

# Optimal Contracts for Experimentation

MARINA HALAC

*Columbia University and University of Warwick*

NAVIN KARTIK

*Columbia University*

and

QINGMIN LIU

*Columbia University*

*First version received September 2013; final version accepted January 2016 (Eds.)*

This paper studies a model of long-term contracting for experimentation. We consider a principal–agent relationship with adverse selection on the agent’s ability, dynamic moral hazard, and private learning about project quality. We find that each of these elements plays an essential role in structuring dynamic incentives, and it is only their interaction that generally precludes efficiency. Our model permits an explicit characterization of optimal contracts.

*Key words:* Innovation, Learning, Adverse selection, Moral hazard

*JEL Codes:* D82, D83, D86

## 1. INTRODUCTION

Agents need to be incentivized to work on, or experiment with, projects of uncertain feasibility. Particularly with uncertain projects, agents are likely to have some private information about their project-specific skills.<sup>1</sup> Incentive design must deal with not only dynamic moral hazard, but also adverse selection (pre-contractual hidden information) and the inherent process of learning. To date, there is virtually no theoretical work on contracting in such settings. How well can a principal incentivize an agent? How do the environment’s features affect the shape of optimal incentive contracts? What distortions, if any, arise? An understanding is relevant not only for motivating research and development, but also for diverse applications like contract farming, technology adoption, and book publishing, as discussed subsequently.

This paper provides an analysis using a simple model of experimentation. We show that the interaction of learning, adverse selection, and moral hazard introduces new conceptual and

1. Other forms of private information, such as beliefs about the project feasibility or personal effort costs, are also relevant; see Subsection 7.4.

analytical issues, with each element playing a role in structuring dynamic incentives. Their interaction affects social efficiency: the principal typically maximizes profits by inducing an agent of low ability to end experimentation inefficiently early, even though there would be no distortion without either adverse selection or moral hazard. Furthermore, despite the intricacy of the problem, intuitive contracts are optimal. The principal can implement the second best by selling the project to the agent and committing to buy back output at time-dated future prices; these prices must increase over time in a manner calibrated to deal with moral hazard and learning.

Our model builds on the now-canonical two-armed “exponential bandit” version of experimentation (Keller *et al.*, 2005).<sup>2</sup> The project at hand may either be good or bad. In each period, the agent privately chooses whether to exert effort (work) or not (shirk). If the agent works in a period and the project is good, the project is successful in that period with some probability; if either the agent shirks or the project is bad, success cannot obtain in that period. In the terminology of the experimentation literature, working on the project in any period corresponds to “pulling the risky arm”, while shirking is “pulling the safe arm”; the opportunity cost of pulling the risky arm is the effort cost that the agent incurs. Project success yields a fixed social surplus, accrued by the principal, and obviates the need for any further effort. We introduce adverse selection by assuming that the probability of success in a period (conditional on the agent working and the project being good) depends on the agent’s ability—either high or low—which is the agent’s ex-ante private information or type. Our baseline model assumes no other contracting frictions, in particular we set aside limited liability and endow the principal with full ex-ante commitment power: she maximizes profits by designing a menu of contracts to screen the agent’s ability.<sup>3</sup>

Since beliefs about the project’s quality decline so long as effort has been exerted but success not obtained, the first-best or socially efficient solution is characterized by a stopping rule: the agent keeps working (so long as he has not succeeded) up until some point at which the project is permanently abandoned. An important feature for our analysis is that the efficient stopping time is a non-monotonic function of the agent’s ability. The intuition stems from two countervailing forces: on the one hand, for any given belief about the project’s quality, a higher-ability agent provides a higher marginal benefit of effort because he succeeds with a higher probability; on the other hand, a higher-ability agent also learns more from the lack of success over time, so at any point he is more pessimistic about the project than the low-ability agent. Hence, depending on parameter values, the first-best stopping time for a high-ability agent may be larger or smaller than that of a low-ability agent (cf. Bobtcheff and Levy, 2015).

Turning to the second best, the key distinguishing feature of our setting from a canonical (static) adverse selection problem is the dynamic moral hazard and its interaction with the agent’s private learning. Recall that in a standard buyer-seller adverse selection problem, there is no issue about what quantity the agent of one type would consume if he were to deviate and take the other type’s contract: it is simply the quantity specified by the chosen contract. In contrast, in our setting, it is not *a priori* clear what “consumption bundle”, *i.e.* effort profile, each agent type will choose after such an off-the-equilibrium path deviation. Dealing with this problem would not pose any conceptual difficulty if there were a systematic relationship between the two types’ effort profiles, for instance if there were a “single-crossing condition” ensuring that the high type always wants to experiment at least as long as the low type. However, given the nature of learning, there is no such systematic relationship in an arbitrary contract. As effort off the equilibrium path is crucial when optimizing over the menu of contracts—because it affects how much “information rent”

2. As surveyed by Bergemann and Välimäki (2008), learning is often modelled in economics as an experimentation or bandit problem since Rothschild (1974).

3. Subsection 7.2 studies the implications of limited liability. The importance of limited liability varies across applications; we also view it as more insightful to separate its effects from those of adverse selection.

the agent gets—and the contracts in turn influence the agent’s off-path behaviour, we are faced with a non-trivial fixed point problem.

Theorem 2 establishes that the principal optimally screens the agent types by offering two distinct contracts, each inducing the agent to work for some amount of time (so long as success has not been obtained) after which the project is abandoned. Compared to the social optimum, an inefficiency typically obtains: while the high-ability type’s stopping time is efficient, the low-ability type experiments too little. This result is reminiscent of the familiar “no distortion at the top but distortion below” in static adverse selection models, but the distortion arises here only from the conjunction of adverse selection and moral hazard; we show that absent either one, the principal would implement the first best (Theorem 1). Moreover, because of the aforementioned lack of a single-crossing property, it is not immediate in our setting that the principal should not have the low type over-experiment to reduce the high type’s information rent, particularly when the first best entails the high type stopping earlier than the low type.

Theorem 2 is indirect in the sense that it establishes the (in)efficiency result without elucidating the form of second-best contracts. Our methodology to characterize such contracts distinguishes between the two orderings of the first-best stopping times. We first study the case in which the efficient stopping time for a high-ability agent is larger than that of a low-ability agent. Here we show that although there is no analogue of the single-crossing condition mentioned above in an arbitrary contract, such a condition must hold in an *optimal* contract for the low type. This allows us to simplify the problem and fully characterize the principal’s solution (Theorem 3 and Theorem 4). The case in which the first-best stopping time for the high-ability agent is lower than that of the low-ability agent proves to be more challenging: now, as suggested by the first best, an optimal contract for the low type is often such that the high type would experiment less than the low type should he take this contract. We are able to fully characterize the solution in this case under no discounting (Theorem 5 and Theorem 6).

The second-best contracts we characterize take simple and intuitive forms, partly owing to the simple underlying primitives. In any contract that stipulates experimentation for  $T$  periods it suffices to consider at most  $T + 1$  transfers. The reason is that the parties share a common discount factor and there are  $T + 1$  possible project outcomes: a success can occur in each of the  $T$  periods or never. One class of contracts are *bonus contracts*: the agent pays the principal an up-front fee and is then rewarded with a bonus that depends on when the project succeeds (if ever). We characterize the unique sequence of time-dependent bonuses that must be used in an optimal bonus contract for the low-ability type.<sup>4</sup> This sequence is increasing over time up until the termination date. The shape, and its exact calibration, arises from a combination of the agent becoming more pessimistic over time (absent earlier success) and the principal’s desire to avoid any slack in the provision of incentives, while crucially taking into account that the agent can substitute his effort across time.

The optimal bonus contract can be viewed as a simple “sale-with-buyback contract”: the principal sells the project to the agent at the outset for some price, but commits to buy back the project’s output (that obtains with a success) at time-dated future prices. It is noteworthy that contract farming arrangements, widely used in developing countries between agricultural companies and farm producers (Barrett *et al.*, 2012), are often sale-with-buyback contracts: the company sells seeds or other technology (*e.g.* fertilizers or pesticides) to the farmer and agrees to buy back the crop at pre-determined prices, conditional on this output meeting certain quality

4. For the high type, there are multiple optimal contracts even within a given class such as bonus contracts. The reason for the asymmetry is that the low type’s contract is pinned down by information rent minimization considerations, unlike the high type’s contract. Of course, the high type’s contract cannot be arbitrary either.

standards and delivery requirements (Minot, 2007). The contract farming setting involves a profit-maximizing firm (principal) and a farmer (agent). Miyata *et al.* (2009) describe the main elements of these environments, focusing on the case of China. It is initially unknown whether the new seeds or technology will produce the desired outcomes in a particular farm, which maps into our project uncertainty.<sup>5</sup> Besides the evident moral hazard problem, there is also adverse selection: farmers differ in unobservable characteristics, such as industriousness, intelligence, and skills.<sup>6</sup> Our analysis not only shows that sale-with-buyback contracts are optimal in the presence of uncertainty, moral hazard, and unobservable heterogeneity, but elucidates why. Moreover, as discussed further in Subsection 5.3, our paper offers implications for the design of such contracts and for field experiments on technology adoption more broadly. In particular, field experiments might test our predictions regarding the rich structure of optimal bonus contracts and how the calibration depends on underlying parameters.<sup>7</sup>

Another class of optimal contracts that we characterize are *penalty contracts*: the agent receives an up-front payment and is then required to pay the principal some time-dependent penalty in each period in which a success does not obtain, up until either the project succeeds or the contract terminates.<sup>8</sup> Analogous to the optimal bonus contract, we identify the unique sequence of penalties that must be used in an optimal penalty contract for the low-ability type: the penalty increases over time with a jump at the termination date. These types of contracts correspond to those used, for example, in arrangements between publishers and authors: authors typically receive advances and are then required to pay the publisher back if they do not succeed in completing the book by a given deadline (Owen, 2013). This application fits into our framework when neither publisher nor author may initially be sure whether a commercially-viable book can be written in the relevant timeframe (uncertain project feasibility); the author will have superior information about his suitability or comparative advantage in writing the book (adverse selection about ability); and how much time he actually devotes to the task is unobservable (moral hazard).<sup>9</sup>

Our results have implications for the extent of experimentation and innovation across different economic environments. An immediate prediction concerns the effects of asymmetric information: we find that environments with more asymmetric information (either moral hazard or adverse selection) should feature less experimentation, lower success rates, and more dispersion of success rates. We also find that the relationship between success rates and the underlying environment can be subtle. Absent any distortions, “better environments” lead to more innovation. Specifically, an increase in the proportion of high-ability agents or an increase in the ability of both types of the agent yields a higher probability of success in the first best. In the presence of moral hazard and adverse selection, however, the opposite can be true: these changes can induce the principal to distort the low-ability type’s experimentation by more, to the extent that the average success probability goes down in the second best. Consequently, observing higher innovation rates in contractual settings like those we study is neither necessary nor sufficient to deduce

5. Besley and Case (1993) study how farmers learn about a new technology over time given the realization of yields from past planting decisions, and how they in turn make dynamic choices.

6. Beaman *et al.* (2015) provide evidence of such unobservable characteristics using a field experiment in Mali.

7. We should highlight that our paper is not aimed at studying all the institutional details of contract farming or technology adoption. For example, we do not address multi-agent experimentation and social learning, which has been emphasized by the empirical literature (*e.g.* Conley and Udry, 2010).

8. There is a flavour here of “clawbacks” that are sometimes used in practice when an agent is found to be negligent. In our setting, it is the lack of project success that is treated like evidence of negligence (*i.e.* shirking); note, however, that in equilibrium the principal knows that the agent is not actually negligent.

9. Not infrequently, authors fail to deliver in a timely fashion (Suddath, 2012). That private information can be a substantive issue is starkly illustrated by the case of Herman Rosenblat, whose contract with Penguin Books to write a Holocaust survivor memoir was terminated when it was discovered that he fabricated his story.

a better underlying environment. As discussed in Subsection 5.3, these results may contribute an agency-theoretic component to the puzzle of low technology adoption rates in developing countries.

**Related literature.** Broadly, this paper fits into literatures on long-term contracting with either dynamic moral hazard and/or adverse selection. Few papers combine both elements, but two recent exceptions are Sannikov (2007) and Gershkov and Perry (2012).<sup>10</sup> These papers are not concerned with learning/experimentation and their settings and focus differ from ours in many ways.<sup>11</sup> More narrowly, starting with Bergemann and Hege (1998, 2005), there is a fast-growing literature on contracting for experimentation. Virtually all existing research in this area addresses quite different issues than we do, primarily because adverse selection is not accounted for.<sup>12</sup> The only exception we are aware of is the concurrent work of Gomes *et al.* (2015). They do not consider moral hazard; instead, they introduce two-dimensional adverse selection. Under some conditions they obtain an “irrelevance result” on the dimension of adverse selection that acts similar to our agent’s ability, a conclusion that is similar to our benchmark that the first best obtains in our model when there is no moral hazard.

Outside a pure experimentation framework, Gerardi and Maestri (2012) analyse how an agent can be incentivized to acquire and truthfully report information over time using payments that compare the agent’s reports with the ex-post observed state; by contrast, we assume the state is never observed when experimentation is terminated without a success. Finally, our model can also be interpreted as a problem of delegated sequential search, as in Lewis and Ottaviani (2008) and Lewis (2011). The main difference is that, in our context, these papers assume that the project’s quality is known and hence there is no learning about the likelihood of success (cf. Subsection 7.3); moreover, they do not have adverse selection.

## 2. THE MODEL

**Environment.** A principal needs to hire an agent to work on a project. The project’s quality—synonymous with the state—may either be good or bad, a binary variable. Both parties are initially uncertain about the project’s quality; the common prior on the project being good is  $\beta_0 \in (0, 1)$ . The agent is privately informed about whether his ability is low or high,  $\theta \in \{L, H\}$ , where  $\theta = H$  represents “high”. The principal’s prior on the agent’s ability being high is  $\mu_0 \in (0, 1)$ . In each period,  $t \in \{1, 2, \dots\}$ , the agent can either exert effort (work) or not (shirk); this choice is never observed by the principal. Exerting effort in any period costs the agent  $c > 0$ . If effort is exerted and the project is good, the project is successful in that period with probability  $\lambda^\theta$ ; if either the agent shirks or the project is bad, success cannot obtain in that period. Success is observable and

10. Some earlier papers with adverse selection and dynamic moral hazard, such as Laffont and Tirole (1988), focus on the effects of short-term contracting. There is also a literature on dynamic contracting with adverse selection and evolving types but without moral hazard or with only one-shot moral hazard, such as Baron and Besanko (1984) or, more recently, Battaglini (2005), Boleslavsky and Said (2013), and Eső and Szentes (2015). Pavan *et al.* (2014) provide a rather general treatment of dynamic mechanism design without moral hazard.

11. Demarzo and Sannikov (2011), He *et al.* (2014), and Prat and Jovanovic (2014) study private learning in moral-hazard models following Holmström and Milgrom (1987), but do not have adverse selection. Sannikov (2013) also proposes a Brownian motion model and a first-order approach to deal with moral hazard when actions have long-run effects, which raises issues related to private learning. Chassang (2013) considers a general environment and develops an approach to find detail-free contracts that are not optimal but instead guarantee some efficiency bounds so long as there is a long horizon and players are patient.

12. See Bonatti and Hörner (2011, 2015), Manso (2011), Klein (2012), Ederer (2013), Hörner and Samuelson (2013), Kwon (2013), Guo (2014), Halac *et al.* (2015), and Moroni (2015).

once a project is successful, no further effort is needed.<sup>13</sup> We assume  $1 > \lambda^H > \lambda^L > 0$ . A success yields the principal a payoff normalized to 1; the agent does not intrinsically care about project success. Both parties are risk neutral, have quasi-linear preferences, share a common discount factor  $\delta \in (0, 1]$ , and are expected-utility maximizers.

