoff of promoting a high-ability agent. If entry- and senior-level jobs are similar, y tends to be larger because a success at the entry level is a reliable predictor of productivity at the senior level. This is not necessarily the case in typical firm organizations where jobs across ranks can differ to a considerable extent, as often discussed in the context of the Peter Principle.[26]

The effect of a change in y is fairly straightforward, as it only affects the break-even belief with no impact on the agent's behavior. A large value of y implies a small value of q_h which gives the principal a stronger incentive to wait for a breakthrough. Given this, the agent also has a stronger incentive to procrastinate, hoping to achieve a breakthrough with low effort. The principal can then unambiguously benefit from committing to a deadline when y is sufficiently large, implying that the value of up-or-out is higher in industries where jobs are similar and there is a strong correlation between performances at different ranks.

Initial screening: When the principal holds a very optimistic belief to begin with, the incentive to wait for a success is stronger, making it harder to stop at the right time. We argue that this is generally the case in professional service industries, for reasons that pertain to their professional nature and well-defined job boundaries. First, job candidates are all trained specialists who have already gone through some selection processes. Second, it is relatively easy to assess an individual's upside potential even at the entry level because there are readily available signals. Moreover, combined with the fact that these industries are typically vertically differentiated, the job market is largely characterized by assortative matching where top organizations attract top prospects: although different organizations have different expectations for incoming workers, each hires a candidate only when there is a reasonable chance for the candidate to clear the hurdle. All of these amount to the fact that the quality of workers is relatively high and homogenous in professional service industries,[27]

[25] Ghosh and Waldman (2010) also raises job similarity as one of the distinctive characteristics of academia and law. Job similarity is also crucial in Kahn and Huberman's (1988) classic argument because if jobs are sufficiently different, we may use promotions to solve the double moral-hazard problem as suggested by Prendergast (1993).

[26] The Peter Principle states that every post tends to be occupied by an employee who is incompetent to carry out its duties, because employees are promoted through positions where they have excelled until they reach a level of incompetence (Peter and Hull, 1969). A premise of this argument is that jobs are inherently different and become progressively harder as one climbs though ranks.

[27] Abowd et al. (2012) show that most workers are of high ability and that technology favors positive

which in turn raises the value of up-or-out.

4.3 Too early or too late?

In the absence of commitment, the principal often fails to stop at the right time because she has a strong incentive to wait for a success. This reasoning seems to indicate that the principal always terminates the project earlier with commitment than without, or equivalently, $p^C \geq p^{NC}$ for any given parameters Θ. As indicated in Figures 3 and 4, however, this is not necessarily true in the current setup. When the principal knows that she cannot stop soon enough, and it is too costly for her, she may terminate the project even when the instantaneous payoff is still strictly positive or, alternatively, when the belief is strictly higher than the break-even point q_h. This is generally the case when λ_l is low, in which case the principal must incur losses in the low-effort phase.

Proposition 7 *$p^{NC} > p^C$ for almost all $p_0 \in (\phi(q_h), 1)$ as λ_l tends to zero.*

Proof: It follows from Proposition 2 that if $p_0 > \phi(q_h)$,

$$\frac{p_0 e^{-\lambda_h k^*}}{1 - p_0 + p_0 e^{-\lambda_h k^*}} \geq p^C.$$

To prove the first part of the proposition, it thus suffices to show that

$$p^{NC} > \frac{p_0 e^{-\lambda_h k^*}}{1 - p_0 + p_0 e^{-\lambda_h k^*}}. \tag{12}$$

We know from the proof of Proposition 3 that (12) holds if $\phi(q^{NC}) > p_0 \geq q^{NC}$: in this case, the agent starts with high effort and the principal stops when the belief reaches q^{NC} in equilibrium, which takes less than k^* units of time by definition. Note that as $\lambda_l \to 0$, we have $q^{NC} \to \phi(q^{NC})$, and thus, this holds for almost all $p_0 \in (\phi(q_h), 1)$. ∎

Two additional implications follow from this result. The first is concerned with the role of λ_l. As in any moral-hazard environment, a low value of λ_l directly implies a high marginal value of effort, which tends to better motivate the agent given the same wage. Aside from this conventional effect, our analysis also suggests an alternative logic by which the principal can benefit from a low λ_l. Suppose that (10) holds, so that the principal has a strong incentive to wait for a success. In fact, if λ_l is sufficiently large, the principal cannot help but wait until the belief reaches q_h, as there is no incentive to stop along the way. A lower value of λ_l can help in this context, because that gives the principal an incentive to stop before the belief

assortative matching in professional service industries.

reaches q_h, thereby prompting the agent to shift effort earlier. In other words, a low value of λ_l may be beneficial for the principal because it works as a strong commitment device.

Second, the result also suggests that under no commitment, the average duration of the project may not be monotonically increasing in the initial belief. When the initial belief is relatively high, it takes long for the belief to reach the lowerbound q_h, and the principal may have to incur large losses if she is to wait until the moment at which the instantaneous payoff equals zero. To avoid this situation and discipline the agent, the principal must act early, i.e., she may need to pull the trigger earlier for *ex ante* more promising agents in some cases.[28]

5 Extensions

5.1 The case with uncertainty

An important cost of making commitment arises from the fact that the principal loses flexibility to adjust to future stochastic shocks. The baseline model does not capture this cost because the principal can gain no relevant information on the equilibrium path, and there is thus no uncertainty regarding the termination date. As a consequence, it is weakly optimal to commit to a deadline in the baseline model. This is apparently unsatisfactory, provided that we live in a world filled with uncertainty where there is clearly value in being flexible. While we focus on the simplest case to illuminate the key insight, it is of crucial interest to examine whether and how our results can be extended to cases where the principal faces some uncertainty that can be resolved only gradually over time.

There are obviously many different ways to introduce uncertainty into the baseline model. Among a plethora of possibilities, we here consider a case where a permanent productivity shock may strike with some probability, which totally changes the nature of the task and subsequently makes the agent unproductive, i.e., $\lambda_h = \lambda_l = 0$.[29] For simplicity, we assume that (i) the shock is permanent and arrives at most once, with a Poisson arrival rate β, and (ii) the arrival of the shock is publicly observable. Other cases, e.g., those with transitory shocks, can in principle be analyzed in an analogous manner although it may be computationally more cumbersome.

The agent: The agent's problem requires a slight modification. The augmented Bellman

[28]Formally, when the principal's termination equilibrium strategy is characterized by multiple termination beliefs, the average duration increases with the initial belief within each segment P^i but jumps down to zero at each p^i. As is clear from this argument, the principal's expected payoff, though continuous, is not monotonic either when the principal's strategy is characterized by multiple termination beliefs. The expected payoff is single-peaked within each segment P^i and equals zero at each p^i.

[29]Once the principal chooses to commit to a deadline, the principal must abide by it and is not allowed to terminate the project, unlike in the discussion in section 4.2.

equation is obtained as

$$rU^j(k) = \max_{a \in \{l,h\}} \Big((b\lambda_a - d_a + w) - (\lambda_a + \beta)U(k) + \beta \tilde{U}^j(k) - \dot{U}(k) \Big), \; j = C, NC$$

where $\tilde{U}^j$ denotes the value function after the shock has arrived. Since the principal's reaction to the shock differs, $\tilde{U}^j$, and hence $U^j(k)$, depend on whether or not the principal makes commitment. As in the baseline model, the incentive compatibility constraint (before the shock strikes) can be written as

$$\Delta_\lambda \big(b - U^j(k) \big) \geq d,$$

although the value function $U^j(k)$ now differs. As such, the agent exerts high effort only when the project is sufficiently close to termination. We let k^C and k^{NC} denote the threshold under commitment and no commitment, respectively.

Proposition 8 *For any Θ and β, $k^{NC} > k^C$. Both k^{NC} and k^C are strictly increasing in β if $\frac{b\lambda_h - d + w}{\lambda_h + \beta + r} > b - \frac{d}{\Delta_\lambda} > 0$.*

Proof: See Appendix A.

It is interesting to note that the possibility of a negative shock generally lowers the expected future payoff and hence induces the agent to exert high effort for longer, either with or without commitment. First, with commitment, the best the agent can do after the shock strikes is to exert low effort and earn the flow payoff w, which is the lowest possible payoff as long as the project survives. The situation is even worse without commitment, however, because the principal can immediately terminate the project, thereby leaving no rent for the agent. The extent of the loss is therefore larger when the principal makes no commitment. The agent is generally better motivated in the face of uncertainty, which in turn favors the no-commitment solution.