**Contracts.** We consider contracting at period zero with full commitment power from the principal. To deal with the agent’s hidden information at the time of contracting, the principal’s problem is, without loss of generality, to offer the agent a *menu* of dynamic contracts from which the agent chooses one. A dynamic contract specifies a sequence of transfers as a function of the publicly observable history, which is simply whether or not the project has been successful to date. To isolate the effects of adverse selection, we do not impose any limited liability constraints until Subsection 7.2. We assume that once the agent has accepted a contract, he is free to work or shirk in any period up until some termination date that is specified by the contract.<sup>14</sup> Throughout, we follow the convention that transfers are from the principal to the agent; negative values represent payments in the other direction.

Formally, a *contract* is given by  $\mathbf{C} = (T, W_0, \mathbf{b}, \mathbf{l})$ , where  $T \in \mathbb{N} \equiv \{0, 1, \dots\}$  is the *termination date* of the contract,  $W_0 \in \mathbb{R}$  is an up-front transfer (or wage) at period zero,  $\mathbf{b} = (b_1, \dots, b_T)$  specifies a transfer  $b_t \in \mathbb{R}$  made at period  $t$  conditional on the project being successful in period  $t$ , and analogously  $\mathbf{l} = (l_1, \dots, l_T)$  specifies a transfer  $l_t \in \mathbb{R}$  made at period  $t$  conditional on the project not being successful in period  $t$  (nor in any prior period).<sup>15, 16</sup> We refer to any  $b_t$  as a *bonus* and any  $l_t$  as a *penalty*. Note that  $b_t$  is not constrained to be positive nor must  $l_t$  be negative; however, these cases will be focal and hence our choice of terminology. Without loss of generality, we assume that if  $T > 0$  then  $T = \max\{t : \text{either } b_t \neq 0 \text{ or } l_t \neq 0\}$ . The agent’s actions are denoted by  $\mathbf{a} = (a_1, \dots, a_T)$ , where  $a_t = 1$  if the agent works in period  $t$  and  $a_t = 0$  if the agent shirks.

**Payoffs.** The principal’s expected discounted payoff at time zero from a contract  $\mathbf{C} = (T, W_0, \mathbf{b}, \mathbf{l})$ , an agent of type  $\theta$ , and a sequence of the agent’s actions  $\mathbf{a}$  is denoted  $\Pi_0^\theta(\mathbf{C}, \mathbf{a})$ , which can be computed as:

$$\Pi_0^\theta(\mathbf{C}, \mathbf{a}) := -W_0 - (1 - \beta_0) \sum_{t=1}^T \delta^t l_t + \beta_0 \sum_{t=1}^T \delta^t \left[ \prod_{s < t} (1 - a_s \lambda^\theta) \right] \left[ a_t \lambda^\theta (1 - b_t) - (1 - a_t \lambda^\theta) l_t \right]. \tag{2.1}$$

Formula (2.1) is understood as follows.  $W_0$  is the up-front transfer made from the principal to the agent. With probability  $1 - \beta_0$  the state is bad, in which case the project never succeeds and hence the entire sequence of penalties  $\mathbf{l}$  is transferred. Conditional on the state being good (which occurs with probability  $\beta_0$ ), the probability of project success depends on both the agent’s effort

13. Subsection 7.1 establishes that our results apply without change if success is privately observed by the agent but can be verifiably disclosed.

14. There is no loss of generality here. If the principal has the ability to block the agent from choosing whether to work in some period—“lock him out of the laboratory”, so to speak—this can just as well be achieved by instead stipulating that project success in that period would trigger a large payment to the principal.

15. We thus restrict attention to deterministic contracts. Throughout, symbols in bold typeface denote vectors.  $W_0$  and  $T$  are redundant because  $W_0$  can be effectively induced by suitable modifications to  $b_1$  and  $l_1$ , while  $T$  can be effectively induced by setting  $b_t = l_t = 0$  for all  $t > T$ . However, it is expositively convenient to include these components explicitly in defining a contract. Furthermore, there is no loss in assuming that  $T \in \mathbb{N}$ ; as we show, it is always optimal for the principal to stop experimentation at a finite time, so she cannot benefit from setting  $T = \infty$ .

16. As the principal and agent share a common discount factor, what matters is only the mapping from outcomes to transfers, not the dates at which transfers are made. Our convention facilitates our exposition.

choices and his ability;  $\prod_{s<t} (1-a_s\lambda^\theta)$  is the probability that a success does not obtain between period 1 and  $t-1$  conditional on the good state. If the project were to succeed at time  $t$ , then the principal would earn a payoff of 1 in that period, and the transfers would be the sequence of penalties  $(l_1, \dots, l_{t-1})$  followed by the bonus  $b_t$ .

Through analogous reasoning, bearing in mind that the agent does not directly value project success but incurs the cost of effort, the agent's expected discounted payoff at time zero given his type  $\theta$ , contract  $\mathbf{C}$ , and action profile  $\mathbf{a}$  is

$$U_0^\theta(\mathbf{C}, \mathbf{a}) := W_0 + (1 - \beta_0) \sum_{t=1}^T \delta^t (l_t - a_t c) + \beta_0 \sum_{t=1}^T \delta^t \left[ \prod_{s<t} (1 - a_s \lambda^\theta) \right] \left[ a_t (\lambda^\theta b_t - c) + (1 - a_t \lambda^\theta) l_t \right]. \quad (2.2)$$

If a contract is not accepted, both parties' payoffs are normalized to zero.

**Bonus and penalty contracts.** Our analysis will make use of two simple classes of contracts. A *bonus contract* is one where aside from any initial transfer there is at most only one other transfer, which occurs when the agent obtains a success. Formally, a bonus contract is  $\mathbf{C} = (T, W_0, \mathbf{b}, \mathbf{I})$  such that  $l_t = 0$  for all  $t \in \{1, \dots, T\}$ . A bonus contract is a *constant-bonus contract* if, in addition, there is some constant  $b$  such that  $b_t = b$  for all  $t \in \{1, \dots, T\}$ . When the context is clear, we denote a bonus contract as just  $\mathbf{C} = (T, W_0, \mathbf{b})$  and a constant-bonus contract as  $\mathbf{C} = (T, W_0, b)$ . By contrast, a *penalty contract* is one where the agent receives no payments for success and instead is penalized for failure. Formally, a penalty contract is  $\mathbf{C} = (T, W_0, \mathbf{b}, \mathbf{I})$  such that  $b_t = 0$  for all  $t \in \{1, \dots, T\}$ . A penalty contract is a *onetime-penalty contract* if, in addition,  $l_t = 0$  for all  $t \in \{1, \dots, T-1\}$ . That is, while in a general penalty contract the agent may be penalized for each period in which he fails to obtain a success, in a onetime-penalty contract the agent is penalized only if a success does not obtain by the termination date  $T$ . We denote a penalty contract as just  $\mathbf{C} = (T, W_0, \mathbf{I})$  and a onetime-penalty contract as  $\mathbf{C} = (T, W_0, l_T)$ .

Although each of these two classes of contracts will be useful for different reasons, there is an isomorphism between them; furthermore, either class is "large enough" in a suitable sense. More precisely, say that two contracts,  $\mathbf{C} = (T, W_0, \mathbf{b}, \mathbf{I})$  and  $\widehat{\mathbf{C}} = (T, \widehat{W}_0, \widehat{\mathbf{b}}, \widehat{\mathbf{I}})$ , are *equivalent* if for all  $\theta \in \{L, H\}$  and  $\mathbf{a} = (a_1, \dots, a_T)$ :  $U_0^\theta(\mathbf{C}, \mathbf{a}) = U_0^\theta(\widehat{\mathbf{C}}, \mathbf{a})$  and  $\Pi_0^\theta(\mathbf{C}, \mathbf{a}) = \Pi_0^\theta(\widehat{\mathbf{C}}, \mathbf{a})$ .

**Proposition 1.** *For any contract  $\mathbf{C} = (T, W_0, \mathbf{b}, \mathbf{I})$  there exist both an equivalent penalty contract  $\widehat{\mathbf{C}} = (T, \widehat{W}_0, \widehat{\mathbf{I}})$  and an equivalent bonus contract  $\widetilde{\mathbf{C}} = (T, \widetilde{W}_0, \widetilde{\mathbf{b}})$ .*

*Proof* See the Supplementary Appendix.  $\parallel$

Proposition 1 implies that it is without loss to focus either on bonus contracts or on penalty contracts. The proof is constructive: given an arbitrary contract, it explicitly derives equivalent penalty and bonus contracts. The intuition is that all that matters in any contract is the induced vector of discounted transfers for success occurring in each possible period (and never), and these transfers can be induced with bonuses or penalties.<sup>17</sup> The proof also shows that when  $\delta = 1$ , onetime-penalty contracts are equivalent to constant-bonus contracts.

17. For example, in a two-period contract  $\mathbf{C} = (2, W_0, \mathbf{b}, \mathbf{I})$ , the agent's discounted transfer is  $W_0 + \delta b_1$  if he succeeds in period one,  $W_0 + \delta l_1 + \delta^2 b_2$  if he succeeds in period two, and  $W_0 + \delta l_1 + \delta^2 l_2$  if he does not succeed in either period. The same transfers are induced by a penalty contract  $\widehat{\mathbf{C}} = (2, \widehat{W}_0, \widehat{\mathbf{I}})$  with  $\widehat{W}_0 = W_0 + \delta b_1$ ,  $\widehat{l}_1 = l_1 - b_1 + \delta b_2$ , and  $\widehat{l}_2 = l_2 - b_2$ , and by a bonus contract  $\widetilde{\mathbf{C}} = (2, \widetilde{W}_0, \widetilde{\mathbf{b}})$  with  $\widetilde{W}_0 = W_0 + \delta l_1 + \delta^2 l_2$ ,  $\widetilde{b}_1 = b_1 - l_1 - \delta l_2$ , and  $\widetilde{b}_2 = b_2 - l_2$ .

3. BENCHMARKS

3.1. *The first best*

Consider the first-best solution, *i.e.* when the agent’s type  $\theta$  is commonly known and his effort in each period is publicly observable and contractible. Since beliefs about the state being good decline so long as effort has been exerted but success not obtained, the first-best solution is characterized by a stopping rule such that an agent of ability  $\theta$  keeps exerting effort so long as success has not obtained up until some period  $t^\theta$ , whereafter effort is no longer exerted.<sup>18</sup> Let  $\beta_t^\theta$  be a generic belief on the state being good at the beginning of period  $t$  (which will depend on the history of effort), and  $\bar{\beta}_t^\theta$  be this belief when the agent has exerted effort in all periods  $1, \dots, t - 1$ . The first-best stopping time  $t^\theta$  is given by

$$t^\theta = \max_{t \geq 0} \left\{ t : \bar{\beta}_t^\theta \lambda^\theta \geq c \right\}, \tag{3.1}$$

where, for each  $\theta$ ,  $\bar{\beta}_0^\theta := \beta_0$ , and for  $t \geq 1$ , Bayes’ rule yields

$$\bar{\beta}_t^\theta = \frac{\beta_0 (1 - \lambda^\theta)^{t-1}}{\beta_0 (1 - \lambda^\theta)^{t-1} + (1 - \beta_0)}. \tag{3.2}$$

Note that (3.1) is only well-defined when  $c \leq \beta_0 \lambda^\theta$ ; if  $c > \beta_0 \lambda^\theta$ , it would be efficient to not experiment at all, *i.e.* stop at  $t^\theta = 0$ . To focus on the most interesting cases, we assume:

**Assumption 1.** *Experimentation is efficient for both types: for  $\theta \in \{L, H\}$ ,  $\beta_0 \lambda^\theta > c$ .*

If parameter values are such that  $\bar{\beta}_{t^\theta}^\theta \lambda^\theta = c$ ,<sup>19</sup> equations (3.1) and (3.2) can be combined to derive the following closed-form solution for the first-best stopping time for type  $\theta$ :

$$t^\theta = 1 + \frac{\log\left(\frac{c}{\lambda^\theta - c} \frac{1 - \beta_0}{\beta_0}\right)}{\log(1 - \lambda^\theta)}. \tag{3.3}$$

Equation (3.3) yields intuitive monotonicity of the first-best stopping time as a function of the prior that the project is good,  $\beta_0$ , and the cost of effort,  $c$ .<sup>20</sup> But it also implies a fundamental non-monotonicity as a function of the agent’s ability,  $\lambda^\theta$ , as shown in Figure 1. (For simplicity, the figure ignores integer constraints on  $t^\theta$ .) This stems from the interaction of two countervailing forces. On the one hand, for any given belief about the state, the expected marginal benefit of effort is higher when the agent’s ability is higher; on the other hand, the higher is the agent’s ability, the more informative is a lack of success in a period in which he works. Hence, at any time  $t > 1$ , a higher-ability agent is more pessimistic about the state (given that effort has been exerted in all prior periods), which has the effect of decreasing the expected marginal benefit of

18. More precisely, the first best can always be achieved using a stopping rule for each type; when and only when  $\delta = 1$ , there are other rules that also achieve the first best. Without loss, we focus on stopping rules.

19. We do not assume this condition in our analysis, but it is convenient for the current discussion.

20. One may also notice that the discount factor,  $\delta$ , does not enter (3.3). In other words, unlike the traditional focus of experimentation models, there is no tradeoff here between “exploration” and “exploitation”, as the first-best strategy is invariant to patience. Our model and subsequent analysis can be generalized to incorporate this tradeoff, but the additional burden does not yield commensurate insight.



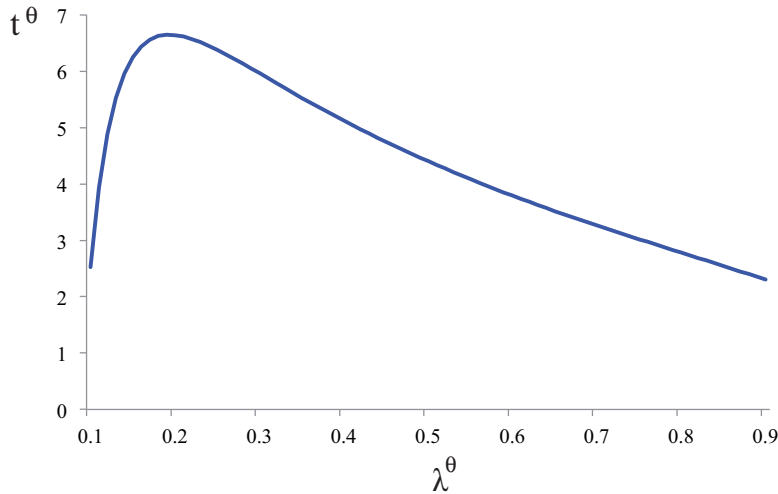


FIGURE 1  
The first-best stopping time.

effort. Altogether, this makes the first-best stopping time non-monotonic in ability; both  $t^H > t^L$  and  $t^H < t^L$  are robust possibilities that arise for different parameters. As we will see, this has substantial implications.

The first-best expected discounted surplus at time zero from type  $\theta$  is

$$\sum_{t=1}^{t^\theta} \delta^t \left[ \beta_0 (1 - \lambda^\theta)^{t-1} (\lambda^\theta - c) - (1 - \beta_0)c \right].$$

### 3.2. No adverse selection or no moral hazard

Our model has two sources of asymmetric information: adverse selection and moral hazard. To see that their interaction is essential, it is useful to understand what would happen in the absence of either one.

Consider first the case without adverse selection, *i.e.* assume the agent's ability is observable but there is moral hazard. The principal can then use a constant-bonus contract to effectively sell the project to the agent at a price that extracts all the (ex-ante) surplus. Specifically, suppose the principal offers the agent of type  $\theta$  a constant-bonus contract  $\mathbf{C}^\theta = (t^\theta, W_0^\theta, 1)$ , where  $W_0^\theta$  is chosen so that conditional on the agent exerting effort in each period up to the first-best termination date (as long as success has not been obtained), the agent's participation constraint at time zero binds:

$$U_0^\theta(\mathbf{C}^\theta, \mathbf{1}) = \sum_{t=1}^{t^\theta} \delta^t \left[ \beta_0 (1 - \lambda^\theta)^{t-1} (\lambda^\theta - c) - (1 - \beta_0)c \right] + W_0^\theta = 0,$$

where the notation  $\mathbf{1}$  denotes the action profile of working in every period of the contract. Plainly, this contract makes the agent fully internalize the social value of success and hence achieves the first-best level of experimentation, while the principal keeps all the surplus.