The principal: We start with the case where the principal commits to a deadline at the outset. Let $V^C(p_0; q)$ denote the principal's expected payoff at time 0, which can be written as

$$
\begin{aligned}
V_0^C(p_0;q) = & \int_0^{\sigma^C(q)} (\lambda_l p_s y - w) e^{-\int_0^s (\lambda_l p_u + \beta + r) du} ds \\
& + e^{-\int_0^{\sigma^C(q)} (\lambda_l p_u + \beta + r) du} \int_{\sigma^C(q)}^{\tau^C} (\lambda_h p_s y - w) e^{-\int_{\sigma^C}^s (\lambda_h p_u + \beta + r) du} ds \\
& - \frac{\beta}{r} \int_0^{\sigma^C(q)} w(1 - e^{-r(\tau - s)}) e^{-\int_0^s (\lambda_l p_u + \beta + r) du} ds \\
& - \frac{\beta}{r} e^{-\int_0^{\bar{t}} (\lambda_l p_u + \beta + r) du} \int_{\sigma^C(q)}^{\tau^C(q)} w(1 - e^{-r(\tau - s)}) e^{-\int_{\sigma^C(q)}^s (\lambda_h p_u + \beta + r) du} ds. \qquad (13)
\end{aligned}
$$

where $\tau^C(q)$ is the time at which the belief reaches q under the agent's best response, and $\sigma^C := \max\{\tau^C(q) - k^C, 0\}$. The commitment solution is obtained as

$$
\tau^C = \text{argmax}_\tau V_0^C(\tau).
$$

As can be seen from (13), the marginal benefit of extending the deadline decreases with β, which prompts the principal to terminate earlier than in the case with no uncertainty.

In contrast, the problem is roughly the same as in the baseline model when the principal makes no commitment because the principal only pays attention to the instantaneous payoff. As in the baseline model, the principal stops when the instantaneous payoff becomes sufficiently low. To be more precise, define

$$
\phi^{NC}(p) := \frac{p}{p + (1-p)e^{-\lambda_h k^{NC}}}.
$$

Given that no shock has arrived, if $\phi^{NC}(q_h) > q_l$, the principal stops when the belief reaches q_h. The only difference is that when the shock strikes, the instantaneous payoff falls to $-w$, which prompts her to stop immediately. The principal's expected payoff is given by

$$
\begin{aligned}
V^{NC}(p;q) = & \int_0^{\sigma^{NC}(q)} (\lambda_l p_s y - w) e^{-\int_0^s (\lambda_l p_u + \beta + r) du} ds \\
& + e^{-\int_0^{\sigma^{NC}(q)} (\lambda_l p_u + \beta + r) du} \int_{\sigma^{NC}(q)}^{\tau^{NC}} (\lambda_h p_s y - w) e^{-\int_{\sigma^{NC}(q)}^s (\lambda_h p_u + \beta + r) du} ds, \qquad (14)
\end{aligned}
$$

where τ^{NC} and σ^{NC} are defined analogously. Note that for the same time horizon, $V^{NC}(p_0;q) > V_0^C(p_0;q)$ for any q.

The optimal incentive scheme: In the baseline model with no uncertainty, any allocation which can be realized without commitment can be realized with commitment as well. This is no longer the case with uncertainty because the principal cannot foresee if and when the shock strikes, in which case she needs to immediately terminate the project.

Proposition 9 *The principal strictly benefits from not committing to a deadline for any $\beta > 0$ if $\phi^{NC}(q_h) \geq p_0$.*

Proof: If $\phi^{NC}(q_h) \geq p_0$, the agent exerts high effort from the beginning under no commitment. The principal's payoff is maximized, and the principal strictly benefits from making no commitment in this case. To see this, observe that $k^{NC} > k^C$ implies $\phi^{NC}(q_h) > \phi^C(q_h)$. If $\phi^{NC}(q_h) \geq p_0 > \phi^C(q_h)$, the agent starts off with low effort, and the principal is forced to set a deadline before the belief reaches q_h. The profit is clearly lower than $V_0^{NC}(\tau^{NC})$. On the other hand, if $\phi^{NC}(q_h) > \phi^C(q_h) \geq p_0$, the agent exerts high effort from the beginning even under commitment. The profit is still lower, however, because the principal must incur a flow loss when the shock strikes whereas she can immediately terminate the project if she makes no commitment. ∎

The presence of uncertainty lowers the value of up-or-out through two channels. First, by making no commitment, the principal can retain the flexibility to adjust to the negative shock by immediately terminating the project whenever it strikes. Aside from this conventional effect of flexibility, there also arises another force that favors the no-commitment solution: in the face of uncertainty, the agent is better motivated by the constant threat of project termination, and thus, he starts exerting high effort earlier. Since $\phi^{NC}(q_h) > \phi(q_h)$ for any $\beta > 0$ by Proposition 8, we could have $\phi^{NC}(q_h) \geq p_0 > \phi(q_h)$, in which case the value of up-or-out is strictly negative with uncertainty whereas it is strictly positive without.

5.2 Endogenous wages

We have thus far assumed that the principal has no control over wage payments and takes the wage scheme (b, w) as given. While we argue that this covers a range of situations of our interest, it is also important to understand how the wage scheme affects the equilibrium allocation when the principal has some control over it. What is especially intriguing in this respect is the role of the flow wage w, as it can directly influence the principal's termination decision and hence can be used as a substitute for the deadline when the commitment cost is prohibitively large.

Here, we briefly analyze the case where the principal can choose to pay any $w \geq 0$ as she likes. To analyze this problem, note that because the flow wage here is a pure transfer payment with no incentive effect on the agent's side, there is no reason for the principal to offer any wage beyond the minimum level, which is normalized at zero, if she can stop at the right time. However, as we have seen, this is not always true when the principal makes no commitment to a deadline. As such, the main question we address here is whether there are circumstances in which she offers any strictly positive w.[30] Since only the case with no commitment is relevant, we focus on this case by assuming that the commitment cost is so large that it is never optimal to commit to a deadline.

In general, an increase in w above zero yields two opposing effects. On one hand, it raises the critical value q_h that may give the principal the incentive to stop earlier than otherwise. On the other hand, it also raises the expected future payoff that in turn lowers k^* and gives the agent the incentive to procrastinate even more. A necessary condition for $w > 0$ is that the first effect dominates the second, or alternatively that $\tau^*(q^{NC}) - k^*$ decreases with w.[31]

Proposition 10 $w^* > 0$ *if* λ_l *is sufficiently small.*

Proof: See Appendix A.

When $w = 0$, the instantaneous payoff is always strictly positive (as in the cooperative case), no matter how unlikely the agent is to succeed. This is the most extreme case of the principal's procrastination, as she can never terminate the project, and given that, the agent never exerts high effort. The principal can alter this structure by raising w above zero because the flow cost of employment is now positive, rendering the instantaneous payoff negative at some point. This can be profit-enhancing for the principal because with the credible threat of termination, she can induce the agent to shift effort earlier, which is especially beneficial when the success probability under low effort is low.

5.3 When the mediocre type too can succeed

We have thus far assumed that only the good type can succeed with positive probability, which amounts to a breakthrough-type specification where one success can resolve all the uncertainty. As is clear from the argument, the assumption that the mediocre type can never succeed is not crucial, as long as the model retains the breakthrough-type structure. Here, we briefly make a remark on this assumption and show that the basic results are independent of the success probabilities of the mediocre type as long as they are sufficiently small.[32]

The success probabilities are now denoted by λ_a^j, where $j = G, M$ indicates the agent's type (G for good and M for mediocre). Let $\Delta_\lambda^j := \lambda_h^j - \lambda_l^j$. We now allow for the possibility

[30] In other words, we assume that the minimum wage level is zero. On the other hand, if the principal can set a negative w, she can immediately screen out the mediocre type by offering $w < 0$. This means that we focus on the situation where the minimum wage is so high that even the mediocre type can earn a positive rent by being employed.

[31] Note that $\tau^*(q^{NC}) - k^*$ is the interval during which the agent exerts low effort. If it is longer with a smaller k^*, the principal's expected payoff is necessarily lower.

[32] In contrast, if the success probabilities are relatively large, the principal may not stop after one success, and the analysis would be much more complicated.

that $\lambda_h^M > 0$ and $\lambda_l^M > 0$, so that the mediocre type may achieve a success with positive probability. We restrict attention to the case where λ_h^M and λ_l^M are both sufficiently small, such that $\lambda_h^G > \lambda_l^G > \lambda_h^M > \lambda_l^M \geq 0$. When the mediocre type too can succeed, the principal's belief never hits the maximum, and hence some uncertainty always remains. We thus need to formally address the optimal stopping problem for the principal, which we sidestepped in the baseline model: after the agent achieves a success, the principal can either continue the project in search of an additional success or to stop the process and "promote" the agent.[33] Suppose that the principal earns payoffs of y_G and $-y_M$ if she promotes the good type and the mediocre type, respectively. The expected net payoff of promoting the agent at any instance is now a function of the current belief p_t and is given by $B(p_t) = y_G p_t - y_M(1-p_t)$. Let $s(p)$ denote the updated belief when a success occurs at belief p, which is given by

$$s(p) = \frac{\lambda^G_{a_t^G} p}{\lambda^G_{a_t^G} p + \lambda^M_{a_t^M}(1-p)},$$

where a_t^j denotes the effort choice of type j at time t.