Consider next the case with adverse selection but no moral hazard: the agent's effort in any period still costs him  $c > 0$  but is observable and contractible. The principal can then implement

the first best and extract all the surplus by using simple contracts that pay the agent for effort rather than outcomes. Specifically, the principal can offer the agent a choice between two contracts that involve no bonuses or penalties, with each paying the agent  $c$  for every period that he works. The termination date is  $t^L$  in the contract intended for the low type and  $t^H$  in the contract intended for the high type. Plainly, the agent's payoff is zero regardless of his type and which contract and effort profile he chooses. Hence, the agent is willing to choose the contract intended for his type and work until either a success is obtained or the termination date is reached.<sup>21</sup>

To summarize:

**Theorem 1.** *If there is either no moral hazard or no adverse selection, the principal optimally implements the first best and extracts all the surplus.*

A proof is omitted in light of the simple arguments preceding the theorem. Theorem 1 also holds when there are many types; that both kinds of information asymmetries are essential to generate distortions is general in our experimentation environment.<sup>22</sup>

#### 4. SECOND-BEST (IN)EFFICIENCY

We now turn to the setting with both moral hazard and adverse selection. In this section, we formalize the principal's problem and deduce the nature of second-best inefficiency. We provide explicit characterizations of optimal contracts in Section 5 and Section 6.

Without loss, we assume that the principal specifies a desired effort profile along with a contract. An optimal menu of contracts maximizes the principal's ex-ante expected payoff subject to incentive compatibility constraints for effort ( $IC_a^\theta$  below), participation constraints ( $IR^\theta$  below), and self-selection constraints for the agent's choice of contract ( $IC^{\theta\theta'}$  below). Denote

$$\alpha^\theta(\mathbf{C}) := \operatorname{argmax}_{\mathbf{a}} U_0^\theta(\mathbf{C}, \mathbf{a})$$

as the set of optimal action plans for the agent of type  $\theta$  under contract  $\mathbf{C}$ . With a slight abuse of notation, we will write  $U_0^\theta(\mathbf{C}, \alpha^\theta(\mathbf{C}))$  for the type- $\theta$  agent's utility at time zero from any contract  $\mathbf{C}$ . The principal's program is:

$$\max_{(\mathbf{C}^H, \mathbf{C}^L, \mathbf{a}^H, \mathbf{a}^L)} \mu_0 \Pi_0^H(\mathbf{C}^H, \mathbf{a}^H) + (1 - \mu_0) \Pi_0^L(\mathbf{C}^L, \mathbf{a}^L)$$

subject to, for all  $\theta, \theta' \in \{L, H\}$ ,

$$\mathbf{a}^\theta \in \alpha^\theta(\mathbf{C}^\theta), \tag{IC_a^\theta}$$

$$U_0^\theta(\mathbf{C}^\theta, \mathbf{a}^\theta) \geq 0, \tag{IR^\theta}$$

$$U_0^\theta(\mathbf{C}^\theta, \mathbf{a}^\theta) \geq U_0^\theta(\mathbf{C}^{\theta'}, \alpha^\theta(\mathbf{C}^{\theta'})). \tag{IC^{\theta\theta'}}$$

Adverse selection is reflected in the self-selection constraints ( $IC^{\theta\theta'}$ ), as is familiar. Moral hazard is reflected directly in the constraints ( $IC_a^\theta$ ) and also indirectly in the constraints ( $IC^{\theta\theta'}$ )

21. The same idea underlies Gomes *et al.*'s (2015) Lemma 2. While this mechanism makes the agent indifferent over the contracts, there are more sophisticated optimal mechanisms, detailed in earlier versions of our paper, that satisfy the agent's self-selection constraint strictly.

22. We note that learning is also important in generating distortions: in the absence of learning (*i.e.* if the project were known to be good,  $\beta_0 = 1$ ), the principal may again implement the first best. For expositional purposes, we defer this discussion to Subsection 7.3.

via the term  $\alpha^\theta(\mathbf{C}^{\theta'})$ . To get a sense of how these matter, consider the agent's incentive to work in some period  $t$ . This is shaped not only by the transfers that are directly tied to success/failure in period  $t$  ( $b_t$  and  $l_t$ ) but also by the transfers tied to subsequent outcomes, through their effect on continuation values. In particular, *ceteris paribus*, raising the continuation value (say, by increasing either  $b_{t+1}$  or  $l_{t+1}$ ) makes reaching period  $t+1$  more attractive and hence reduces the incentive to work in period  $t$ : this is a dynamic agency effect.<sup>23</sup> Note moreover that the continuation value at any point in a contract depends on the agent's type and his effort profile; hence it is not sufficient to consider a single continuation value at each period. Furthermore, besides having an effect on continuation values, the agent's type also affects current incentives for effort because the expected marginal benefit of effort in any period differs for the two types. Altogether, the optimal plan of action will generally be different for the two types of the agent, *i.e.* for an arbitrary contract  $\mathbf{C}$ , we may have  $\alpha^H(\mathbf{C}) \cap \alpha^L(\mathbf{C}) = \emptyset$ .<sup>24</sup>

Our result on second-best (in)efficiency is as follows:

**Theorem 2.** *In any optimal menu of contracts, each type  $\theta \in \{L, H\}$  is induced to work for some number of periods,  $\bar{t}^\theta$ . Relative to the first-best stopping times,  $t^H$  and  $t^L$ , the second best has  $\bar{t}^H = t^H$  and  $\bar{t}^L \leq t^L$ .*

*Proof* See Appendix A. ||

Theorem 2 says that relative to the first best, there is no distortion in the amount of experimentation by the high-ability agent, whereas the low-ability agent may be induced to under-experiment. It is interesting that this is a familiar “no distortion (only) at the top” result from static models of adverse selection, even though the inefficiency arises here from the conjunction of adverse selection and dynamic moral hazard (cf. Theorem 1). Moral hazard generates an “information rent” for the high type but not for the low type. As will be elaborated subsequently, reducing the low type's amount of experimentation allows the principal to reduce the high type's information rent. The optimal  $\bar{t}^L$  trades off this information rent with the low type's efficiency. For typical parameters, it will be the case that  $\bar{t}^L \in \{1, \dots, t^L - 1\}$ , so that the low type engages in some experimentation but not as much as socially efficient; however, it is possible that the low type is induced to not experiment at all ( $\bar{t}^L = 0$ ) or to experiment for the first-best amount of time ( $\bar{t}^L = t^L$ ). The former possibility arises for reasons akin to exclusion in the standard model (*e.g.* the prior,  $\mu_0$ , on the high type is sufficiently high); the latter possibility is because time is discrete. Indeed, if the length of each time interval shrinks and one takes a suitable continuous-time limit, then there will be some distortion, *i.e.*  $\bar{t}^L < t^L$ .

The proof of Theorem 2 does not rely on characterizing second-best contracts.<sup>25</sup> We establish  $\bar{t}^H = t^H$  by proving that the low type's self-selection constraint can always be satisfied without creating any distortions. The idea is that the principal can exploit the two types' differing probabilities of success by making the high type's contract “risky enough” to deter the low

23. Mason and Välimäki (2011), Bhaskar (2012, 2014), Hörner and Samuelson (2013), and Kwon (2013) also highlight dynamic agency effects, but in settings without adverse selection.

24. Related issues arise in static models that allow for both adverse selection and moral hazard; see for example the discussion in Laffont and Martimort (2001, Chapter 7).

25. Note that when  $\delta < 1$ , efficiency requires each type to use a “stopping strategy” (*i.e.* work for a consecutive sequence of periods beginning with period one). The proof technique for Theorem 2 does not allow us to establish that the low type uses a stopping strategy in the second-best solution; however, it shows that one can take the high type to be doing so. That the low type can also be taken to use a stopping strategy (with the second-best stopping time) will be deduced subsequently in those cases in which we are able to characterize second-best contracts.

type from taking it, while still satisfying all other constraints.<sup>26</sup> We establish  $\bar{t}^L \leq t^L$  by showing that any contract for the low type inducing  $\bar{t}^L > t^L$  can be modified by “removing” the last period of experimentation in this contract and concurrently reducing the information rent for the high type. Due to the lack of structure governing the high type’s behaviour upon deviating to the low type’s contract, we prove the information-rent reduction no matter what action plan the high type would choose upon taking the low type’s contract. It follows that inducing over-experimentation by the low type cannot be optimal: not only would that reduce social surplus but it would also increase the high type’s information rent.

While Theorem 2 has implications for the extent of experimentation and innovation in different economic environments, we postpone such discussion to Subsection 5.3, after describing optimal contracts and their comparative statics.

### 5. OPTIMAL CONTRACTS WHEN $t^H > t^L$

We characterize optimal contracts by first studying the case in which the first-best stopping times are ordered  $t^H > t^L$ , *i.e.* when the speed-of-learning effect that pushes the first-best stopping time down for a higher-ability agent does not dominate the productivity effect that pushes in the other direction. Any of the following conditions on the primitives is sufficient for  $t^H > t^L$ , given a set of other parameters: (1)  $\beta_0$  is small enough, (2)  $\lambda^L$  and  $\lambda^H$  are small enough, or (3)  $c$  is large enough. We maintain the assumption that  $t^H > t^L$  implicitly throughout this section.

#### 5.1. The solution

A class of solutions to the principal’s program described in Section 4 when  $t^H > t^L$  is as follows:

**Theorem 3.** *Assume  $t^H > t^L$ . There is an optimal menu in which the principal separates the two types using penalty contracts. In particular, the optimum can be implemented using a onetime-penalty contract for type H,  $C^H = (t^H, W_0^H, l_{t^H}^H)$  with  $l_{t^H}^H < 0 < W_0^H$ , and a penalty contract for type L,  $C^L = (t^L, W_0^L, l^L)$ , such that:*

(1) For all  $t \in \{1, \dots, \bar{t}^L\}$ ,

$$l_t^L = \begin{cases} -(1-\delta) \frac{c}{\beta_t \lambda^L} & \text{if } t < \bar{t}^L, \\ -\frac{c}{\beta_{\bar{t}^L} \lambda^L} & \text{if } t = \bar{t}^L; \end{cases} \tag{5.1}$$

(2)  $W_0^L > 0$  is such that the participation constraint,  $(IR^L)$ , binds;

26. Specifically, given an optimal contract for the high type, the principal can increase the magnitude of the penalties while adjusting the time-zero transfer so that the high type’s expected payoff and effort profile do not change. Making the penalties severe enough (*i.e.* negative enough) then ensures that the low type’s payoff from taking the high type’s contract is negative and hence  $(IC^{LH})$  is satisfied at no cost. Crucially, an analogous construction would not work for the high type’s self-selection constraint: the high type’s payoff under the low type’s contract cannot be lower than the low type’s, as the high type can always generate the same distribution of project success as the low type by suitably mixing over effort. From the point of view of correlated-information mechanism design (Cremer and McLean, 1985, 1988; Riordan and Sappington, 1988), the issue is that because of moral hazard, the signal correlated with the agent’s type is not independent of the agent’s report. In a different setting, Obara (2008) has also noted this effect of hidden actions. While Obara (2008) shows that in his setting approximate full surplus extraction may be achieved by having agents randomize over their actions, this is not generally possible here because the feasible set of distributions of project success for the high type is a superset of that of the low type.

- (3) Type  $H$  gets an information rent:  $U_0^H(\mathbf{C}^H, \boldsymbol{\alpha}^H(\mathbf{C}^H)) > 0$ ;  
 (4)  $\mathbf{1} \in \boldsymbol{\alpha}^H(\mathbf{C}^H)$ ;  $\mathbf{1} \in \boldsymbol{\alpha}^L(\mathbf{C}^L)$ ; and  $\mathbf{1} = \boldsymbol{\alpha}^H(\mathbf{C}^L)$ .

Generically, the above contract is the unique optimal contract for type  $L$  within the class of penalty contracts.

*Proof* See Appendix B.  $\parallel$

The optimal contract for the low type characterized by (5.1) is a penalty contract in which the magnitude of the penalty is increasing over time, with a “jump” in the contract’s final period. The jump highlights dynamic agency effects: by obtaining a success in a period  $t$ , the agent not only avoids the penalty  $l_t^L$  but also the penalty  $l_{t+1}^L$  and those after. The last period’s penalty needs to compensate for the absence of future penalties. Figure 2 depicts the low type’s contract; the comparative statics seen in the figure will be discussed subsequently. Only when there is no discounting does the low type’s contract reduce to a *onetime*-penalty contract where a penalty is paid only if the project has not succeeded by  $\bar{t}^L$ . For any discount factor, the high type’s contract characterized in Theorem 3 is a *onetime*-penalty contract in which he only pays a penalty to the principal if there is no success by the first-best stopping time  $t^H$ . On the equilibrium path, both types of the agent exert effort in every period until their respective stopping times; moreover, were the high type to take the low type’s contract (off the equilibrium path), he would also exert effort in every period of the contract. This implies that the high type gets an information rent because he would be less likely than the low type to incur any of the penalties in  $\mathbf{C}^L$ .

Although the optimal contract for the low type is (generically) unique among penalty contracts, there are a variety of optimal penalty contracts for the high type. The reason is that the low type’s optimal contract is pinned down by the need to simultaneously incentivize the low type’s effort and yet minimize the information rent obtained by the high type. This leads to a sequence of penalties for the low type, given by (5.1), that make him indifferent between working and shirking in each period of the contract, as we explain further in Subsection 5.2. On the other hand, the high type’s contract only needs to be made unattractive to the low type subject to incentivizing effort from the high type and providing the high type a utility level given by his information rent. There is latitude in how this can be done: the *onetime* penalty in the high type’s contract of Theorem 3 is chosen to be severe enough so that this contract is “too risky” for the low type to accept.

**Remark 1.** *The proof of Theorem 3 provides a simple algorithm to solve for an optimal menu of contracts. For any  $\hat{t} \in \{0, \dots, t^L\}$ , we characterize an optimal menu that solves the principal’s program subject to an additional constraint that the low type must experiment until period  $\hat{t}$ . The low type’s contract in this menu is given by (5.1) with the termination date  $\hat{t}$  rather than  $\bar{t}^L$ . An optimal (unconstrained) menu is then obtained by maximizing the principal’s objective function over  $\hat{t} \in \{0, \dots, t^L\}$ .*

The characterization in Theorem 3 yields the following comparative statics:

**Proposition 2.** *Assume  $t^H > t^L$  and consider changes in parameters that preserve this ordering. The second-best stopping time for type  $L$ ,  $\bar{t}^L$ , is weakly increasing in  $\beta_0$  and  $\lambda^L$ , weakly decreasing in  $c$  and  $\mu_0$ , and can increase or decrease in  $\lambda^H$ . The distortion in this stopping time, measured by  $t^L - \bar{t}^L$ , is weakly increasing in  $\mu_0$  and can increase or decrease in  $\beta_0$ ,  $\lambda^L$ ,  $\lambda^H$ , and  $c$ .*

*Proof* See the Supplementary Appendix.  $\parallel$

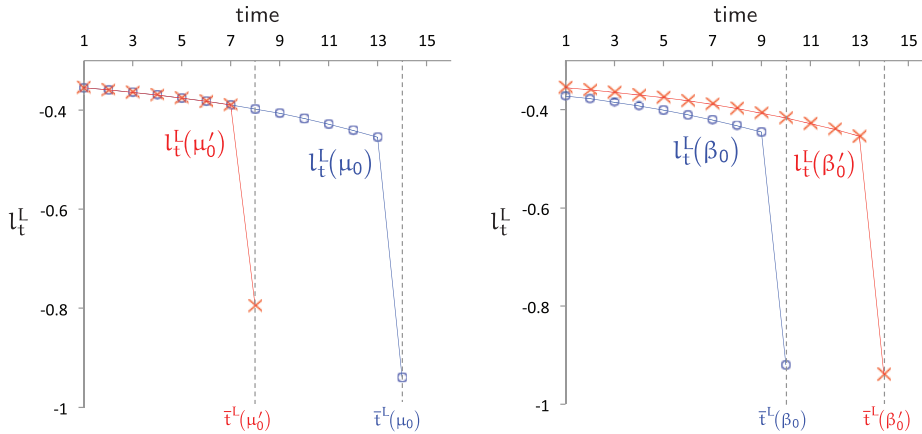


FIGURE 2

The optimal penalty contract for type  $L$  under different values of  $\mu_0$  and  $\beta_0$ . Both graphs have  $\delta=0.5$ ,  $\lambda^L=0.1$ ,  $\lambda^H=0.12$ , and  $c=0.06$ . The left graph has  $\beta_0=0.89$ ,  $\mu_0=0.3$ , and  $\mu'_0=0.6$ ; the right graph has  $\beta_0=0.85$ ,  $\beta'_0=0.89$ , and  $\mu_0=0.3$ . The first best entails  $t^L=15$  on the left graph, and  $t^L=12$  (for  $\beta_0$ ) and  $t^L=15$  (for  $\beta'_0$ ) on the right graph.