In what follows, we solve for equilibria where the principal always promotes the agent after one success; in other words, one success is still informative enough. Note first that the principal's belief declines monotonically over time for any effort profile if $\lambda_l^G > \lambda_h^M$, which holds by assumption. The agent's best response thus remains the same, where there exists some threshold k^* such that the agent exerts high effort if and only if the remaining time is less than k^*. With slight abuse of notation, let $V^*(p, q)$ denote the principal's value function when the agent adopts the best response, and the principal always promotes the agent after one success. For some given termination belief q, the principal promotes the agent after one success if

$$B(s(p)) \geq V^*(s(p), q). \tag{15}$$

There exists an equilibrium where the principal always promotes the agent after one success if this condition holds for all $p \in [q, p_0]$ (on top of all the other equilibrium conditions). Clearly, this is the case when λ_h^M is sufficiently small even if it is strictly bounded away from zero.

6 Conclusion

This paper presents a simple model of a tenure-clock problem with particular emphasis on the role of commitment to a deadline. We consider an environment where the principal chooses

[33] In principle, the principal can promote the agent at any instance. We only consider the moment at which a success occurs because it is never optimal to do so at any other instance (except at time 0).

whether to commit to a deadline and if so, at what time. Within this framework, we establish the generic uniqueness of the equilibrium and obtain a necessary and sufficient condition for committing to a deadline to be optimal. The simple framework also allows us to conduct various comparative statics exercises to give insight into why up-or-out contracts are more prevalent in some industries than others.

As a final note, since our model is highly stylized, there are several avenues to extend the current analysis. Of particular interest is the case with symmetric information where the agent himself does not know his own type and needs to uncover it through experiments. In this symmetric-information case, there is an additional effect that is absent in our asymmetric-information setup: having observed no success, the agent gradually loses confidence in his own ability and is eventually discouraged to work hard as time progresses. This seems to be an important aspect of many professional industries and is worth a comprehensive investigation. We leave this for the future.

References

Abowd, J.M., Kramarz, F., Perez-Duarte, S. and Schmutte, I.M., 2012, A Formal Test of Assortative Matching in the Labor Market, mimeo.

Bergemann, D. and Hege, U., 1998, Venture Capital Financing, Learning, and Moral Hazard, *Journal of Banking and Finance*, 22, 703-35.

Bergemann, D. and Hege, U., 2005, The Financing of Innovation: Learning and Stopping, *RAND Journal of Economics*, 36, 719-52.

Bergemann, D. and Välimäki, J., 2008, Bandit Problems, in Durlauf, S. and Blume, L., eds, *The New Palgrave Dictionary of Economics*, 2nd edition, Macmillan Press.

Bolton, P. and Harris, C., 1999, Strategic Experimentation, *Econometrica*, 67, 349-74.

Bonatti, A. and Hörner, J., 2011, Collaborating, *American Economic Review*, 101, 632-63.

Bonatti, A. and Hörner, J., 2013, Career Concerns and Market Structure, mimeo.

Campbell, A., Ederer, F. and Spinnewijn, J., 2014, Delay and Deadlines: Freeriding and Information Revelation in Partnerships, *American Economic Journal: Microeconomics*, 6, 163-204.

Chen, C.-H., 2012, Name Your Own Price at Priceline.com: Strategic Bidding and Lockout Periods, *Review of Economic Studies*, 79, 1341-69.

Damiano, E., Li, H. and Suen, W., 2012, Optimal Deadlines for Agreements, *Theoretical Economics*, 7, 357-93.

Fershtman, C. and Seidmann, D.J., 1993, Deadline Effects and Inefficient Delay in Bargaining with Endogenous Commitment, *Journal of Economic Theory*, 60, 306-21.

Fuchs, W. and Skrzypacz, A., 2013, Bargaining with Deadlines and Private Information, *American Economic Journal: Microeconomics*, 5, 219-43.

Gerardi, D. and Maestri, L., 2012, A Principal-Agent Model of Sequential Testing, *Theoretical Economics*, 7, 425-63.

Ghosh, S. and Waldman, M., 2010, Standard Promotion Practices Versus Up-or-Out Contracts, *RAND Journal of Economics*, 41, 301-25.

Guo, Y., 2014, Dynamic Delegation of Experimentation, mimeo.

Halac, M., Kartik, N., and Liu, Qingmin, 2013, Optimal Contracts for Experimentation, mimeo.

Hendricks, K., Weiss, A. and Wilson, C., 1988, The War of Attrition in Continuous Time with Complete Information, *International Economic Review*, 29, 663-80.

Hörner, J. and Samuelson, L., 2011, Managing Strategic Buyers, *Journal of Political Economy*, 119, 379-425.

Hörner, J. and Samuelson, L., 2013, Incentives for Experimenting Agents, *RAND Journal of Economics*, 44, 632-63.

Kahn, C. and Huberman, G., 1988, Two-Sided Uncertainty and "Up-or-Out" Contracts, *Journal of Labor Economics*, 6, 423-44.

Keller G., Rady, S. and Cripps, M., 2005, Strategic Experimentation with Exponential Bandits, *Econometrica*, 73, 39-68.

Klein, N.A. and Rady, S., 2011, Negatively Correlated Bandits, *Review of Economic Studies*, 78, 693-732.

Manso, G., 2011, Motivating Innovation, *Journal of Finance*, 66, 1823-69.

Mason, R. and Välimäki, J., 2014, Getting It Done: Dynamic Incentives to Complete a Project, forthcoming in *Journal of the European Economic Association.*

O'Flaherty, B. and Siow, A., 1992, On the Job Screening, Up or Out Rules, and Firm Growth, *Canadian Journal of Economics*, 25, 346-68.

Peter, L.J. and Hull, R., 1969, *The Peter Principle: Why Things Always Go Wrong*, New York: Morrow.

Prendergast, C., 1993, The Role of Promotion in Inducing Specific Human Capital Acquisition, *Quarterly Journal of Economics*, 108, 523-34.

Rothschild, M., 1997, A Two-Armed Bandit Theory of Market Pricing, *Journal of Economic Theory*, 9, 185-202.

Spier, K.E., 1992, The Dynamics of Pretrial Negotiation, *Review of Economic Studies*, 59, 93-108.

Torres, D., 1991, What, If Anthining, is Professionalism? Institutions and the Problem of Change, *Research in the Sociology of Organizations: Organizations and Occupations*, 8, 43-68.

von Bordenflycht, A., 2010, What is a Professional Serive Firm? Towards a Theory and Taxonomy of Knowledge Intensive Firms, *Academy of Management Review*, 35, 155-74.

Waldman, M., 1990, Up-or-Out Contracts: A Signaling Perspective, *Journal of Labor Economics*, 8, 230-50.

Appendix A: Proofs

Proof of Proposition 1: It is intuitively clear that the value function is strictly increasing in k. Formally, define

$$U_a(k) = \Big((\lambda_a b - d + w) - \lambda_a U_a(k) - \dot{U}_a(k)\Big) \tag{16}$$

as the value function when the effort level is fixed at $a = l, h$. Solving this differential equation and imposing the boundary condition $U(0) = 0$, we derive

$$U_h(k) = \frac{\lambda_h b - d + w}{\lambda_h + r}(1 - e^{-(\lambda_h + r)k}).$$

If $\Delta_\lambda b > d$, it is optimal for the agent to exert high effort as long as $U_h(k)$ is sufficiently small. Then, k^* is obtained as a solution to $U_h(k^*) = b - \frac{d}{\Delta_\lambda}$, which can be written as

$$\frac{\lambda_h b - d + w}{\lambda_h + r}(1 - e^{-(\lambda_h + r)k^*}) = b - \frac{d}{\Delta_\lambda}. \tag{17}$$

Note that the agent always exerts high effort if

$$b - \frac{d}{\Delta_\lambda} \geq \frac{\lambda_h b - d + w}{\lambda_h + r},$$

while he never does so if $\frac{d}{\Delta_\lambda} \geq b$. For in-between cases, solving (17) yields

$$k^* = -\frac{1}{\lambda_h + r}\ln\Big(1 - \frac{(\lambda_h + r)(b - \frac{d}{\Delta_\lambda})}{\lambda_h b - d + w}\Big),$$

so that the agent exerts high effort when the remaining time is less than or equal to k^*. For $k > k^*$, the agent exerts low effort because $U(k) \geq U_h(k) > b - \frac{d}{\Delta_\lambda}$. ∎

Proof of Proposition 2: First, if $\phi(q_h) \geq p_0 > q_h$ (the latter inequality holds by Assumption 2), there exists a unique deadline $T \in (0, k^*]$ that satisfies

$$\frac{p_0 e^{-\lambda_h T}}{1 - p_0 + p_0 e^{-\lambda_h T}} = q_h.$$

There is clearly no reason to wait beyond this point because the instantaneous payoff is strictly negative for any effort choice. It is also not optimal to stop before q_h because the instantaneous payoff is still strictly positive. The optimal termination belief is then q_h and it is unique in this range.

Second, suppose that $p_0 > \phi(q_h)$. In this case, it is optimal to set a deadline at k^* if $\lim_{q \to \psi_0 - 0} \frac{\partial V^*}{\partial q} \leq 0$. The condition for this can be written as

$$p_0\Big(\pi_l - \frac{\lambda_l + r}{\lambda_h + r}\pi_h(1 - e^{-(\lambda_h + r)k^*})\Big) \leq (1 - p_0) w e^{-rk^*},$$

which can also be written as

$$w e^{-rk^*} \geq p_0\Big(\pi_l - \frac{\lambda_l + r}{\lambda_h + r}\pi_h(1 - e^{-(\lambda_h + r)k^*}) + w e^{-rk^*}\Big).$$

This condition holds for any p_0 if

$$\frac{\lambda_l + r}{\lambda_h + r}\pi_h(1 - e^{-(\lambda_h + r)k^*}) \geq \pi_l. \tag{18}$$

If this holds, it is optimal to set the deadline at $T = k^*$. The optimal termination belief is given by

$$q^C = \frac{p_0 e^{-\lambda_h k^*}}{1 - p_0 + p_0 e^{-\lambda_h k^*}},$$

which is again unique in this range.

Finally, if (18) fails to hold, there exists some $\Pi \in (\phi(q_h), 1)$ such that

$$w e^{-rk^*} = \Pi\Big(\pi_l - \frac{\lambda_l + r}{\lambda_h + r}\pi_h(1 - e^{-(\lambda_h + r)k^*}) + w e^{-rk^*}\Big).$$

For $p_0 > \Pi$, we can find a unique $\hat{q}$ that solves

$$p_0 e^{-(\lambda_l + r)\sigma^*(\hat{q})}\Big(\pi_l - \frac{\lambda_l + r}{\lambda_h + r}\pi_h(1 - e^{-(\lambda_h + r)k^*})\Big) = (1 - p_0) w e^{-r\tau^*(\hat{q})}.$$

This is the unique termination belief when $p_0 > \Pi$. ∎

Proof of Proposition 3: We first establish the following two facts that will be useful for the subsequent argument.

Lemma 2 *In any equilibrium, $q_h \in Q^*$ and $A^*(p) = h$ for $\phi(q_h) \geq p \geq q_h$.*

Proof: Suppose otherwise, i.e., that the agent chooses low effort for some interval $[t + dt)$ when $p_t \leq \phi(q_h)$. To satisfy the incentive condition, the agent must exert high effort at least for $[t + dt, t + k^* + dt)$, but then the belief necessarily falls below q_h at time $t + k^* + dt$, at which point the principal's instantaneous payoff is negative for any effort choice. This is therefore a contradiction. ∎

The lemma indicates that the agent always exerts high effort once the belief dips below $\phi(q_h)$. Note that this result holds both on and off the equilibrium path. Of course, this does not rule out the possibility that the agent exerts high effort for $p > \phi(q_1)$; as we will see below, there exists an equilibrium in which the agent exerts high effort before the belief reaches $\phi(q_h)$.

Lemma 3 *For each $q^i \in Q^*$, $q^i \in \{q_h\} \cup (\phi(q_h), q_l]$.*

Proof: We have already argued that it is a dominant strategy to terminate the project for any $p_t < q_h$ because the instantaneous payoff is negative for any effort choice. Likewise, it is a dominated strategy to stop at any $p_t > q_l$ because the instantaneous payoff is positive for any effort choice, which shows that $q^i \in [q_h, q_l]$. Note also that the instantaneous payoff under high effort is strictly positive for all $p \in (q_h, \phi(q_h))$, such that Condition T is not satisfied. ∎

Given these results, we show the existence and uniqueness of the equilibrium by construction. There are three cases we need to consider, depending on the initial belief p_0.

Case 1: $\phi(q_h) \geq p_0$

Lemmas 2 and 3 directly imply $Q^* = \{q_h\}$ in equilibrium. In the unique equilibrium, the agent always exerts high effort on the equilibrium path and the principal stops at $p_t = q_h$. The no-commitment solution coincides with the commitment solution and is characterized by $q^{NC} = q_h$.

Case 2: $p_0 > \phi(q_h) \geq q_l$

This is the case where the equilibrium may involve a phase during which the agent exerts low effort on the equilibrium path. When $\phi(q_h) \geq q_l$, however, we can show the existence and uniqueness of the equilibrium in essentially the same manner as in Case 1.

An obvious candidate strategy in this case is that the agent exerts low effort until the belief reaches $\phi(q_h)$ and then switches to high effort from that point on. To check whether this strategy constitutes an equilibrium, note that once the belief reaches $\phi(q_h)$, the unique continuation equilibrium is that the agent exerts high effort and the principal stops when the belief reaches q_h, as shown by Lemmas 2 and 3. Moving backwards, it is optimal for the principal to exert low effort before the belief reaches $\phi(q_h)$. It follows from Lemma 3 that the principal never stops before the belief reaches $\phi(q_h)$ because $\phi(q_h) > q_l$. The no-commitment solution is then given by $q^{NC} = q_h$ as above. It is also clear from this argument that the constructed equilibrium must be unique: on one hand, there is only one equilibrium once the belief reaches $\phi(q_h)$; on the other hand, it is never optimal to stop before the belief reaches $\phi(q_h)$ because $\phi(q_h) > q_l$.

Case 3: $p_0 > \phi(q_h)$ and $q_l > \phi(q_h)$

This is the case where if the agent starts out with low effort, the belief reaches q_l before switching to high effort. This can entirely change the structure of the equilibrium as the principal may be tempted to stop when the belief reaches q_l; in other words, Condition C may now bind.

We show the existence and uniqueness of the equilibrium by checking Condition C sequentially until it is satisfied. We start with $Q^1 = \{q_h\}$ and the corresponding best response $A^*_{Q^1}$ such that

$$A^*_{Q^1}(p) = \begin{cases} l & \text{for } p_0 \geq p > \phi(q), \\ h & \text{for } \phi(q) \geq p \geq q_h. \end{cases} \tag{19}$$

This pair of strategies constitute an equilibrium if $V^*(p; Q^1) > 0$ for all $p \in (q_h, p_0]$. It is obvious that this holds for $p \in (q_h, \phi(q_h)]$ during which the agent exerts high effort. It thus suffices to show that $V^*(p; Q^1) > 0$ for all $p \in (\phi(q_h), p_0]$. The following fact is convenient to this end.

Lemma 4 *Fix the principal's strategy $Q^1 = \{q_h\}$ and the agent's best response $A^*_{Q^1}$. Then, $V^*(p; Q^1) > 0$ for all $p \in (\phi(q_h), p_0]$ if and only if $V^*(\min\{p_0, q_l\}; Q^1) > 0$.*

Proof: Since the necessity is evident by definition, we only prove the sufficiency part. Suppose first that $p_0 \geq q_l$. In this case, $V^*(q_l; Q^1) > 0$ implies $V^*(p; Q^1) > 0$ for all $p > q_l$ because the instantaneous payoff is strictly positive for any effort choice when the belief is above q_l. Given this, it suffices to show that $V^*(p;Q^1) > 0$ for all $p \in (\phi(q_h), q_l)$ if $V^*(q_l;Q^1) > 0$. To show this, suppose on the contrary that $V^*(q_l;Q^1) > 0$ but $V^*(p;Q^1) \leq 0$ for some $p \in (\phi(q_h), q_l)$. If this is the case, there must exist some $\hat{p} \in (\phi(q_h), q_l)$ such that $V^*(\hat{p};Q^1) = 0$. Since

$$\begin{aligned} V^*(p;Q^1) = &\frac{p(\lambda_l y - w)}{\lambda_l + r}\big(1 - e^{-(\lambda_l+r)(\tau^*(p')-\tau^*(p))}\big) - \frac{(1-p)w}{r}\big(1 - e^{-r(\tau^*(p')-\tau^*(p))}\big) \\ &+ (1 - p + pe^{-\lambda_l(\tau^*(p')-\tau^*(p))})e^{-r(\tau^*(p')-\tau^*(p))}V^*(p';Q^1), \end{aligned}$$

for any $p > p' > \phi(q_h)$, due to the recursive structure of the expected payoff, we have

$$\begin{aligned} V^*(q_l;Q^1) = &\frac{q_l(\lambda_l y - w)}{\lambda_l + r}\big(1 - e^{-(\lambda_l+r)(\tau^*(\hat{p})-\tau^*(q_l))}\big) - \frac{(1-q_l)w}{r}\big(1 - e^{-r(\tau^*(\hat{p})-\tau^*(q_l))}\big) \\ &+ (1 - q_l + q_l e^{-\lambda_l(\tau^*(\hat{p})-\tau^*(q_l))})e^{-r(\tau^*(\hat{p})-\tau^*(q_l))}V^*(\hat{p};Q^1) \\ = &\frac{q_l(\lambda_l y - w)}{\lambda_l + r}\big(1 - e^{-(\lambda_l+r)(\tau^*(\hat{p})-\tau^*(q_l))}\big) - \frac{(1-q_l)w}{r}\big(1 - e^{-r(\tau^*(\hat{p})-\tau^*(q_l))}\big), \end{aligned}$$

which is strictly negative by definition, and therefore a contradiction.