Figure 2 illustrates some of the conclusions of Proposition 2. The comparative static of  $\bar{t}^L$  in  $\mu_0$  is intuitive: the higher the ex-ante probability of the high type, the more the principal benefits from reducing the high type’s information rent and hence the more she shortens the low type’s experimentation. Matters are more subtle for other parameters. Consider, for example, an increase in  $\beta_0$ . On the one hand, this increases the social surplus from experimentation, which suggests that  $\bar{t}^L$  should increase. But there are two other effects: holding fixed  $\bar{t}^L$ , penalties of lower magnitude can be used to incentivize effort from the low type because the project is more likely to succeed (cf. equation (5.1)), which has an effect of decreasing the information rent for the high type; yet, a higher  $\beta_0$  also has a direct effect of increasing the information rent because the differing probability of success for the two types is only relevant when the project is good. Nevertheless, Proposition 2 establishes that it is optimal to (weakly) increase  $\bar{t}^L$  when  $\beta_0$  increases.

Since the high type’s information rent is increasing in  $\lambda^H$ , one may expect the principal to reduce the low type’s experimentation when  $\lambda^H$  increases. However, a higher  $\lambda^H$  means that the high type is likely to succeed earlier when deviating to the low type’s contract. For this reason, an increase in  $\lambda^H$  can reduce the incremental information-rent cost of extending the low type’s contract, to the extent that the gain in efficiency from the low type makes it optimal to increase  $\bar{t}^L$ .

Turning to the magnitude of distortion,  $t^L - \bar{t}^L$ : since the first-best stopping time  $t^L$  does not depend on the probability of a high type,  $\mu_0$ , while  $\bar{t}^L$  is decreasing in this parameter, it is immediate that the distortion is increasing in  $\mu_0$ . The time  $t^L$  is also independent of the high type’s ability,  $\lambda^H$ ; thus, since  $\bar{t}^L$  may increase or decrease in  $\lambda^H$ , the same is true for  $t^L - \bar{t}^L$ . Finally, with respect to  $\beta_0$ ,  $\lambda^L$ , and  $c$ , the distortion’s ambiguous comparative statics stem from the fact that  $t^L$  and  $\bar{t}^L$  move in the same direction when these parameters change. For example, increasing  $\beta_0$  can reduce  $t^L - \bar{t}^L$  when  $\mu_0$  is low but increase  $t^L - \bar{t}^L$  when  $\mu_0$  is high; the reason is that a larger ex-ante probability of the high type makes increasing  $\bar{t}^L$  more costly in terms of information rent.

Theorem 3 utilizes penalty contracts in which the agent is required to pay the principal when he fails to obtain a success. While these contracts prove analytically convenient (as explained in Subsection 5.2), a weakness is that they do not satisfy *interim participation constraints*: in the

implementation of Theorem 3, the agent of either type  $\theta$  would “walk away” from his contract in any period  $t \in \{1, \dots, \bar{t}^\theta\}$  if he could. The following result provides a remedy:

**Theorem 4.** *Assume  $t^H > t^L$ . The second best can also be implemented using a menu of bonus contracts. Specifically, the principal offers type  $L$  the bonus contract  $\mathbf{C}^L = (\bar{t}^L, W_0^L, \mathbf{b}^L)$  wherein for any  $t \in \{1, \dots, \bar{t}^L\}$ ,*

$$b_t^L = \sum_{s=t}^{\bar{t}^L} \delta^{s-t} (-l_s^L), \quad (5.2)$$

where  $\mathbf{l}^L$  is the penalty sequence in the optimal penalty contract given in Theorem 3, and  $W_0^L$  is chosen to make the participation constraint,  $(\mathbf{IR}^L)$ , bind. For type  $H$ , the principal can use a constant-bonus contract  $\mathbf{C}^H = (t^H, W_0^H, \mathbf{b}^H)$  with a suitably chosen  $W_0^H$  and  $b^H > 0$ .

Generically, the above contract is the unique optimal contract for type  $L$  within the class of bonus contracts. This implementation satisfies interim participation constraints in each period for each type, i.e. each type  $\theta$ 's continuation utility at the beginning of any period  $t \in \{1, \dots, \bar{t}^\theta\}$  in  $\mathbf{C}^\theta$  is non-negative.

A proof is omitted because the proof of Proposition 1 can be used to verify that each bonus contract in Theorem 4 is equivalent to the corresponding penalty contract in Theorem 3, and hence the optimality of those penalty contracts implies the optimality of these bonus contracts. Using (5.1), it is readily verified that in the bonus sequence (5.2),

$$b_{\bar{t}^L}^L = \frac{c}{\beta_{\bar{t}^L}^L \lambda^L} \text{ and } b_t^L = \frac{(1-\delta)c}{\beta_t^L \lambda^L} + \delta b_{t+1}^L \text{ for any } t \in \{1, \dots, \bar{t}^L - 1\}, \quad (5.3)$$

and hence the reward for success increases over time. When  $\delta = 1$ , the low type's bonus contract is a constant-bonus contract, analogous to the penalty contract in Theorem 3 being a onetime-penalty contract.

An interpretation of the bonus contracts in Theorem 4 is that the principal initially sells the project to the agent at some price (the up-front transfer  $W_0$ ) with a commitment to buy back the output generated by a success at time-dated future prices (the bonuses  $\mathbf{b}$ ).

## 5.2. Sketch of the proof

We now sketch in some detail how we prove Theorem 3. The arguments reveal how the interaction of adverse selection, dynamic moral hazard, and private learning jointly shape optimal contracts. This subsection also serves as a guide to follow the formal proof in Appendix B.

While we have defined a contract as  $\mathbf{C} = (T, W_0, \mathbf{b}, I)$ , it will be useful in this subsection alone (so as to parallel the formal proof) to consider a larger space of contracts, where a contract is given by  $\mathbf{C} = (\Gamma, W_0, \mathbf{b}, I)$ . The first element here is a set of periods,  $\Gamma \subseteq \mathbb{N} \setminus \{0\}$ , at which the agent is not “locked out,” i.e. at which he is allowed to choose whether to work or shirk. As discussed in fn. 14, this additional instrument does not yield the principal any benefit, but it will be notationally convenient in the proof. The termination date of the contract is now 0 if  $\Gamma = \emptyset$  and otherwise  $\max \Gamma$ . We say that a contract is *connected* if  $\Gamma = \{1, \dots, T\}$  for some  $T$ ; in this case we refer to  $T$  as the *length* of the contract, and  $T$  is also the termination date. The agent's actions are denoted by  $\mathbf{a} = (a_t)_{t \in \Gamma}$ .

As justified by Proposition 1, we solve the principal's problem (stated at the outset of Section 4) by restricting attention to menus of penalty contracts: for each  $\theta \in \{L, H\}$ ,  $\mathbf{C}^\theta = (\Gamma^\theta, W_0^\theta, \mathbf{l}^\theta)$ .

Penalty contracts are analytically convenient to deal with the combination of adverse selection and dynamic moral hazard for reasons explained in Step 4 below.

**Step 1:** We simplify the principal's program by (1) focussing on contracts for type  $L$  that induce him to work in every non-lockout period, *i.e.* on contracts in the set  $\{\mathbf{C}^L : \mathbf{1} \in \alpha^L(\mathbf{C}^L)\}$ ; and (2) ignoring the constraints  $(\text{IR}^H)$  and  $(\text{IC}^{LH})$ . It is established in the proof of Theorem 2 that a solution to this simplified program also solves the original program.<sup>27</sup> Call this program [P1].

It is not obvious *a priori* what action plan the high type may use when taking the low type's contract. Accordingly, we tackle a relaxed program, [RP1], that replaces  $(\text{IC}^{HL})$  in program [P1] by a relaxed version, called (Weak- $\text{IC}^{HL}$ ), that only requires type  $H$  to prefer taking his contract and following an optimal action plan over taking type  $L$ 's contract and working in every period. Formally,  $(\text{IC}^{HL})$  requires  $U_0^H(\mathbf{C}^H, \alpha^H(\mathbf{C}^H)) \geq U_0^H(\mathbf{C}^L, \alpha^H(\mathbf{C}^L))$ , whereas (Weak- $\text{IC}^{HL}$ ) requires only  $U_0^H(\mathbf{C}^H, \alpha^H(\mathbf{C}^H)) \geq U_0^H(\mathbf{C}^L, \mathbf{1})$ . We emphasize that this restriction on type  $H$ 's action plan under type  $L$ 's contract is not without loss for an arbitrary contract  $\mathbf{C}^L$ ; *i.e.* given an arbitrary  $\mathbf{C}^L$  with  $\mathbf{1} \in \alpha^L(\mathbf{C}^L)$ , it need not be the case that  $\mathbf{1} \in \alpha^H(\mathbf{C}^L)$ —it is in this sense that there is no “single-crossing property” in general. The reason is that because of their differing probabilities of success from working in future periods (conditional on the good state), the two types trade off current and future penalties differently when considering exerting effort in the current period. In particular, the desire to avoid future penalties provides more of an incentive for the low type to work in the current period than the high type.<sup>28</sup>

Relaxing  $(\text{IC}^{HL})$  to (Weak- $\text{IC}^{HL}$ ) is motivated by a conjecture that even though the high type may choose to work less than the low type in an arbitrary contract, this will not be the case in an *optimal* contract for the low type. This relaxation is a critical step in making the program tractable because it severs the knot in the fixed point problem of optimizing over the low type's contract while not knowing what action plan the high type would follow should he take this contract. The relaxation works because of the efficiency ordering  $t^H > t^L$ , as elaborated subsequently.

In the relaxed program [RP1], it is straightforward to show that (Weak- $\text{IC}^{HL}$ ) and  $(\text{IR}^L)$  must bind at an optimum: otherwise, time-zero transfers in one of the two contracts can be profitably lowered without violating any of the constraints. Consequently, one can substitute from the binding version of these constraints to rewrite the objective function as the sum of total surplus less an information rent for the high type, as in the standard approach. We are left with a relaxed program, [RP2], which maximizes this objective function and whose only constraints are the direct moral hazard constraints  $(\text{IC}_a^H)$  and  $(\text{IC}_a^L)$ , where type  $L$  must work in all periods. This program is tractable because it can be solved by separately optimizing over each type's penalty

27. The idea for (1) is as follows: fix any contract,  $\mathbf{C}^L$ , in which there is some period,  $t \in \Gamma^L$ , such that it would be suboptimal for type  $L$  to work in period  $t$ . Since type  $L$  will not succeed in period  $t$ , one can modify  $\mathbf{C}^L$  to create a new contract,  $\widehat{\mathbf{C}}^L$ , in which  $t \notin \widehat{\Gamma}^L$ , and  $t_t^L$  is “shifted up” by one period with an adjustment for discounting. This ensures that the incentives for type  $L$  in all other periods remain unchanged, and critically, that no matter what behavior would have been optimal for type  $H$  under contract  $\mathbf{C}^L$ , the new contract is less attractive to type  $H$ .

As for (2), we show that type  $H$  always has an optimal action plan under contract  $\mathbf{C}^L$  that yields him a higher payoff than that of type  $L$  under  $\mathbf{C}^L$ , and hence  $(\text{IR}^H)$  is implied by  $(\text{IC}^{HL})$  and  $(\text{IR}^L)$ . Finally, we show that  $(\text{IC}^{LH})$  can always be satisfied while still satisfying the other constraints in the principal's program by making the high type's contract “risky enough” to deter the low type from taking it.

28. To substantiate this point, consider any two-period penalty contract under which it is optimal for both types to work in each period. It can be verified that changing the first-period penalty by  $\varepsilon_1 > 0$  while simultaneously changing the second period penalty by  $-\varepsilon_2 < 0$  would preserve type  $\theta$ 's incentive to work in period one if and only if  $\varepsilon_1 \leq (1 - \lambda^\theta) \delta \varepsilon_2$ . Note that because  $-\varepsilon_2 < 0$ , both types will continue to work in period two independent of their action in period one. Consequently, the initial contract can always be modified in a way that preserves optimality of working in both periods for the low type, but makes it optimal for the high type to shirk in period one and work in period two.



contract. The following steps 2–5 derive an optimal contract for type  $L$  in program [RP2] that has useful properties.

**Step 2:** We show that there is an optimal penalty contract for type  $L$  that is connected. A rough intuition is as follows.<sup>29</sup> Because type  $L$  is required to work in all non-lockout periods, the value of the objective function in program [RP2] can be improved by removing any lockout periods in one of two ways: either by “shifting up” the sequence of effort and penalties or by terminating the contract early (suitably adjusting for discounting in either case). Shifting up the sequence of effort and penalties eliminates inefficient delays in type  $L$ ’s experimentation, but it also increases the rent given to type  $H$ , because the penalties—which are more likely to be borne by type  $L$  than type  $H$ —are now paid earlier. Conversely, terminating the contract early reduces the rent given to type  $H$  by lowering the total penalties in the contract, but it also shortens experimentation by type  $L$ . It turns out that either of these modifications may be beneficial to the principal, but at least one of them will be if the initial contract is not connected.

**Step 3:** Given any termination date  $T^L$ , there are many penalty sequences that can be used by a connected penalty contract of length  $T^L$  to induce the low-ability agent to work in each period  $1, \dots, T^L$ . We construct the unique sequence, call it  $\bar{l}(T^L)$ , that ensures the low type’s incentive constraint for effort binds in each period of the contract, *i.e.* in any period  $t \in \{1, \dots, T^L\}$ , the low type is indifferent between working (and then choosing any optimal effort profile in subsequent periods) and shirking (and then choosing any optimal effort profile in subsequent periods), given the past history of effort. The intuition is straightforward: in the final period,  $T^L$ , there is obviously a unique such penalty as it must solve  $\bar{l}_{T^L}^L(T^L) = -c + (1 - \bar{\beta}_{T^L}^L \lambda^L) \bar{l}_{T^L}^L(T^L)$ . Iteratively working backward using a one-step deviation principle, this pins down penalties in each earlier period through the (forward-looking) incentive constraint for effort in each period. Naturally, for any  $T^L$  and  $t \in \{1, \dots, T^L\}$ ,  $\bar{l}_t^L(T^L) < 0$ , *i.e.* as suggested by the term “penalty”, the agent pays the principal each time there is a failure.

**Step 4:** We show that any connected penalty contract for type  $L$  that solves program [RP2] must use the penalty structure  $\bar{l}^L(\cdot)$  of Step 3. The idea is that any slack in the low type’s incentive constraint for effort in any period can be used to modify the contract to strictly reduce the high type’s expected payoff from taking the low type’s contract (without affecting the low type’s behaviour or expected payoff), based on the high type succeeding with higher probability in every period when taking the low type’s contract.<sup>30</sup>

Although this logic is intuitive, a formal argument must deal with the challenge that modifying a transfer in any period to reduce slack in the low type’s incentive constraint for effort in that period has feedback on incentives in every prior period—the dynamic agency problem. Our focus on penalty contracts facilitates the analysis here because penalty contracts have the property that reducing the incentive to exert effort in any period  $t$  by decreasing the severity of the penalty in period  $t$  has a *positive feedback* of also reducing the incentive for effort in earlier periods, since the continuation value of reaching period  $t$  increases. Due to this positive feedback, we are able to show that the low type’s incentive for effort in a given period of a connected penalty contract can

29. For the intuition that follows, assume that all penalties being discussed are negative transfers, *i.e.* transfers from the agent to the principal.

30. This is because the constraint (Weak-IC<sup>HL</sup>) in program [RP2] effectively constrains the high type in this way, even though, as previously noted, it may not be optimal for the high type to work in each period when taking an arbitrary contract for the low type.

be modified without affecting his incentives in any other period by solely adjusting the penalties in that period and the previous one. In particular, in an arbitrary connected penalty contract  $\mathbf{C}^L$ , if type  $L$ 's incentive constraint is slack in some period  $t$ , we can increase  $l_t^L$  and reduce  $l_{t-1}^L$  in a way that leaves type  $L$ 's incentives for effort unchanged in every period  $s \neq t$  while still being satisfied in period  $t$ . We then verify that this “local modification” strictly reduces the high type’s information rent.<sup>31</sup>

**Step 5:** In light of Steps 2–4, all optimal connected penalty contracts for type  $L$  in program [RP2] can be found by just optimizing over the length of connected penalty contracts with the penalty structure  $\bar{l}^L(\cdot)$ . By Theorem 2, the optimal length,  $\bar{t}^L$ , cannot be larger than the first-best stopping time:  $\bar{t}^L \leq t^L$ . In this step, we further establish that  $\bar{t}^L$  is generically unique, and that generically there is no optimal penalty contract for type  $L$  that is not connected.

**Step 6:** Let  $\bar{\mathbf{C}}^L$  be the contract for type  $L$  identified in Steps 2–5.<sup>32</sup> Recall that [RP1] differs from the principal’s original program [P1] in that it imposes (Weak- $\text{IC}^{HL}$ ) rather than ( $\text{IC}^{HL}$ ). In this step, we show that any solution to [RP1] using  $\bar{\mathbf{C}}^L$  satisfies ( $\text{IC}^{HL}$ ) and hence is also a solution to program [P1]. Specifically, we show that  $\alpha^H(\bar{\mathbf{C}}^L) = \mathbf{1}$ , *i.e.* if type  $H$  were to take contract  $\bar{\mathbf{C}}^L$ , it would be uniquely optimal for him to work in all periods  $1, \dots, \bar{t}^L$ . The intuition is as follows: under contract  $\bar{\mathbf{C}}^L$ , type  $H$  has a higher expected probability of success from working in any period  $t \leq \bar{t}^L$ , no matter his prior choices of effort, than does type  $L$  in period  $t$  given that type  $L$  has exerted effort in all prior periods (recall  $\mathbf{1} \in \alpha^L(\bar{\mathbf{C}}^L)$ ). The argument relies on Theorem 2 having established that  $\bar{t}^L \leq t^L$ , because  $t^H > t^L$  then implies that for any  $t \in \{1, \dots, \bar{t}^L\}$ ,  $\beta_t^H \lambda^H > \bar{\beta}_t^L \lambda^L$  for any history of effort by type  $H$  in periods  $1, \dots, t-1$ . Using this property, we verify that because  $\bar{\mathbf{C}}^L$  makes type  $L$  indifferent between working and shirking in each period up to  $\bar{t}^L$  (given that he has worked in all prior periods), type  $H$  would find it strictly optimal to work in each period up to  $\bar{t}^L$  no matter his prior history of effort.

### 5.3. Implications and applications

**Asymmetric information and success.** Our results offer predictions on the extent of experimentation and innovation. An immediate implication concerns the effects of asymmetric information. Compare a setting with either no moral hazard or no adverse selection, as in Theorem 1, with a setting where both features are present, as in Theorem 2. The theorems reveal that, other things equal, the amount of experimentation will be lower in the latter, and, consequently, the average probability of success will also be lower. Furthermore, because low-ability agents’ experimentation is typically distorted down whereas that of high-ability agents is not, we predict a larger dispersion in success rates across agents and projects when both forms of asymmetric information are present.<sup>33</sup>

31. In contrast, bonuses have a *negative feedback*: reducing the bonus in a period  $t$  increases the incentive to work in prior periods because the continuation value of reaching period  $t$  decreases. Consequently, keeping incentives for effort in earlier periods unchanged after reducing the bonus in period  $t$  would require a “global modification” of reducing the bonus in all prior periods, not just the previous period. This makes the analysis with bonus contracts less convenient.

32. The initial transfer in  $\bar{\mathbf{C}}^L$  is set to make the participation constraint for type  $L$  bind. In the non-generic cases where there are multiple optimal lengths of contract,  $\bar{\mathbf{C}}^L$  uses the largest one.

33. While this is readily evident when  $t^H > t^L$ , it is also true when  $t^H \leq t^L$ . In the latter case, even though the second best may narrow the gap in the types’ duration of experimentation, the gap in their success rates widens.

Our analysis also bears on the relationship between innovation rates and the quality of the underlying environment. Absent any distortions, “better environments” lead to more success. In particular, an increase in the agent’s average ability,  $\mu_0\lambda^H + (1 - \mu_0)\lambda^L$ , yields a higher probability of success in the first best.<sup>34</sup> However, contracts designed in the presence of moral hazard and adverse selection need not produce this property. The reason is that an improvement in the agent’s average ability can make it optimal for the principal to distort experimentation by more: as shown in Proposition 2,  $\bar{t}^L$  decreases in  $\mu_0$  and, for some parameter values, in  $\lambda^H$ . Such a reduction in  $\bar{t}^L$  can decrease the second-best average success probability when the agent’s average ability increases. Consequently, observing higher innovation rates in contractual settings is neither necessary nor sufficient to deduce a better underlying environment.

**Contract farming and technology adoption.** Though our model is not developed to explain a particular application, our framework speaks to contract farming and, more broadly, technology adoption in developing countries. Technology adoption is inherently a dynamic process of experimentation and learning. Understanding the adoption of agricultural innovations in low-income countries, and the obstacles to it, has been a central topic in development economics (Feder *et al.*, 1985; Foster and Rosenzweig, 2010). Practitioners, policymakers, and researchers have long recognized the importance of contractual arrangements to provide proper incentives, because farmers typically do not internalize the broader benefits of their experimentation.

As described in the Introduction, contract farming is a common practice in developing countries; it involves a profit-maximizing firm, which is typically a large-scale buyer such as an exporter or a food processor, and a farmer, who may be a small or a large grower. The contractual environment features not only learning about the quality of new seeds or a new technology, but also moral hazard and unobservable heterogeneity (Miyata *et al.*, 2009).<sup>35</sup> The arrangements used between agricultural firms and farmers resemble the contracts characterized in Theorem 4, with firms committing to time-dated future prices for an output of a certain quality delivered by a given deadline. Our analysis shows why such contracts are optimal in the presence of uncertainty, moral hazard, and unobservable heterogeneity, and how the shape of the contract hinges on the interaction of these three key features. Theorem 4 and formula (5.3) reveal how an optimal pattern of outcome-contingent buyback prices should be determined. In principle, these predicted contracts could be subject to empirical testing.

Much of the recent research on technology adoption uses controlled field experiments to study the incentives of potential adopters. Our results may inform the design of experimental work, particularly with regards to dynamic considerations, which are receiving increasing attention. For example, Jack *et al.* (2014) use a field experiment to study both the initial take-up decision and the subsequent investment (follow-through) decisions in the context of agricultural technology (tree species) adoption in Zambia. The authors consider simple contracts to investigate the interplay between the uncertainty of a technology’s profitability, the self-selection of farmers, and learning of new information. In their experimental design, contracts specify the initial price of the technology and an outcome-contingent payment tied to the survival of trees by the end of one year. The study uses variation of the contracts in the two dimensions (initial price and contingent payment) to evaluate their performance. The authors find that 35% of farmers who

34. Although  $t^L$  and  $t^H$  may increase or decrease in  $\lambda^L$  and  $\lambda^H$  respectively, one can show that the first-best probability of success is always increasing in  $\mu_0$ ,  $\lambda^L$ , and  $\lambda^H$ .

35. Using a field experiment, Kelsey (2013) shows that landholders have private information relevant to their performance under a contract that offers incentives for afforestation, and that efficiency can be increased by using an allocation mechanism that induces self-selection.

pay a positive price for take-up have no trees one year later; in addition, among farmers who follow-through, the tree survival rate responds to learning over time.

The contract form used in Jack *et al.* (2014) shares features with what emerges as an optimal contract in our model, and their basic findings are also consistent with our results. Their controlled experiment is simple in that performance is assessed and a reward is paid only at the end of one year. Our model shows that to optimally incentivize experimentation, agents must be compensated with continual rewards contingent on the time of success, up until an optimally chosen termination date which may differ from the efficient stopping time. Moreover, perhaps counterintuitively, Theorem 4 shows that higher rewards must be offered for later success, with the rate of increase depending on the rate of learning (and other factors).<sup>36</sup> Our results thus point to a new dimension that can improve follow-through rates; this could be tested in future field experiments.

Finally, many scholars study the puzzle of low technology adoption rates and its potential solutions (*e.g.* Suri, 2011, and the references therein). Our paper adds to the discussion by relating adoption rates to the underlying contractual environment. As mentioned earlier, we predict less experimentation, lower success rates, and more dispersion of success rates in settings with more asymmetric information; the lower (and more dispersed) success rates translate into lower (and more dispersed) adoption rates. We also find that the relationship between adoption rates and the underlying environment can be subtle, with “better environments” possibly leading to less experimentation and lower adoption in the second best. Our results thus provide a novel explanation for the low adoption rate puzzle. Empirical researchers have recently been interested in how agency contributes to the puzzle (*e.g.* Atkin *et al.*, 2015); our work contributes to the theoretical background for such lines of inquiry.

Naturally, there are dimensions of contract farming and technology adoption that our analysis does not cover. For example, social learning among farmers affects adoption (Conley and Udry, 2010), and agricultural companies will want to take this into account when designing contracts.<sup>37</sup> A deeper understanding of optimal contracts for multiple experimenting agents who can learn from each other would be useful for this application.<sup>38</sup> While this and similar extensions may yield new insights, we expect our main results to be robust: to reduce the information rent of high-ability types, the principal will benefit from distorting the length of experimentation of low-ability types, and from setting payments so that their incentive constraint for effort binds at each time. This suggests that, under appropriate conditions, an agent will still receive a higher reward for succeeding later rather than earlier.

**Book contracts.** As mentioned in the Introduction, some contractual relationships between a publisher and author have the features we study: it is initially uncertain whether a satisfactory book can be written in the relevant timeframe; the author may be privately informed about his suitability for the task; and how much time the author spends on this is not observable to the publisher. It is common for real-world publishing contracts to resemble the penalty contracts characterized in Theorem 3: book contracts pay an advance to the author that the publisher can recoup if the author fails to deliver on time (according to a delivery-of-manuscript clause) or if the book is unacceptable (according to a satisfactory-manuscript clause); see Bunnin (1983) and Fowler (1985). There is substantial dispersion in both the deadlines and the advances that

36. In particular, formula (5.3) reveals that rewards will optimally increase more sharply over time, up until the contract termination, if the rate of learning is higher.

37. Another aspect is the choice of farmer size: as discussed in Miyata *et al.* (2009), there are different advantages to contracting with small versus large growers, and the optimal farmer size for a firm may change as parties experiment and learn over time.

38. Recent work on this agenda, albeit without adverse selection, includes Frick and Ishii (2015) and Moroni (2015).

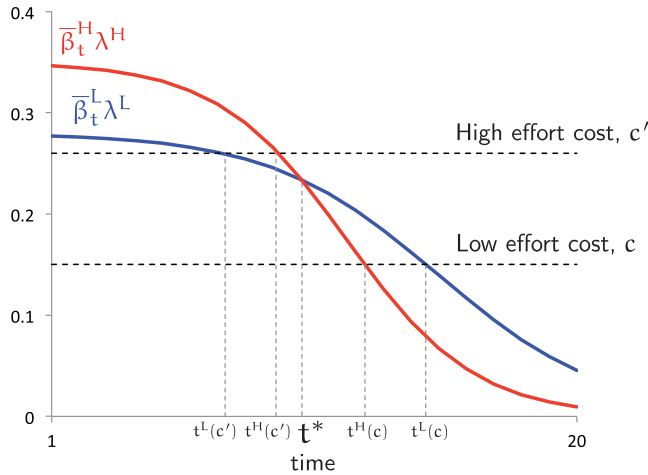


FIGURE 3

No-shirk expected marginal product curves with  $\beta_0 = 0.99, \lambda^L = 0.28, \lambda^H = 0.35$ .

authors are given; Kuzyk (2006) notes that publishing houses try to assess an author's chances of succeeding when determining these terms.

### 6. OPTIMAL CONTRACTS WHEN $t^H \leq t^L$

We now turn to characterizing optimal contracts when the first-best stopping times are ordered  $t^H \leq t^L$ . Any of the following conditions on the primitives is sufficient for this case given a set of other parameters: (1)  $\beta_0$  is large enough, (2)  $\lambda^H$  is large enough, or (3)  $c$  is small enough.

The principal's program remains as described in Section 4, but solving the program is now substantially more difficult than when  $t^H > t^L$ . To understand why, consider Figure 3, which depicts the two types' "no-shirk expected marginal product" curves,  $\bar{\beta}_t^\theta \lambda^\theta$ , as a function of time. (For simplicity, the figure is drawn ignoring integer constraints.) For any parameters, these curves cross exactly once as shown in the figure; the crossing point  $t^*$  is the unique solution to

$$\bar{\beta}_{t^*}^H \lambda^H - \bar{\beta}_{t^*}^L \lambda^L \geq 0 > \bar{\beta}_{t^*+1}^H \lambda^H - \bar{\beta}_{t^*+1}^L \lambda^L.$$

Parameters under which  $t^H > t^L$  entail  $t^L < t^*$ , as seen with the high effort cost in Figure 3. When  $t^L \leq t^*$ , it holds at any  $t \leq t^L$  that the high type has a higher expected marginal product than the low type conditional on the agent working in all prior periods. It is this fact that allowed us to prove Theorem 3 by conjecturing that the high type would work in every period when taking the low type's contract.

By contrast,  $t^H \leq t^L$  implies  $t^L \geq t^*$ , as seen with the low effort cost in Figure 3. Since the second-best stopping time for the low type can be arbitrarily close to his first-best stopping time (e.g. if the prior on the low type,  $1 - \mu_0$ , is sufficiently large), it is no longer valid to conjecture that the high type will work in every period when taking the low type's optimal contract—in this sense, "single crossing" need not hold even at the optimum. The reason is that at some period after  $t^*$ , given that both types have worked in each prior period, the high type can be sufficiently more pessimistic than the low type that the high type finds it optimal to shirk in some or all of the remaining periods, even though  $\lambda^H > \lambda^L$  and the low type would be willing to work for the

contract’s duration.<sup>39</sup> Indeed, this will necessarily be true in the last period of the low type’s contract if this period is later than  $t^*$  and the contract makes the low type just indifferent between working and shirking in this period as in the characterization of Theorem 3.

Solving the principal’s program without being able to restrict attention to some suitable subset of action plans for the high type when he takes the low type’s contract appears intractable. For an arbitrary  $\delta$ , we have been unable to find a valid restriction. The following example elucidates the difficulties.