Now suppose that $q_l > p_0$. As above, it suffices to show that $V^*(p;Q^1) > 0$ for all $p \in (\phi(q_h), p_0)$ if $V^*(p_0;Q^1) > 0$. We can apply the same reasoning to show that this holds. ∎

The lemma shows that we only need to check the continuation payoff at $x := \min\{p_0, q_l\}$ to see whether a given strategy is sequentially rational. We start with the candidate strategy considered above, in which the agent exerts low effort until the belief reaches $\phi(q_h)$ and switches to high effort from that point on (until the belief reaches q_h). The principal has no incentive to deviate and stop if the expected payoff at x is positive, i.e.,

$$\begin{aligned} V^*(x;Q^1) = &\int_{\tau^*(x)}^{\sigma^*(q_h)} (\lambda_l p_s y - w)e^{-\int_{\tau^*(x)}^{s}(\lambda_l p_u + r)du} ds \\ &+ e^{-\int_{\tau^*(x)}^{\sigma^*(q_h)}(\lambda_l p_u + r)du} \int_{\sigma^*(q_h)}^{\tau^*(q_h)} (\lambda_h p_s y - w)e^{-\int_{\sigma^*(q_h)}^{s}(\lambda_h p_u + r)du} ds > 0. \end{aligned}$$

Letting $\Delta_\tau = \tau^*(\phi(q_h)) - \tau^*(x)$, the condition can now be written as

$$\begin{aligned} &\frac{r(\lambda_l y - w)}{\lambda_l + r}\big(1 - e^{-(\lambda_l+r)\Delta_\tau}\big) + \frac{r(\lambda_h y - w)}{\lambda_h + r}e^{-(\lambda_l+r)\Delta_\tau}\big(1 - e^{-(\lambda_h+r)k^*}\big) \\ &> \frac{(1-r)w}{r}\big(1 - e^{-r(\Delta_\tau + k^*)}\big). \end{aligned} \tag{20}$$

If (20) holds, we can apply the same argument as in Case 2 to show the existence and uniqueness of the equilibrium.

The situation becomes more complicated when (20) fails to hold, in which case the principal has an incentive to deviate and stop at some point, and the candidate strategy no longer constitutes an equilibrium. In this case, there instead exists a critical belief that satisfies (20) with equality, because $V^*(p; Q^1)$ is continuous in p with $V^*(\phi(q_h); Q^1) > 0 > V^*(r; Q^1)$. Define q^2 such that $q^2 := \min\{q \in (\phi(q_h), x] : V^*(q; Q^1) = 0\}$ under the candidate strategy $A^*_{Q^1}(p)$ in (19). Given the critical belief, we can then consider $Q^2 = \{q^1, q^2\}$, where $q^1 = q_h$, and the best response

$$A^*_{Q^2}(p) = \begin{cases} 0 & \text{for } p_0 \geq p > \min\{p_0, \phi(q^2)\}, \\ 1 & \text{for } \min\{p_0, \phi(q^2)\} \geq p \geq q_2, \\ 0 & \text{for } q^2 > p > \phi(q_h), \\ 1 & \text{for } \phi(q_h) > p \geq q_1 = q_h. \end{cases}$$

Note that if q^2 exists, $q^2 > \phi(q^1) > q^1 = q_h$.

The game is now divided into two segments (P^1, P^2). First, we can show that the equilibrium is unique in P^1 (as done in Case 2). For P^2, we can apply the same procedure to the truncated interval $[q^2, p_0]$ to see whether Q^2 can be an equilibrium strategy. More precisely, we need to check if $V^*(p; Q^2) > 0$ for all $p \in [p_0, \phi(q^2))$. If this condition holds, $Q^* = Q^2$. If not, we can find yet another critical belief $q^3 \in (\phi(q^2), x]$ and repeat the same process until we find a strategy that can satisfy the equilibrium conditions. The existence of an equilibrium is assured because $q^{i+1} > \phi(q^i) > q^i$ for all $i = 1, 2, ..., n-1$, so that we are bound to have q^i at some point that satisfies either (i) $\phi(q^i) \geq p_0 \geq q^i$ or (ii) $p_0 > \phi(q^i) \geq q_l$. The uniqueness is also evident in this setup: (i) the equilibrium is unique in P^1; (ii) moving backwards, we can also show that the equilibrium in each P^i is unique for $i = 2, 3, ..., n$. ∎

Proof of Proposition 4: The agent always exerts high effort in the noncooperative case if (2) fails to hold. An analogous condition for the cooperative case is given by

$$\frac{\lambda_h(y+b) - d}{\lambda_h + r} > y + b - \frac{d}{\Delta_\lambda}.$$

Applying the same argument as in Proposition 1, we obtain

$$\tilde{k}^* = -\frac{1}{\lambda_h + r} \ln\left(1 - \frac{(\lambda_h + r)(y + b - \frac{d}{\Delta_\lambda})}{\lambda_h(y+b) - d}\right).$$

Observe that $\tilde{k}^* > 0$ under Assumption 1. It follows from this that $\tilde{k}^* > k^*$ if

$$-\frac{1}{\lambda_h + r} \ln\left(1 - \frac{(\lambda_h + r)(y + b - \frac{d}{\Delta_\lambda})}{\lambda_h(y+b) - d}\right) > -\frac{1}{\lambda_h + r} \ln\left(1 - \frac{(\lambda_h + r)(b - \frac{d}{\Delta_\lambda})}{\lambda_h b - d + w}\right),$$

which is reduced to

$$\frac{y + b - \frac{d}{\Delta_\lambda}}{\lambda_h(y+b) - d} > \frac{b - \frac{d}{R_\lambda}}{\lambda_h b - d + w}.$$

With some algebra, we obtain

$$\frac{\lambda_l d y}{\Delta_\lambda} + w\big(y + b - \tfrac{d}{\Delta_\lambda}\big) > 0,$$

which always holds. ∎

Proof of Proposition 5: Given some threshold $\tilde{k}^*$, the only possibility is to set a deadline at time $\tilde{k}^*$, so that the agent exerts high effort from the beginning. This is because any other finite deadline yields a strictly lower payoff. Let $\tilde{U}_\infty := \tilde{V}^*(p;0)$ denote the continuation payoff when the project continues indefinitely. We show that we cannot have $\tilde{U}(k) > \tilde{U}_\infty$ for any finite k. Since, by definition,

$$y + b - \frac{d}{\Delta_\lambda} = \tilde{U}(\tilde{k}^*) = \frac{\big(\lambda_h(y+b) - d\big)}{\lambda_h + r}\big(1 - e^{-(\lambda_h + r)\tilde{k}^*}\big).$$

it is optimal to set a deadline at time $\tilde{k}^*$ if and only if

$$y + b - \frac{d}{\Delta_\lambda} > \tilde{U}_\infty.$$

This is the case, however, where (8) fails to hold, and the agent always exerts high effort, i.e., $\tilde{k}^* \to \infty$. On the other hand, if (8) holds, it is strictly better to let the project continue indefinitely as $\tilde{U}_\infty \geq U(\tilde{k}^*)$. This shows that no finite deadline can improve the social surplus in the cooperative case. ∎

Proof of Proposition 8: We first consider the case where the principal commits to a deadline. Note that the principal cannot terminate the project even if the shock arrives before the deadline. Since the agent has no incentive to exert high effort, we have $\tilde{U}^C(k) = \frac{w(1-e^{-rk})}{r}$. Let $U_a^C(k)$ denote the value function under commitment when the effort level is fixed at a. We then obtain

$$U_h^C(k) = \frac{b\lambda_h - d + w}{\lambda_h + \beta + r}\big(1 - e^{-(\lambda_h+\beta+r)k}\big) + \frac{\beta w}{r}\Big(\frac{1 - e^{-(\lambda_h+\beta+r)k}}{\lambda_h + \beta + r} - \frac{e^{-rk}\big(1 - e^{-(\lambda_h+\beta)k}\big)}{\lambda_h + \beta}\Big).$$

The threshold under commitment is a solution to $U_h^C(k^C) = b - \frac{d}{\Delta_\lambda}$.

The problem is much more straightforward when the principal makes no commitment. In this case, the optimal choice for the principal is to terminate the project as soon as the shock arrives, which implies $\tilde{U}^{NC}(k) = 0$ for all k. As above, define $U_a^{NC}(k)$ as the value function under no commitment with fixed effort a, which is given by

$$U_h^{NC}(k) = \frac{b\lambda_h - d + w}{\lambda_h + \beta + r}\big(1 - e^{-(\lambda_h+\beta+r)k}\big).$$

If $\frac{b\lambda_h - d + w}{\lambda_h+\beta+r} > b - \frac{d}{\Delta_\lambda} > 0$, there exists an interior threshold k^{NC} given by

$$k^{NC} = -\frac{1}{\lambda_h+\beta+r}\ln\Big(1 - \frac{(\lambda_h+\beta+r)(b-\frac{d}{\Delta_\lambda})}{b\lambda_h - d + w}\Big).$$

It is evident that $k^{NC} > k^C$ because the expected payoff must be higher under commitment for a given horizon. Formally, we need to show that

$$\frac{1-e^{-(\lambda_h+\beta+r)k}}{\lambda_h+\beta+r} > \frac{e^{-rk}(1-e^{-(\lambda_h+\beta)k})}{\lambda_h+\beta},$$

which can also be written as

$$\int_0^k e^{-(\lambda_h+\beta+r)s}ds > \int_0^k e^{-(\lambda_h+\beta)s-rk}ds.$$

This condition holds because

$$e^{-(\lambda_h+\beta+r)s} > e^{-(\lambda_h+\beta)s-rk} \;\Leftrightarrow\; e^{-rs} > e^{-rk},$$

for any $s < k$.