**Example 1.** For an open and dense set of parameters  $\{\beta_0, c, \lambda^L, \lambda^H\}$  with  $\bar{t}^L = t^L = 3$ ,<sup>40</sup> there is a  $\delta^* \in (0, 1)$  such that the optimal penalty contract for type  $L$  as a function of the discount factor,  $\mathbf{C}^L(\delta) = (3, W_0^L(\delta), I^L(\delta))$ , has the property that the optimal action plans for type  $H$  under this contract are given by

$$\alpha^H(\mathbf{C}^L(\delta)) = \begin{cases} \{(1, 1, 0), (1, 0, 1)\} & \text{if } \delta \in (0, \delta^*) \\ \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\} & \text{if } \delta = \delta^* \\ \{(1, 0, 1), (0, 1, 1)\} & \text{if } \delta \in (\delta^*, 1) \\ \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\} & \text{if } \delta = 1. \end{cases}$$

Figure 4 depicts the contract and type  $H$ ’s optimal action plans as a function of  $\delta$  for a particular set of other parameters.<sup>41</sup> Notice that the only action plan that is optimal for type  $H$  for all  $\delta$  is the non-consecutive-work plan  $(1, 0, 1)$ , but for each value of  $\delta$  at least one other plan is also optimal. Interestingly, the stopping strategy  $(1, 1, 0)$  is not optimal for type  $H$  when  $\delta \in (\delta^*, 1)$  although it is when  $\delta = 1$ . The lack of lower hemi-continuity of  $\alpha^H(\mathbf{C}^L(\delta))$  at  $\delta = 1$  is not an accident, as we will discuss subsequently.

Nevertheless, we are able to solve the problem when  $\delta = 1$ .

**Theorem 5.** Assume  $\delta = 1$  and  $t^H \leq t^L$ . There is an optimal menu in which the principal separates the two types using onetime-penalty contracts,  $\mathbf{C}^H = (t^H, W_0^H, I_{t^H}^H)$  with  $I_{t^H}^H < 0 < W_0^H$  for type  $H$  and  $\mathbf{C}^L = (t^L, W_0^L, I_{t^L}^L)$  for type  $L$ , such that:

- (1)  $I_{t^L}^L = \min \left\{ -\frac{c}{\bar{\beta}_3^L \lambda^L}, -\frac{c}{\bar{\beta}_{t^L}^H \lambda^H} \right\}$ , where  $t^{HL} := \max_{\mathbf{a} \in \alpha^H(\mathbf{C}^L)} \# \{n : a_n = 1\}$ ;
- (2)  $W_0^L > 0$  is such that the participation constraint,  $(\mathbf{IR}^L)$ , binds;

39. More precisely, the relaxed program, [RP1], described in Step 1 of the proof sketch of Theorem 3 can yield a solution that is not feasible in the original program, because the constraint  $(\mathbf{IC}^{HL})$  is violated; the high type would deviate from accepting his contract to accepting the low type’s contract and then shirk in some periods.

40. It suffices for the parameters to satisfy the following four conditions:

- (1) The first-best stopping time for type  $L$  is  $t^L = 3$  (i.e.  $\bar{\beta}_3^L \lambda^L > c > \bar{\beta}_4^L \lambda^L$ ) and the probability of type  $L$  is large enough (i.e.  $\mu_0$  is sufficiently small) that it is not optimal to distort the stopping time of type  $L$ :  $\bar{t}^L = t^L = 3$ .
- (2) The expected marginal product for type  $H$  after one period of work is less than that of type  $L$  after one period of work, but larger than that of type  $L$  after two periods of work:  $\bar{\beta}_3^L \lambda^L < \bar{\beta}_2^H \lambda^H < \bar{\beta}_2^L \lambda^L$ .
- (3) Ex-ante, type  $H$  is more likely to succeed by working in one period than type  $L$  is by working in two periods:  $1 - \lambda^H < (1 - \lambda^L)^2$ .
- (4) There is some  $\delta^* \in (0, 1)$  such that  $\frac{1}{\beta_0 \lambda^L} - \frac{1}{\beta_0 \lambda^H} = \delta^*(1 - \lambda^L) \left( \frac{1}{\bar{\beta}_2^H \lambda^H} - \frac{1}{\bar{\beta}_2^L \lambda^L} \right)$ .

41. The initial transfer  $W_0^L$  in each case is determined by making the participation constraint of type  $L$  bind.

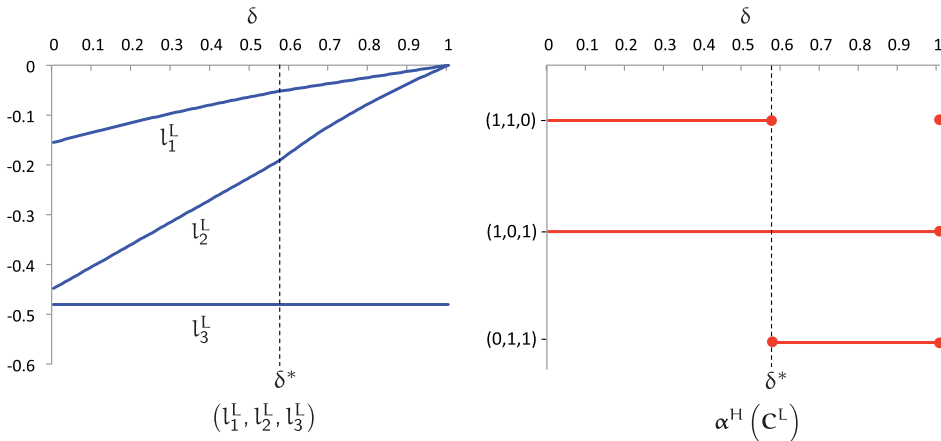


FIGURE 4

The optimal penalty contract for type  $L$  in Example 1 with  $\beta_0=0.86$ ,  $c=0.1$ ,  $\lambda^L=0.75$ ,  $\lambda^H=0.95$  (left graph) and the optimal action profiles for type  $H$  under this contract (right graph).

- (3) Type  $H$  gets an information rent:  $U_0^H(C^H, \alpha^H(C^H)) > 0$ ;
- (4)  $\mathbf{1} \in \alpha^H(C^H)$ ;  $\mathbf{1} \in \alpha^L(C^L)$ .

*Proof* See Appendix C. ||

For  $\delta=1$ , the optimal menus of penalty contracts characterized in Theorem 5 for  $t^H \leq t^L$  share some common properties with those characterized in Theorem 3 for  $t^H > t^L$ : in both cases, a onetime-penalty contract is used for the low type and the high type earns an information rent. On the other hand, part 1 of Theorem 5 points to two differences: (1) it will generally be the case in the optimal  $C^L$  that when  $t^H \leq t^L$ ,  $\mathbf{1} \notin \alpha^H(C^L)$ , whereas for  $t^H > t^L$ ,  $\alpha^H(C^L) = \mathbf{1}$ ; and (2) when  $t^H \leq t^L$ , it can be optimal for the principal to induce the low type to work in each period by satisfying the low type’s incentive constraint for effort with slack (*i.e.* with strict inequality), whereas when  $t^H > t^L$ , the penalty sequence makes this effort constraint bind in each period.

The intuition for these differences derives from information rent minimization considerations. The high type earns an information rent because by following the same effort profile as the low type he is less likely to incur any penalty for failure, and hence has a higher utility from any penalty contract than the low type.<sup>42</sup> Minimizing the rent through this channel suggests minimizing the magnitude of the penalties that are used to incentivize the low type’s effort; it is this logic that drives Theorem 3 and for  $\delta=1$  leads to a onetime-penalty contract with

$$l_{t^L}^L = -\frac{c}{\beta_{t^L}^L \lambda^L}. \tag{6.1}$$

However, when  $t^L > t^*$  (which is only possible when  $t^H \leq t^L$ ), the high type would find it optimal under this contract to work only for some  $T < \bar{t}^L$  number of periods. It is then possible—and is true for an open and dense set of parameters—that  $T$  is such that the high type is more likely to incur the onetime penalty than the low type. But in such a case, the penalty given in (6.1) would

42. Strictly speaking, this intuition applies so long as  $l_t \leq 0$  for all  $t$  in the penalty contract.

not be optimal because the principal can lower  $l_t^L$  (i.e. increase the magnitude of the penalty) to reduce the information rent, which she can keep doing until the high type finds it optimal to work for more periods and becomes less likely to incur the onetime penalty than the low type. This explains part 1 of Theorem 5.

We should note that this possibility arises because time is discrete. It can be shown that when the length of time intervals vanishes, in real-time the  $t^{HL}$  and  $\bar{t}^L$  in the statement of Theorem 5 are such that  $\bar{\beta}_{\bar{t}^L}^L \lambda^L \leq \bar{\beta}_{t^{HL}}^H \lambda^H$  (in particular,  $\bar{\beta}_{\bar{t}^L}^L \lambda^L = \bar{\beta}_{t^{HL}}^H \lambda^H$  when  $\bar{t}^L > t^{HL}$ , or equivalently when  $\bar{t}^L > t^*$ ), and hence  $l_{\bar{t}^L}^L = -\frac{c}{\bar{\beta}_{\bar{t}^L}^L \lambda^L}$  is optimal, just as in Theorem 3 when  $\delta = 1$ . Intuitively, because learning is smooth in continuous time, the high type would always work long enough upon deviating to the low type's contract that he is less likely to incur the onetime penalty  $l_{\bar{t}^L}^L$  than the low type. Thus, by the logic above, lowering the onetime penalty below that in (6.1) would only increase the information rent of the high type in the continuous-time limit.

**Remark 2.** *The proof of Theorem 5 provides an algorithm to solve for an optimal menu of contracts when  $t^H \leq t^L$  and  $\delta = 1$ . For each pair of integers  $(s, t)$  such that  $0 \leq s \leq t \leq t^L$ , one can compute the principal's payoff from using the onetime-penalty contract for type L given by Theorem 5 when  $\bar{t}^L$  is replaced by  $t$  and  $t^{HL}$  is replaced by  $s$ . Optimizing over  $(s, t)$  then yields an optimal (unconstrained) menu.*

How do we prove Theorem 5 in light of the difficulties described earlier of finding a suitable restriction on the high type's behaviour when taking the low-type's contract? The answer is that when  $\delta = 1$ , one can conjecture that the optimal contract for the low type must be a onetime-penalty contract (as was also true when  $t^H > t^L$ ). Notice that because of no discounting, any onetime-penalty contract would make the agent of either type indifferent among all action plans that involve the same number of periods of work. In particular, a stopping strategy—an action plan that involves consecutive work for some number of periods followed by shirking thereafter—is always optimal for either type in a onetime-penalty contract. The heart of the proof of Theorem 5 establishes that it is without loss of generality to restrict attention to penalty contracts for the low type under which the high type would find it optimal to use a stopping strategy (see Appendix C.4 in Appendix C). With this in hand, we are then able to show that a onetime-penalty contract for the low type is indeed optimal (see Appendix C.5). Finally, the rent-minimization considerations described above are used to complete the argument. Observe that optimality of a onetime-penalty contract for the low type and that of a stopping strategy for the high type under such a contract is consistent with the solution in Example 1 for  $\delta = 1$ , as seen in Figure 4. Moreover, the example plainly shows that such a strategy space restriction will not generally be valid when  $\delta < 1$ .<sup>43</sup>

We provide a bonus-contracts implementation of Theorem 5:

**Theorem 6.** *Assume  $\delta = 1$  and  $t^H \leq t^L$ . The second-best can also be implemented using a menu of constant-bonus contracts:  $\mathbf{C}^L = (\bar{t}^L, W_0^L, b^L)$  with  $b^L = -l_{\bar{t}^L}^L > 0 > W_0^L$  where  $l_{\bar{t}^L}^L$  is given in Theorem 5, and  $\mathbf{C}^H = (t^H, W_0^H, b^H)$  with a suitably chosen  $W_0^H$  and  $b^H > 0$ .*

43. Due to the agent's indifference over all action plans that involve the same number of periods of work in a onetime-penalty contract when  $\delta = 1$ , the correspondence  $\alpha^H(\mathbf{C}^L(\delta))$  will generally fail lower hemi-continuity at  $\delta = 1$ . In particular, the low type's optimal contract for  $\delta$  close to 1 may be such that a stopping strategy is not optimal for the high type under this contract. However, the correspondence  $\alpha^H(\mathbf{C}^L(\delta))$  is upper hemi-continuous and the optimal contract is continuous at  $\delta = 1$ . All these points can be seen in Figure 4.



A proof is omitted since this result follows directly from Theorem 5 and the proof of Proposition 1 (using  $\delta = 1$ ). For similar reasons to those discussed around Theorem 4, the implementation in Theorem 6 satisfies interim participation constraints whereas that of Theorem 5 does not.

We end this section by emphasizing that although we are unable to characterize second-best optimal contracts when  $\delta < 1$  and  $t^H \leq t^L$ , the (in)efficiency conclusions from Theorem 2 apply for all parameters.

## 7. DISCUSSION

### 7.1. *Private observability and disclosure*

Suppose that project success is privately observed by the agent but can be verifiably disclosed. The principal's payoff from project success obtains here only when the agent discloses it, and contracts are conditioned not on project success but rather the disclosure of project success. Private observability introduces additional constraints for the principal because the agent must also now be incentivized to not withhold project success. For example, in a bonus contract where  $\delta b_{t+1} > b_t$ , an agent who obtains success in period  $t$  would strictly prefer to withhold it and continue to period  $t+1$ , shirk in that period, and then reveal the success at the end of period  $t+1$ . Nevertheless, we show in the Supplementary Appendix that private observability does not reduce the principal's payoff compared to our baseline setting: in each of the menus identified in Theorems 3–6, each of the contracts would induce the agent (of either type) to reveal project success immediately when it is obtained, so these menus remain optimal and implement the same outcome as when project success is publicly observable.<sup>44</sup>

### 7.2. *Limited liability*

To focus on the interaction of adverse selection and moral hazard in experimentation, we have abstracted away from limited-liability considerations. Consider introducing the requirement that all transfers must be above some minimum threshold, say zero. The Supplementary Appendix shows how such a limited-liability constraint alters the second-best solution for the case of  $t^H > t^L$  and  $\delta = 1$ . This constraint results in both types of the agent acquiring a rent, so long as they are both induced to experiment. Three points are worth emphasizing. First, each type's second-best stopping time is no larger than his first-best stopping time. The logic precluding over-experimentation, however, is somewhat different—and simpler—than without limited liability: inducing over-experimentation requires paying a bonus of more than one (the principal's value of success) in the last period of the contract in which the agent works, implying a loss for the principal which under limited liability cannot be offset through an up-front payment. Second, while both types' second-best stopping times are now (typically) distorted, their ordering is the same as without limited liability (*i.e.*  $\bar{t}^L \leq \bar{t}^H$ ). The reason is that the principal could otherwise improve upon the menu by just offering both types the low type's contract, which would induce the high type to experiment longer without increasing the high type's payoff. Third, the principal can implement the second-best stopping time for the low type by using a constant-bonus contract of the form described in Theorem 4 (with  $\delta = 1$ ). This contract ensures that the low type's incentive

44. However, unlike the menus of Theorems 3–6, not every optimal menu under public observability is optimal under private observability. In this sense, these menus have a desirable robustness property that other optimal menus need not.

constraint for effort binds in each period, and thus it minimizes both the rent that the low type obtains from his contract and the high type's payoff from taking the low type's contract.

We should note that in our dynamic setting, there are less severe forms of limited liability that may be relevant in applications. For example, one may only require that the sum of penalties at any point do not exceed the initial transfer given to the agent.<sup>45</sup> We conjecture that similar conclusions to those discussed above would also emerge under such a requirement, as both types of the agent will again acquire a rent.

### 7.3. *The role of learning*

We have assumed that  $\beta_0 \in (0, 1)$ . If instead  $\beta_0 = 1$  then there would be no learning about the project quality and the first best would entail both types working until project success has been obtained. How is the second best affected by  $\beta_0 = 1$ ?