Finally, note that $k^C, k^{NC} \in (0,\infty)$ when $\frac{b\lambda_h-d+w}{\lambda_h+\beta+r} > b - \frac{d}{\Delta_\lambda} > 0$. We can then show that k^j, $j = C, NC$, is increasing in β if $U_a^j(k)$ is decreasing in β for $k \le k^j$. Straightforward computation shows that $U_h^{NC}(k)$ is strictly decreasing in β. In contrast, it is a little more involved to show the same for $U_h^C(k)$. To see this, with some computation, we obtain

$$\begin{aligned}
U_h^C(k) &= \frac{b\lambda_h - d}{\lambda_h+\beta+r}(1-e^{-(\lambda_h+\beta+r)k}) - \frac{\lambda_h w}{r}\Big(\frac{1-e^{-(\lambda_h+\beta+r)k}}{\lambda_h+\beta+r} - \frac{e^{-rk}(1-e^{-(\lambda_h+\beta)k})}{\lambda_h+\beta}\Big)\\
&\quad + \frac{w(1-e^{-rk})}{r}\\
&= \frac{b\lambda_h - d - \lambda_h\frac{w(1-e^{-rk})}{r}}{\lambda_h+\beta+r}(1-e^{-(\lambda_h+\beta+r)k}) + \frac{\lambda_h w e^{-rk}}{r}\Big(\frac{1-e^{-(\lambda_h+\beta)k}}{\lambda_h+\beta} - \frac{1-e^{-(\lambda_h+\beta+r)k}}{\lambda_h+\beta+r}\Big)\\
&\quad + \frac{w(1-e^{-rk})}{r}.
\end{aligned}$$

First, it is easy to verify that $\frac{1-e^{-\alpha k}}{\alpha}$ is strictly decreasing for any $\alpha > 0$. Further, observe that the second term is strictly decreasing in β. These two facts imply that it suffices to show that

$$b\lambda_h - d > \lambda_h\frac{w(1-e^{-rk})}{r},$$

for $k \le k^C$. To show this, since $b\lambda_h - d > b\lambda_h - \frac{\lambda_h}{\Delta_\lambda}d \ge \lambda_h U_h^C(k)$ for $k \le k^C$, we have

$$\begin{aligned}
U_h^C(k) &> \frac{\lambda_h(U_h^C(k) - \frac{w(1-e^{-rk})}{r})}{\lambda_h+\beta+r}(1-e^{-(\lambda_h+\beta+r)k}) + \frac{\lambda_h w e^{-rk}}{r}\Big(\frac{1-e^{-(\lambda_h+\beta)k}}{\lambda_h+\beta} - \frac{1-e^{-(\lambda_h+\beta+r)k}}{\lambda_h+\beta+r}\Big)\\
&\quad + \frac{w(1-e^{-rk})}{r},
\end{aligned}$$

for $k \le k^C$. This is further reduced to

$$\Big(1 - \frac{\lambda_h(1-e^{-(\lambda_h+\beta+r)k})}{\lambda_h+\beta+r}\Big)\Big(U_h^C(k) - \frac{w(1-e^{-rk})}{r}\Big) > \frac{\lambda_h w e^{-rk}}{r}\Big(\frac{1-e^{-(\lambda_h+\beta)k}}{\lambda_h+\beta} - \frac{1-e^{-(\lambda_h+\beta+r)k}}{\lambda_h+\beta+r}\Big),$$

which gives $U_h^C(k) > \frac{w(1-e^{-rk})}{r}$. ∎

Proof of Proposition 10: As $w \to 0$, we have $q_h \to 0$, which implies $\tau^*(q^{NC}) \to \infty$. The expected payoff converges to

$$\lim_{w\to 0} V^*(p_0, q^{NC}) = \frac{p_0\lambda_l y}{\lambda_l + r}, \tag{21}$$

for any p_0. Let

$$\mu(w) := \frac{\lambda_h y - w}{w}\Big(1 - \frac{(\lambda_h+r)(b-\frac{d}{\Delta_\lambda})}{\lambda_h b - d + \tilde{w}}\Big)^{\frac{\lambda_h}{\lambda_h+r}}.$$

Since $\lim_{w\to 0}\mu(w) = \infty$ and $\mu(\lambda_h y) = 0$, there must exist at least one $\tilde{w}$ such that

$$\mu(\tilde{w}) = \frac{1-p_0}{p_0}.$$

If we set $w = \tilde{w}$, then $\tau^*(q^{NC}) = k^* > 0$ by definition, and the expected payoff is

$$V(p_0, q^{NC}) = \int_0^{k^*} (\lambda_h y p_s y - w)e^{-\int_0^s(\lambda_h p_u + r)du}ds,$$

which is always strictly positive and larger than (21) if λ_l is sufficiently small. ∎

Appendix B: A discrete-time model

Our continuous-time setup is particularly convenient as it allows us to obtain a complete characterization of equilibrium in a very tractable manner. The nature of the equilibrium remains the same, however, even when actions can be taken only at discrete points in time. In this appendix, we provide a discrete counterpart of our model to show how we can construct an equilibrium and also that the constructed equilibrium converges to the continuous-time equilibrium as the time interval vanishes to zero. Since the commitment case is essentially the same, we only consider the no-commitment case in this appendix.

Setup: We consider the same setting as in the continuous-time case, except that actions can now be taken only at a fixed interval $\Delta > 0$. We refer to each interval as a "period," each denoted by an integer $t = 1, 2, 3, ...$ The timing of events within each period goes as follows. At the beginning of each period, the principal first chooses whether to continue the project. If the principal chooses to continue, the agent then chooses the effort level, and the outcome is realized. The game ends if the agent attains a success. If not, the game proceeds to the next period. The common discount factor is replaced by $\delta := e^{-r\Delta}$. Further, define $\theta_a := \frac{1-e^{-\lambda_a \Delta}}{\Delta}$ as the success probability (per unit time) for a given effort level $a = l, h$.

Agent's effort decision: Let K denote the number of remaining periods, which is an analogous concept to the remaining time in the continuous-time case: at the beginning of period t, if the principal is to terminate the project at the beginning of period T, $K = T - t$. Let $U(K)$ denote the agent's value function when there are K remaining periods. Similarly to the continuous-time case, we show that the good type exerts high effort only in the last few periods.

Proposition B.1 *There exists K^* such that the good type exerts high effort if and only if the number of remaining periods is fewer than or equal to K^*.*

Proof: We first show that $U(K)$ is strictly increasing in K. Note that

$$U(1) = \max\{\theta_h b - d + w, \theta_l b + w\}\Delta > U(0) = 0.$$

It follows from this that

$$\begin{aligned} U(2) = &\max\{(\theta_h b - d + w)\Delta + (1 - \theta_h \Delta)\delta U(1), (\theta_l b + w)\Delta + (1 - \theta_l \Delta)\delta U(1)\} \\ &> U(1) = \max\{(\theta_h b - d + w)\Delta, (\theta_l b + w)\Delta\}. \end{aligned}$$

Applying the same argument backward, we obtain $U(K) > U(K-1)$.

Given that there are K remaining periods, the good type exerts high effort if

$$(\theta_h b - d + w)\Delta + (1 - \theta_h \Delta)\delta U(K-1) \geq (\theta_l b + w)\Delta + (1 - \theta_l \Delta)\delta U(K-1).$$

This holds when $U(K)$ is small enough. Since $U(K)$ increases with K, there exists K^* such that the good type exerts high effort if and only if the number of remaining periods is fewer than or equal to K^*.[34] ∎

Principal's termination decision: Since the belief makes discrete jumps on the equilibrium path, it is more convenient to define the principal's continuation payoff as a function of the number of remaining periods K rather than of the termination beliefs as we did in the continuous-time case. The principal's strategy is represented by a set of termination dates

[34] When $\Delta \to 0$, $K^* \to \infty$, but $K^*\Delta$ converges to k^*.

$S := \{T^i\}_{i=1}^n$. Let $V_t(p; S, A)$ denote the principal's continuation payoff at the beginning of period t under strategies A and S. Straightforward computation yields

$$V_t(p_t; S, A) = \sum_{s=t}^{t+K_t-1} \Big(\delta^{s-t} \prod_{u=0}^{s-t-1} (1 - p_{t+u}\theta_{a_{t+u}}\Delta)\Big)(p_s\theta_{a_s} b - da_s + w)\Delta,$$

where K_t is the number of remaining periods in period t, subject to

$$p_{t+1} = D_{a_t}(p_t) := \frac{p_t e^{-\lambda_{a_t}\Delta}}{1 - p_t + p_t e^{-\lambda_{a_t}\Delta}} = \frac{p_t(1 - \theta_{a_t}\Delta)}{1 - p_t\theta_{a_t}\Delta}.$$

We can characterize the principal's termination strategy in essentially the same manner with some technical modifications. One difference is that since the belief now makes discrete jumps, the termination belief in the discrete-time case is given by a set of continuous intervals rather than of discrete points. As in the continuous-time case, it is a strictly dominated strategy to continue the project if $p_t \le q_h$. More precisely, the game must end if $p_t \in \Theta_h := (D_h(q_h), q_h]$, which is the lowerbound of the belief for the discrete-time case. Notice that the lowerbound degenerates to a single point q_h, the lowerbound for the continuous-time case, as $\Delta \to 0$.

As in the continuous-time model, the principal may also need to stop before the belief reaches Θ_h. The principal terminates the project at the beginning of period T^i if the continuation payoff is weakly negative in period T^i but is positive from the next period on. More precisely, the principal terminates the project if

$$V_{T^i}(p_{T^i}; S, A) \le 0 \text{ and } V_t(p_t; S, A) > 0,$$

for all $t = T^i + 1, ..., T^{i+1} - 1$.

Equilibrium construction: The fact that the principal's termination strategy is characterized by intervals gives rise to a technical issue that would not arise in the continuous-time case. To see this, suppose that there exists some T such that $p_T \in \Theta_h$, so that the principal terminates the project at the beginning of period T. The agent then starts exerting high effort from period $T - K^*$, and the principal has no incentive to terminate the project until period T. Before period $T - K^*$, however, there is an interval during which the agent exerts low effort and the principal's continuation payoff may become negative at some point. Letting $K := T - t$ denote the number of remaining periods, consider the agent's strategy $\hat{A}$ given by

$$a_t = \hat{A}(K) = \begin{cases} h & \text{if } K^* \ge K, \\ l & \text{if } K > K^*, \end{cases}$$

and define a backward operator $B_K(p)$, where $p_{T-K} = B_K(p)$ for some $p \in \Theta_h$ under the strategy $\hat{A}$. Let $\hat{V}(p; K) := V_{T-K}(p; S, \hat{A})$.

To construct an equilibrium, it is convenient to define the following notations. First, define $\nu_K := (B_K(D_h(q_h)), B_K(q_h)]$: if a continuation game starts from some $p \in \nu_K$, the belief reaches the lowerbound Θ_h after K periods under $\hat{A}$. Note that $\nu_K \cap \nu_{K-1}$ is generally nonempty, and when the belief falls in this overlapped range, there are (at least) two ways to reach the lowerbound Θ_h. Second, for a given $p \in \Theta_h$, we may find some $\kappa(p) \in [K^*+1, T-1]$ such that

$$\hat{V}(B_{\kappa(p)}(p); \kappa(p)) \leq 0 \text{ and } \hat{V}(D_l(B_{\kappa(p)}(p)); \kappa(p) - 1) > 0.$$

If there is no such $\kappa(p)$, there is no incentive for the principal to terminate the project until period T and we let $\kappa(p) = T$ in this case. Moreover, define $\bar{p} := \max\{p \in \Theta_h : \kappa(p) < T\}$, $\bar{\kappa} := \kappa(\bar{p})$, and $\underline{\kappa} := \kappa(D_h(q_h))$. Note that $\underline{\kappa} \geq K^* + 1$ because the continuation payoff must be positive if the number of remaining periods is less than or equal to K^*.

Let $\bar{p}_K := \max\{p \in \Theta_h : \kappa(p) = K\}$ for $K \in [\underline{\kappa}, \bar{\kappa}]$ and $\bar{p}_{\underline{\kappa}-1} = D_h(q_h)$. We construct a continuation equilibrium for the interval $\nu_{\bar{\kappa}}$ in which the principal may need to terminate the project before period T.[35] To this end, note first that if $q_h > \bar{p}$, there exists a nonempty interval $(B_{\bar{\kappa}}(\bar{p}), B_{\bar{\kappa}}(q_h)]$ in which the principal's continuation payoff is invariably positive until period T under $\hat{A}$: in this case, there exists a well-behaved continuation equilibrium in which the principal never terminates the project until the belief reaches Θ_h. It thus suffices to consider a (possibly truncated) interval $\bar{\nu}_{\bar{\kappa}} := (B_{\bar{\kappa}}(D_h(q_h)), B_{\bar{\kappa}}(\bar{p})]$.

For $p \in \bar{\nu}_{\bar{\kappa}}$, the principal needs to terminate the project at some point. More precisely, we go through the following procedure, starting from step 0, until we reach a point where the principal has an incentive to stop.

Step j

Consider a continuation game with $\bar{\kappa} - j$ remaining periods. Any $p \in \nu_{\bar{\kappa}-j}$ is classified into one of the following four cases.

Case 1: $p \in (B_{\bar{\kappa}-j}(\bar{p}_{\bar{\kappa}-1-j}), B_{\bar{\kappa}-j}(\bar{p}_{\bar{\kappa}-1})]$

There always exists a nonempty interval $(B_{\bar{\kappa}-j}(\bar{p}_{\kappa-1-j}), B_{\bar{\kappa}-j}(\bar{p}_{\bar{\kappa}-j})]$ in which the continuation payoff is weakly negative but is positive from the next period. When the current belief falls in this interval, there exists a well-behaved continuation equilibrium in which the principal continues the project until period T and the players play for $\bar{\kappa}$ periods. The principal thus has an incentive to stop when the belief is in this interval.

Case 2: $p \in (B_{\bar{\kappa}-1-j}(q_h), B_{\bar{\kappa}-j}(\bar{p}_{\kappa-1-j})]$

[35] Note that each ν_K is by definition sufficiently long, so that any belief path starting from above must eventually fall in this interval under an arbitrary strategy.

For $p \in (B_{\bar{\kappa}-1-j}(q_h), B_{\bar{\kappa}-j}(\bar{p}_{\bar{\kappa}-1-j})]$, which may or may not be empty, the continuation payoff in the next period is also negative under $\hat{A}$. The principal must then continue the project for one more period, and the agent exerts high effort. Since $D_h(p) \in \nu_{\bar{\kappa}-1-j}$ by definition, we proceed to step $j+1$ with the updated belief $D_h(p)$.

Case 3: $p \in (\max\{B_{\bar{\kappa}-1-j}(\bar{p}_{\kappa-1-j}), B_{\bar{\kappa}-j}(D_h(q_h))\}, \min\{B_{\bar{\kappa}-1-j}(q_h), B_{\bar{\kappa}-j}(\bar{p}_{\kappa-1-j})\}]$

Note that $p \in \nu_{\bar{\kappa}-j} \cap \nu_{\bar{\kappa}-1-j}$ in this case, meaning that there are (at least) two ways to reach Θ_h. The problem is that neither constitutes an equilibrium by itself: if the players play for $\bar{\kappa}-j$ periods,

$$\hat{V}(p; \bar{\kappa}-j) < 0 \text{ and } \hat{V}(D_l(p); \bar{\kappa}-1-j) \leq 0,$$

whereas if they play for $\bar{\kappa}-1-j$ periods,

$$\hat{V}(p; \bar{\kappa}-1-j) > 0 \text{ and } \hat{V}(D_l(p; \bar{\kappa}-2-j) > 0.$$

Our equilibrium construction in this interval exploits the idea of public randomization, much in the spirit of repeated games. Since

$$\hat{V}(D_l(p); \bar{\kappa}-2-j) > \hat{V}(p; \bar{\kappa}-1-j) > 0 > V(D_l(p); \bar{\kappa}-1-j) > \hat{V}(p; \bar{\kappa}-j),$$

there exists a range of γ such that

$$\begin{aligned}
&\gamma\hat{V}(p; \bar{\kappa}-j) + (1-\gamma)\hat{V}(D_l(p); \bar{\kappa}-1-j) \leq 0, \\
&\gamma\hat{V}(p; \bar{\kappa}-1-j) + (1-\gamma)V(D_l(p); \bar{\kappa}-2-j) > 0.
\end{aligned}$$

That is, possibly with the aid of some public randomization device, we can construct a well-behaved continuation equilibrium where the players play for $\bar{\kappa}-j$ periods with probability γ and for $\bar{\kappa}-1-j$ periods with probability $1-\gamma$. Given this, the principal has an incentive to stop when the belief falls in this interval.

Case 4: $p \in (B_{\bar{\kappa}-j}(D_h(q_h)), B_{\bar{\kappa}-1-j}(\bar{p}_{\kappa-1-j})]$

This is an interval, if any, where $\hat{V}(p; \bar{\kappa}-1-j) \leq 0$, meaning that $\kappa(p) < \bar{\kappa}-j$. We directly proceed to step $j+1$ with the current belief.

In cases 2 and 4, the process goes to the next step. This process must stop eventually as it cannot go beyond $K^*+1 > \bar{\kappa}-j$. To see this, suppose that there exists a well-defined $\underline{\kappa} = K^*+1$ and the belief falls in $(B_{K^*+1}(D_h(q_h)), B_{K^*+1}(\bar{p}_{K^*+1})]$ in step $\bar{\kappa}-K^*-1$. Since $\bar{p}_{K^*} = D_h(q_h)$, however, this must correspond to Case 1.