Suppose, for simplicity, that there is some (possibly large) exogenous date  $\bar{T}$  at which the game ends. The first-best stopping times are then  $t^L = t^H = \bar{T}$ . The principal's program can be solved here just as in Section 5, because  $\bar{\beta}_t^H \lambda^H = \lambda^H > \bar{\beta}_t^L \lambda^L = \lambda^L$  for all  $t \leq \bar{T}$ .<sup>46</sup> In the absence of learning, the social surplus from the low type working is constant over time. So long as parameters are such that it is not optimal for the principal to exclude the low type (*i.e.*  $\bar{t}^L > 0$ ), it turns out that there is no distortion:  $\bar{t}^L = \bar{t}^H = \bar{T}$ . We provide a more complete argument in the Supplementary Appendix, but to see the intuition consider a large  $\bar{T}$ . Then, even though both types are likely to succeed prior to  $\bar{T}$ , the probability of reaching  $\bar{T}$  without a success is an order of magnitude higher for the low type because  $\left(\frac{1-\lambda^L}{1-\lambda^H}\right)^t \rightarrow \infty$  as  $t \rightarrow \infty$ . Hence, it would not be optimal to locally distort the length of experimentation from  $\bar{T}$  because such a distortion would generate a larger efficiency loss from the low type than a gain from reducing the high type's information rent. By contrast, when  $\beta_0 < 1$  and there is learning, this logic fails because the incremental social surplus from the low type working vanishes over time. Therefore, learning from experimentation plays an important role in our results: for any parameters with  $\beta_0 < 1$  under which there is distortion of the low type's length of experimentation without entirely excluding him, there would instead be no distortion were  $\beta_0 = 1$ .

### 7.4. *Adverse selection on other dimensions*

Another important modelling assumption in this paper is that pre-contractual hidden information is about the agent's ability. An alternative is to suppose that the agent has hidden information about his cost of effort but his ability is commonly known; specifically, the low type's cost of working in any period is  $c^L > 0$ , whereas the high type's cost is  $c^H \in (0, c^L)$ . It is immediate that the first-best stopping time for the high type would always be larger than that of the low type because there is no speed-of-learning effect. Hence, the problem can be solved following our approach in Section 5 for  $t^H > t^L$ .<sup>47</sup> However, not only would this alternative model miss the

45. Biais *et al.* (2010) study such a limited-liability requirement in a setting without adverse selection or learning, where large losses arrive according to a Poisson process whose intensity is determined by the agent's effort.

46. It should be clear that nothing would have changed in the analysis in Section 5 if we had assumed existence of a suitably large end date, in particular so long as  $\bar{T} \geq \max\{t^H, t^L\}$ .

47. This applies to binary effort choices. Another alternative would be for the agent to choose effort from a richer set, *e.g.*  $\mathbb{R}_+$ , and effort costs be convex with one type having a lower marginal cost than the other. The speed-of-learning effect would emerge in this setting because the two types would generally choose different effort levels in any period. Analysing such a problem is beyond the scope of this paper.

considerations involved with  $t^H \leq t^L$ , but furthermore, it also obviates interesting features of the problem even when  $t^H > t^L$ . For example, in this setting it would be optimal for the high type to work in all periods in any contract in which it is optimal for the low type to work in all periods; recall that this is not true in our model even when  $t^H > t^L$  (cf. fn. 28).

Another source of adverse selection would be private information about project quality. Specifically, suppose that the agent's ability is commonly known but, prior to contracting, he receives a private signal about the true project quality: there is a high type whose belief that the state is good is  $\beta_0^H \in (0, 1)$  and a low type whose belief is  $\beta_0^L \in (0, \beta_0^H)$ .<sup>48</sup> Again, the first-best stopping times here would always have  $t^H > t^L$  and the problem can be studied following our approach to this case.

## APPENDICES: NOTATION AND TERMINOLOGY

It is convenient in proving our results to work with an apparently larger set of contracts than that defined in the main text. Specifically, in the Appendices, we assume that the principal can stipulate binding "lockout" periods in which the agent is prohibited from working. As discussed in fn. 14 of the main text, this instrument does not yield any benefit to the principal because suitable transfers can be used to ensure that the agent shirks in any desired period regardless of his type and action history. Nevertheless, stipulating lockout periods simplifies the phrasing of our arguments; we use it, in particular, to prove that an optimal contract for the low type never induces him to shirk before termination.

Accordingly, we denote a general contract by  $\mathbf{C} = (\Gamma, W_0, \mathbf{b}, I)$ , where all the elements are as introduced in the main text, except that instead of having the termination date of the contract in the first component, we now have a set of periods,  $\Gamma \subseteq \mathbb{N} \setminus \{0\}$ , at which the agent is not locked out, *i.e.* at which he is allowed to choose whether to work or shirk. Note that, without loss,  $\mathbf{b} = (b_t)_{t \in \Gamma}$  and  $I = (l_t)_{t \in \Gamma}$ ,<sup>49</sup> and the agent's actions are denoted by  $\mathbf{a} = (a_t)_{t \in \Gamma}$ , where  $a_t = 1$  if the agent works in period  $t \in \Gamma$  and  $a_t = 0$  if the agent shirks. The termination date of the contract is 0 if  $\Gamma = \emptyset$  and is otherwise  $\max \Gamma$ , which we require to be finite.<sup>50</sup> We say that a contract is *connected* if  $\Gamma = \{1, \dots, T\}$  for some  $T$ ; in this case we refer to  $T$  as the *length* of the contract,  $T$  is also the termination date, and we write  $\mathbf{C} = (T, W_0, \mathbf{b}, I)$ .

Given some program for the principal, we say that a simplified program entails *no loss of optimality* if the value of the two programs is the same.

### A. PROOF OF THEOREM 2

Without loss by Proposition 1, we focus on penalty contracts throughout the proof.

#### A.1. Step 1: Low type always works

We show that it is without loss of optimality to focus on contracts for the low type,  $\mathbf{C}^L = (\Gamma^L, W_0^L, I^L)$ , in which the low type works in all periods  $t \in \Gamma^L$ . Denote the set of penalty contracts by  $\mathcal{C}$ , and recall that the principal's program, with the restriction to penalty contracts, is:

$$\max_{(\mathbf{C}^H \in \mathcal{C}, \mathbf{C}^L \in \mathcal{C}, \mathbf{a}^H, \mathbf{a}^L)} \mu_0 \Pi_0^H(\mathbf{C}^H, \mathbf{a}^H) + (1 - \mu_0) \Pi_0^L(\mathbf{C}^L, \mathbf{a}^L)$$

subject to, for all  $\theta, \theta' \in \{L, H\}$ ,

$$\mathbf{a}^\theta \in \alpha^\theta(\mathbf{C}^\theta), \quad (\text{IC}^\theta_a)$$

$$U_0^\theta(\mathbf{C}^\theta, \mathbf{a}^\theta) \geq 0, \quad (\text{IR}^\theta)$$

$$U_0^\theta(\mathbf{C}^\theta, \mathbf{a}^\theta) \geq U_0^\theta(\mathbf{C}^{\theta'}, \alpha^\theta(\mathbf{C}^{\theta'})). \quad (\text{IC}^{\theta\theta'})$$

Suppose there is a solution to this program,  $(\mathbf{C}^H, \mathbf{C}^L, \mathbf{a}^H, \mathbf{a}^L)$ , with  $\mathbf{a}^L \neq \mathbf{1}$  and  $\mathbf{C}^L = (\Gamma^L, W_0^L, I^L)$ . It suffices to show that there is another solution to the program,  $(\mathbf{C}^H, \widehat{\mathbf{C}}^L, \mathbf{a}^H, \mathbf{1})$ , where  $\widehat{\mathbf{C}}^L = (\widehat{\Gamma}^L, \widehat{W}_0^L, \widehat{I}^L)$  is such that:

48. Private information about project quality is studied by Gomes *et al.* (2015) in experimentation without moral hazard, and in a different setting by Gerardi and Maestri (2012). Another possibility would be non-common priors between the principal and the agent, which would involve quite distinct considerations.

49. There is no loss in not allowing for transfers in lockout periods.

50. One can show that this restriction does not hurt the principal.

- (i)  $\mathbf{1} \in \alpha^L(\widehat{\mathbf{C}}^L)$ ;
- (ii)  $U_0^L(\mathbf{C}^L, \mathbf{a}^L) = U_0^L(\widehat{\mathbf{C}}^L, \mathbf{1})$ ;
- (iii)  $\Pi_0^L(\mathbf{C}^L, \mathbf{a}^L) = \Pi_0^L(\widehat{\mathbf{C}}^L, \mathbf{1})$ ; and
- (iv)  $U_0^H(\mathbf{C}^L, \alpha^H(\mathbf{C}^L)) \geq U_0^H(\widehat{\mathbf{C}}^L, \alpha^H(\widehat{\mathbf{C}}^L))$ .

To this end, let  $t = \min\{s : a_s = 0\}$  and denote the largest preceding period in  $\Gamma^L$  as

$$p(t) = \begin{cases} \max \Gamma^L \setminus \{t, t+1, \dots\} & \text{if } \exists s \in \Gamma^L \text{ s.t. } s < t, \\ 0 & \text{otherwise.} \end{cases}$$

Construct  $\widehat{\mathbf{C}}^L = (\widehat{\Gamma}^L, \widehat{W}_0^L, \widehat{I}^L)$  as follows:

$$\begin{aligned} \widehat{\Gamma}^L &= \Gamma^L \setminus \{t\}; \\ \widehat{I}_s^L &= \begin{cases} I_s^L & \text{if } s \neq p(t) \text{ and } s \in \widehat{\Gamma}^L, \\ I_s^L + \delta^{t-p(t)} I_t^L & \text{if } s = p(t) > 0; \end{cases} \\ \widehat{W}_0^L &= \begin{cases} W_0^L & \text{if } p(t) > 0, \\ W_0^L + \delta^t I_t^L & \text{if } p(t) = 0. \end{cases} \end{aligned}$$

Notice that under contract  $\mathbf{C}^L$ , the profile  $\mathbf{a}^L$  has type  $L$  shirking in period  $t$  and thus receiving  $I_t^L$  with probability one conditional on not succeeding before this period; the new contract  $\widehat{\mathbf{C}}^L$  just locks the agent out in period  $t$  and shifts the payment  $I_t^L$  up to the preceding non-lockout period, suitably discounted. It follows that the incentives for effort for type  $L$  remain unchanged in any other period; moreover, since  $a_t^L = 0$ , both the principal's payoff from type  $L$  under this contract and type  $L$ 's payoff do not change. Finally, observe that for type  $H$ , no matter which action he would take at  $t$  in any optimal action plan under  $\mathbf{C}^L$  (whether it is work or shirk), his payoff from  $\widehat{\mathbf{C}}^L$  must be weakly lower because the lockout in period  $t$  is effectively as though he has been forced to shirk in period  $t$  and receive  $I_t^L$ .

Performing this procedure repeatedly for each period in which the original profile  $\mathbf{a}^L$  prescribes shirking yields a final contract  $\widetilde{\mathbf{C}}^L$  which satisfies all the desired properties.

### A.2. Step 2: Simplifying the principal's problem

By Step 1, we can focus on the following program [P]:

$$\max_{(\mathbf{C}^H \in \mathcal{C}, \mathbf{C}^L \in \mathcal{C}, \mathbf{a}^H)} \mu_0 \Pi_0^H(\mathbf{C}^H, \mathbf{a}^H) + (1 - \mu_0) \Pi_0^L(\mathbf{C}^L, \mathbf{1}) \tag{P}$$

subject to

$$\mathbf{1} \in \alpha^L(\mathbf{C}^L), \tag{IC_a^L}$$

$$\mathbf{a}^H \in \alpha^H(\mathbf{C}^H), \tag{IC_a^H}$$

$$U_0^L(\mathbf{C}^L, \mathbf{1}) \geq 0, \tag{IR^L}$$

$$U_0^H(\mathbf{C}^H, \mathbf{a}^H) \geq 0, \tag{IR^H}$$

$$U_0^L(\mathbf{C}^L, \mathbf{1}) \geq U_0^L(\mathbf{C}^H, \alpha^L(\mathbf{C}^H)), \tag{IC^{LH}}$$

$$U_0^H(\mathbf{C}^H, \mathbf{a}^H) \geq U_0^H(\mathbf{C}^L, \alpha^H(\mathbf{C}^L)). \tag{IC^{HL}}$$

We first show that it is without loss of optimality to ignore constraints (IR<sup>H</sup>) and (IC<sup>LH</sup>).

**Step 2a:** Consider (IR<sup>H</sup>). Define a stochastic action plan  $\sigma = (\sigma_t)_{t \in \Gamma^L}$  for type  $H$  under contract  $\mathbf{C}^L$  as follows:  $\sigma_t \in \Delta(\{0, 1\})$  with  $\sigma_t(1) \equiv \frac{\lambda^L}{\lambda^H}$  and  $\sigma_t(0) \equiv 1 - \frac{\lambda^L}{\lambda^H}$  for all  $t \in \Gamma^L$ . In other words, under  $\sigma$ , the agent works in any period of  $\Gamma^L$  (so long as he not succeeded before) with probability  $\lambda^L/\lambda^H$ . Note that these probabilities are independent across periods. By construction, it holds for all  $t \in \Gamma^L$  that  $\mathbb{E}_\sigma[a_t] = \lambda^L$ , where  $\mathbb{E}_\sigma$  is the ex-ante expectation with respect to the probability measure induced by  $\sigma$ .

Type  $H$ 's expected payoff under contract  $C^L$  given stochastic action plan  $\sigma$  is

$$\begin{aligned}
 U_0^H(C^L, \sigma) &= \beta_0 \sum_{t \in \Gamma^L} \delta^t \mathbb{E}_\sigma \left\{ \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - a_s \lambda^H) \right] [(1 - a_t \lambda^H) l_t^L - a_t c] \right\} + (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t \mathbb{E}_\sigma [(l_t^L - a_t c)] + W_0^L \\
 &= \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} \mathbb{E}_\sigma (1 - a_s \lambda^H) \right] \mathbb{E}_\sigma [(1 - a_t \lambda^H) l_t^L - a_t c] + (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t \mathbb{E}_\sigma [(l_t^L - a_t c)] + W_0^L \\
 &= \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - \lambda^L c] + (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t (l_t^L - \lambda^L c) + W_0^L \\
 &\geq \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - c] + (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t (l_t^L - c) + W_0^L \\
 &= U_0^L(C^L, \mathbf{1}),
 \end{aligned}$$

where the second equality follows from the independence of  $\sigma_t$  and  $\sigma_s$  for all  $t, s \in \Gamma^L$ , the third equality follows from the fact that  $\mathbb{E}_\sigma [a_t] = \lambda^L$  for all  $t \in \Gamma^L$ , and the inequality follows from  $\lambda^L < 1$ .<sup>51</sup>

It follows immediately from the above string of (in)equalities that there exists a pure action plan  $\mathbf{a} = (a_t)_{t \in \Gamma^L}$  such that

$$U_0^H(C^L, \mathbf{a}) \geq U_0^H(C^L, \sigma) \geq U_0^L(C^L, \mathbf{1}) \geq 0,$$

where the last inequality follows from  $(IR^L)$ . Therefore,  $(IC^{HL})$  implies that

$$U_0^H(C^H, \mathbf{a}^H) \geq U_0^H(C^L, \alpha^H(C^L)) \geq U_0^H(C^L, \mathbf{a}) \geq 0,$$

which establishes  $(IR^H)$ .