Once we identify when to stop, we can again go back in time from that point and see if the continuation payoff stays positive on the equilibrium path. If there is a point at which the continuation payoff becomes negative, we go through the same procedure again to construct a continuation equilibrium.

Finally, it is clear that as $\Delta \to 0$, Θ_h degenerates to q_h. As such, all the termination intervals eventually converge to q^2, q^3,..., q^n.

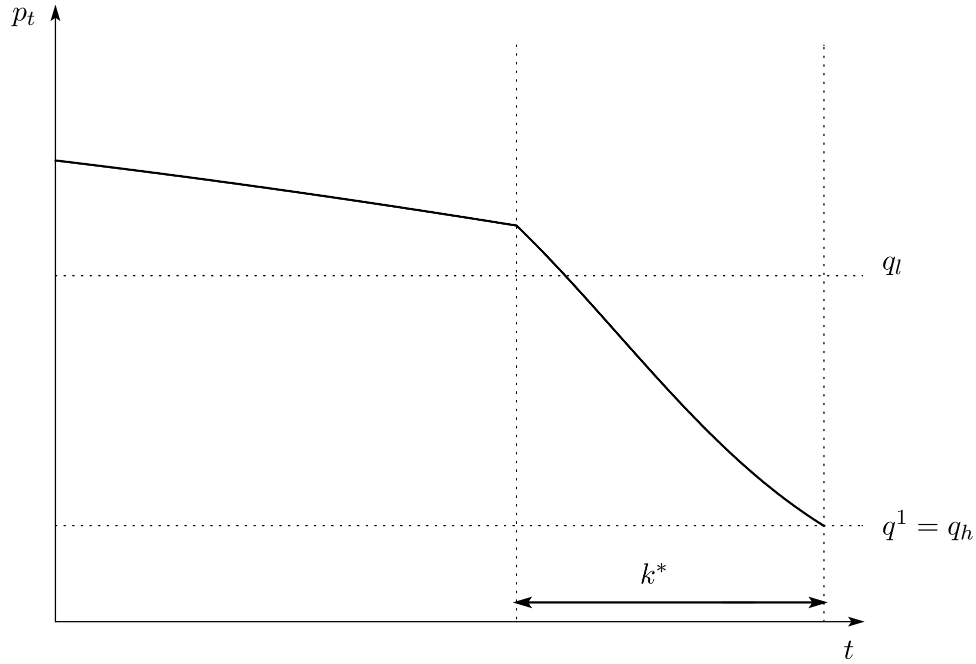

Figure 1: Evolution of the belief on the equilibrium path

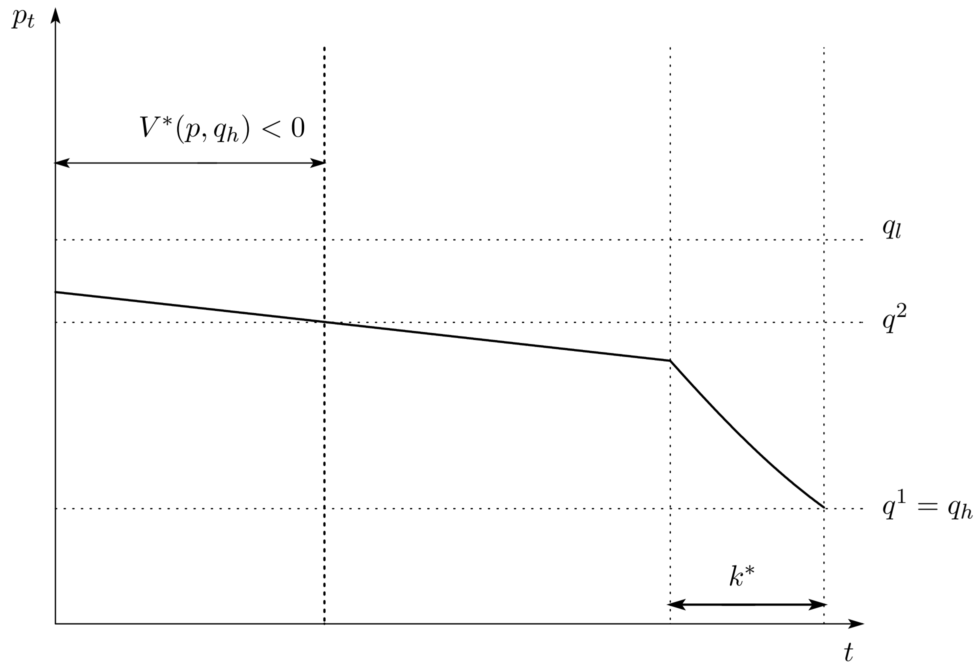

Figure 2: A violation of Condition C: the belief

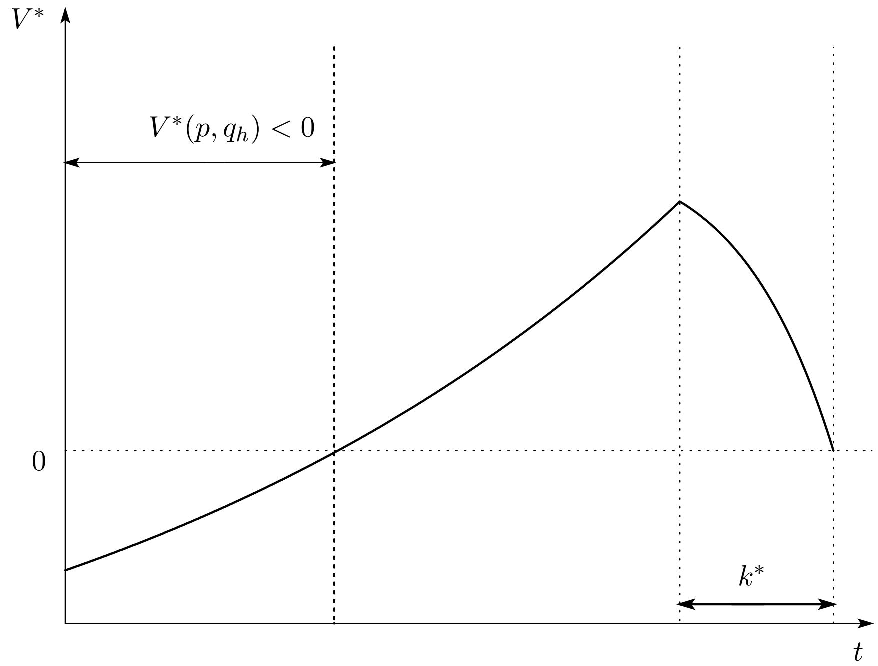

Figure 3: A violation of Condition C: the continuation payoff

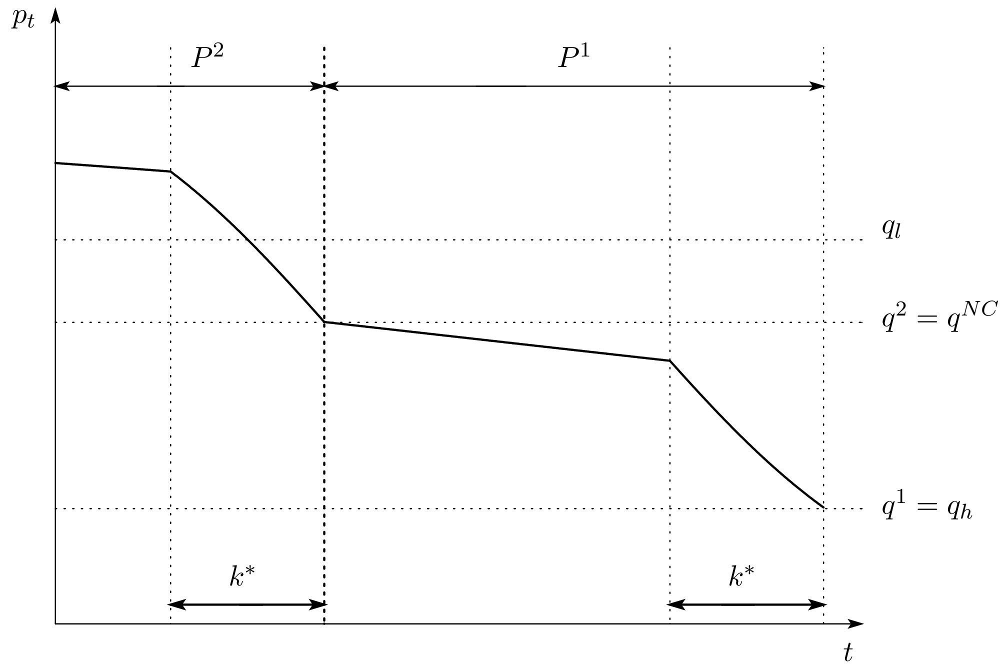

Figure 4: Evolution of the belief, on and off the equilibrium path