**Step 2b:** Consider next  $(IC^{LH})$ . By the same arguments as in Step 1, without loss of optimality we can restrict attention to contracts for the high type  $C^H$  in which the high type works in all periods  $t \in \Gamma^H$ . If an optimal contract  $C^H$  has  $\Gamma^H = \emptyset$ ,  $(IC^{LH})$  is trivially satisfied.<sup>52</sup> Thus, assume an optimal contract  $C^H$  has  $\Gamma^H \neq \emptyset$ . Let  $T^H = \max \Gamma^H$  and denote type  $H$ 's expected payoff under  $C^H$  by  $\bar{U}_0^H$ . We show that there exists a onetime-penalty contract that yields the principal the same expected payoff as  $C^H$  and satisfies  $(IC^{LH})$ . Consider a family of contracts  $\hat{C}^H = (\Gamma^H, \hat{W}_0^H, \hat{l}_{T^H}^H)$ , where  $\hat{l}_{T^H}^H$  and  $\hat{W}_0^H$  jointly ensure that type  $H$  works in all periods  $t \in \Gamma^H$  and his expected payoff under  $\hat{C}^H$  is equal to  $\bar{U}_0^H$ :

$$\left[ \prod_{t \in \Gamma^H} (1 - \lambda^H) \beta_0 + (1 - \beta_0) \right] \delta^{T^H} \hat{l}_{T^H}^H - c \left[ \beta_0 \sum_{t \in \Gamma^H} \delta^t \prod_{s \in \Gamma^H, s < t} (1 - \lambda^H) - (1 - \beta_0) \sum_{t \in \Gamma^H} \delta^t \right] + \hat{W}_0^H = \bar{U}_0^H. \tag{A.1}$$

It is immediate that any such contract  $\hat{C}^H$  yields the principal the same expected payoff from type  $H$  as the original contract  $C^H$ , as it leaves both type  $H$ 's action plan and type  $H$ 's expected payoff under the new contract unchanged from the original contract. Furthermore, note that the penalty  $\hat{l}_{T^H}^H$  can be chosen to be severe enough (*i.e.* sufficiently negative) to ensure that it is also optimal for type  $L$  to work in all periods after accepting contract  $\hat{C}^H$ ; *i.e.* we can choose  $\hat{l}_{T^H}^H$  so that for all  $\theta \in \{L, H\}$ ,  $\alpha^\theta(\hat{C}^H) = \mathbf{1}$ . All that remains is to show that a sufficiently severe  $\hat{l}_{T^H}^H$  and its corresponding  $\hat{W}_0^H$

51. As a notational convention, the expression  $\prod_{s \in \Gamma^L} (1 - \lambda^L)$  means  $(1 - \lambda^L)^{|\Gamma^L|}$ , and analogously for similar expressions.

52. If an optimal contract  $C^H$  excludes type  $H$ , then without loss it can be taken to involve no transfers at all, which ensures that it would yield type  $L$  a zero payoff, and hence  $(IC^{LH})$  follows from  $(IR^L)$ .

(determined by (A.1)) also satisfy (IC<sup>LH</sup>) given that  $\alpha^L(\widehat{\mathbf{C}}^H) = \mathbf{1}$ . To show this, note that type  $L$ 's expected payoff from taking contract  $\widehat{\mathbf{C}}^H$  and working in all periods  $t \in \Gamma^H$  is

$$U_0^L(\mathbf{C}^H, \mathbf{1}) = \left[ \prod_{t \in \Gamma^H} (1 - \lambda^L) \beta_0 + (1 - \beta_0) \right] \delta^{T^H} \widehat{I}_{T^H}^H - c \left[ \beta_0 \sum_{t \in \Gamma^H} \delta^t \prod_{s \in \Gamma^H, s < t} (1 - \lambda^L) + (1 - \beta_0) \sum_{t \in \Gamma^H} \delta^t \right] + \widehat{W}_0^H. \quad (\text{A.2})$$

It follows from (A.2) and (A.1) that

$$U_0^L(\mathbf{C}^H, \mathbf{1}) = \left[ \prod_{t \in \Gamma^H} (1 - \lambda^L) - \prod_{t \in \Gamma^H} (1 - \lambda^H) \right] \beta_0 \delta^{T^H} \widehat{I}_{T^H}^H + c \beta_0 \sum_{t \in \Gamma^H} \delta^t \left[ \prod_{s \in \Gamma^H, s < t} (1 - \lambda^H) - \prod_{s \in \Gamma^H, s < t} (1 - \lambda^L) \right] + \overline{U}_0^H.$$

Since  $\Gamma^H \neq \emptyset$  and  $\prod_{t \in \Gamma^H} (1 - \lambda^L) - \prod_{t \in \Gamma^H} (1 - \lambda^H) > 0$ ,  $\widehat{I}_{T^H}^H$  can be chosen sufficiently negative such that  $U_0^L(\mathbf{C}^H, \mathbf{1}) < 0$ , establishing (IC<sup>LH</sup>).

**Step 2c:** By Step 2a and Step 2b, it is without loss of optimality to ignore (IR<sup>H</sup>) and (IC<sup>LH</sup>) in program [P]. Ignoring these two constraints yields the following program [P1]:

$$\max_{(\mathbf{C}^H \in \mathcal{C}, \mathbf{C}^L \in \mathcal{C}, \mathbf{a}^H)} \mu_0 \Pi_0^H(\mathbf{C}^H, \mathbf{a}^H) + (1 - \mu_0) \Pi_0^L(\mathbf{C}^L, \mathbf{1}) \quad (\text{P1})$$

subject to

$$\mathbf{1} \in \alpha^L(\mathbf{C}^L), \quad (\text{IC}_a^L)$$

$$\mathbf{a}^H \in \alpha^H(\mathbf{C}^H), \quad (\text{IC}_a^H)$$

$$U_0^L(\mathbf{C}^L, \mathbf{1}) \geq 0, \quad (\text{IR}^L)$$

$$U_0^H(\mathbf{C}^H, \mathbf{a}^H) \geq U_0^H(\mathbf{C}^L, \alpha^H(\mathbf{C}^L)). \quad (\text{IC}^{HL})$$

It is clear that in any solution to program [P1], (IR<sup>L</sup>) must be binding; otherwise, the initial time-zero transfer from the principal to the agent in the contract  $\mathbf{C}^L$  can be reduced slightly to strictly improve the second term of the objective function while not violating any of the constraints. Similarly, (IC<sup>HL</sup>) must also bind because otherwise the time-zero transfer in the contract  $\mathbf{C}^H$  can be reduced to improve the first term of the objective function without violating any of the constraints.

Using these two binding constraints, substituting in the formulae from equations (2.1) and (2.2), and letting the principal select the optimal action plan the high type should use when taking the low type's contract ( $\mathbf{a}^{HL} \in \alpha^H(\mathbf{C}^L)$ ), we can rewrite the objective function (P1) as the expected total surplus less type  $H$ 's "information rent", obtaining the following program that we call [P2]:

$$\max_{\substack{(\mathbf{C}^H \in \mathcal{C}, \\ \mathbf{C}^L \in \mathcal{C}, \\ \mathbf{a}^{HL} \in \alpha^H(\mathbf{C}^L), \mathbf{a}^H)}} \left\{ \begin{aligned} & \mu_0 \left\{ \beta_0 \sum_{t \in \Gamma^H} \delta^t \left[ \prod_{\substack{s \in \Gamma^H \\ s \leq t-1}} (1 - a_s^H \lambda^H) \right] a_t^H (\lambda^H - c) - (1 - \beta_0) \sum_{t \in \Gamma^H} \delta^t a_t^H c \right\} \\ & + (1 - \mu_0) \left\{ \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] (\lambda^L - c) - (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t c \right\} \\ & - \mu_0 \left\{ \begin{aligned} & \beta_0 \sum_{t \in \Gamma^L} \delta^t I_t^L \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t}} (1 - a_s^{HL} \lambda^H) - \prod_{\substack{s \in \Gamma^L \\ s \leq t}} (1 - \lambda^L) \right] \\ & - \beta_0 c \sum_{t \in \Gamma^L} \delta^t a_t^{HL} \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - a_s^{HL} \lambda^H) - \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] \\ & + c \sum_{t \in \Gamma^L} \delta^t (1 - a_t^{HL}) \left[ 1 - \beta_0 + \beta_0 \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] \end{aligned} \right\} \end{aligned} \right. \quad (\text{P2})$$

Information rent of type  $H$

subject to

$$\mathbf{1} \in \arg \max_{(a_t)_{t \in \Gamma^L}} \left\{ \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - a_s \lambda^L) \right] \left[ (1 - a_t \lambda^L) l_t^L - a_t c \right] + (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t \left( l_t^L - a_t c \right) + W_0^L \right\}, \quad (\text{IC}_a^L)$$

$$\mathbf{a}^H \in \arg \max_{(a_t)_{t \in \Gamma^H}} \left\{ \beta_0 \sum_{t \in \Gamma^H} \delta^t \left[ \prod_{\substack{s \in \Gamma^H \\ s \leq t-1}} (1 - a_s \lambda^H) \right] \left[ (1 - a_t \lambda^H) l_t^H - a_t c \right] + (1 - \beta_0) \sum_{t \in \Gamma^H} \delta^t \left( l_t^H - a_t c \right) + W_0^H \right\}. \quad (\text{IC}_a^H)$$

Program [P2] is separable, *i.e.* it can be solved by maximizing (P2) with respect to  $(\mathbf{C}^L, \mathbf{a}^{HL})$  subject to  $(\text{IC}_a^L)$  and separately maximizing (P2) with respect to  $(\mathbf{C}^H, \mathbf{a}^H)$  subject to  $(\text{IC}_a^H)$ .

We denote the information rent of type  $H$  by  $R(\mathbf{C}^L, \mathbf{a}^{HL})$ . Note that given any action plan  $\mathbf{a}$  that type  $H$  uses when taking type  $L$ 's contract,

$$R(\mathbf{C}^L, \mathbf{a}) = U_0^H(\mathbf{C}^L, \mathbf{a}) - U_0^L(\mathbf{C}^L, \mathbf{1}).$$

Hence,  $R(\mathbf{C}^L, \mathbf{a}) = R(\mathbf{C}^L, \hat{\mathbf{a}})$  whenever both  $\mathbf{a}, \hat{\mathbf{a}} \in \alpha^H(\mathbf{C}^L)$ .

It will be convenient at various places to consider the difference in information rents under contracts  $\hat{\mathbf{C}}^L$  and  $\mathbf{C}^L$  and corresponding action plans  $\hat{\mathbf{a}}$  and  $\mathbf{a}$ :

$$\begin{aligned} R(\hat{\mathbf{C}}^L, \hat{\mathbf{a}}) - R(\mathbf{C}^L, \mathbf{a}) &= U_0^H(\hat{\mathbf{C}}^L, \hat{\mathbf{a}}) - U_0^L(\hat{\mathbf{C}}^L, \mathbf{1}) - \left( U_0^H(\mathbf{C}^L, \mathbf{a}) - U_0^L(\mathbf{C}^L, \mathbf{1}) \right) \\ &= \left( U_0^H(\hat{\mathbf{C}}^L, \hat{\mathbf{a}}) - U_0^H(\mathbf{C}^L, \mathbf{a}) \right) - \left( U_0^L(\hat{\mathbf{C}}^L, \mathbf{1}) - U_0^L(\mathbf{C}^L, \mathbf{1}) \right). \end{aligned} \quad (\text{A.3})$$

Moreover, when the action plan does not change across contracts (*i.e.*  $\mathbf{a} = \hat{\mathbf{a}}$  above), (A.3) specializes to

$$R(\hat{\mathbf{C}}^L, \mathbf{a}) - R(\mathbf{C}^L, \mathbf{a}) = \beta_0 \sum_{t \in \Gamma^L} \delta^t (\hat{l}_t^L - l_t^L) \left[ \prod_{s \in \Gamma^L, s \leq t} (1 - a_s \lambda^H) - \prod_{s \in \Gamma^L, s \leq t} (1 - \lambda^L) \right]. \quad (\text{A.4})$$

### A.3. Step 3: Under-experimentation by the low type

Suppose per contra that  $\mathbf{C}^L = (\Gamma^L, W_0^L, l^L)$  is an optimal contract for the low type inducing him to work for  $|\Gamma^L| > t^L$  periods. This implies  $\Gamma^L \neq \emptyset$ . We show that there exists  $\hat{\mathbf{C}}^L = (\hat{\Gamma}^L, \hat{W}_0^L, \hat{l}^L)$  that induces the low type to work for  $|\hat{\Gamma}^L| = |\Gamma^L| - 1$  periods and strictly increases the principal's payoff.

Let  $T = \max \Gamma^L$  and  $\hat{T} = \max \Gamma^L \setminus \{T\}$  be respectively the last and the second to the last non-lockout periods in contract  $\mathbf{C}^L$ . Consider contract  $\hat{\mathbf{C}}^L$  defined as follows:

$$\hat{\Gamma}^L = \Gamma^L \setminus \{T\},$$

$$\hat{l}_t^L = \begin{cases} l_t^L & \text{if } t \in \hat{\Gamma}^L \text{ and } t < \hat{T}, \\ l_{\hat{T}}^L + \delta^{T-\hat{T}}(1 - \lambda^L)l_T^L - \delta^{T-\hat{T}}c & \text{if } t = \hat{T}, \end{cases}$$

and  $\hat{W}_0$  is such that  $(\text{IR}^L)$  binds in contract  $\hat{\mathbf{C}}^L$ . Note that by construction,  $\hat{\mathbf{C}}^L$  gives the agent a continuation payoff in  $\hat{T}$  which is the same the low-type agent would obtain if, given no success in  $\hat{T}$ , the agent were to work in period  $T$ . We proceed in two sub-steps.

#### Step 3a: Type $L$ works in all periods of $\hat{\mathbf{C}}^L$

We first show that type  $L$  works in all periods  $t \in \hat{\Gamma}^L$  in  $\hat{\mathbf{C}}^L$ . Specifically, we show that type  $L$ 's "incentive to work" in any period  $t \in \hat{\Gamma}^L$  under  $\hat{\mathbf{C}}^L$  is the same as his incentive to work in that period  $t$  under the original contract  $\mathbf{C}^L$ ; hence, the fact that type  $L$  is willing to work in all periods  $t \in \Gamma^L$  under  $\mathbf{C}^L$  given that he works in all future periods (by Step 1) implies that is willing to work in all periods  $t \in \hat{\Gamma}^L$  under  $\hat{\mathbf{C}}^L$  given that he works in all future periods.

Type  $L$ 's incentive to work in period  $\hat{T}$  under the original contract  $\mathbf{C}^L$ , given that he works in period  $T$  under such contract, is given by the difference between his continuation payoff from working and his continuation payoff from shirking in  $\hat{T}$ :

$$\begin{aligned} & \left\{ -c + \beta_{\hat{T}}^L (1 - \lambda^L) \left[ l_{\hat{T}}^L + \delta^{T-\hat{T}} (1 - \lambda^L) l_T^L - \delta^{T-\hat{T}} c \right] + (1 - \beta_{\hat{T}}^L) \left( l_{\hat{T}}^L + \delta^{T-\hat{T}} l_T^L - \delta^{T-\hat{T}} c \right) \right\} \\ & - \left\{ l_{\hat{T}}^L + \delta^{T-\hat{T}} \left[ -c + \beta_{\hat{T}}^L (1 - \lambda^L) l_T^L + (1 - \beta_{\hat{T}}^L) l_T^L \right] \right\}. \end{aligned} \quad (\text{A.5})$$

Note that  $\beta_T^L = \bar{\beta}_{|\Gamma^L|-1}^L$  if the agent works in all periods prior to  $\hat{T}$ . Type  $L$  works in period  $\hat{T}$  only if expression (A.5) is non-negative. With some algebra, expression (A.5) can be simplified to

$$-\beta_T^L \lambda^L (l_{T-1}^L + \delta(1-\lambda^L)l_T^L) + \beta_T^L \lambda^L \delta c - c.$$

More generally, type  $L$ 's incentive to work in any period  $t \in \Gamma^L$ ,  $t < T$ , under contract  $\mathbf{C}^L$ , given work in all future periods in  $\Gamma^L$  and a belief  $\beta_t^L$  in period  $t$ , is

$$-\beta_t^L \lambda^L \sum_{\tau \in \Gamma^L, \tau \geq t} \delta^{\tau-t} \left[ \prod_{s \in \Gamma^L, t \leq s \leq \tau} (1-\lambda^L) \right] l_\tau^L + \beta_t^L \lambda^L \sum_{\tau \in \Gamma^L, \tau > t} \delta^{\tau-t} \left[ \prod_{s \in \Gamma^L, t < s \leq \tau} (1-\lambda^L) \right] c - c. \quad (\text{A.6})$$

Note that  $\beta_t^L = \bar{\beta}_{\{s \leq t, s \in \Gamma^L\}}^L$  if the low type works in all periods prior to  $t$ . Under contract  $\hat{\mathbf{C}}^L$ , type  $L$ 's incentive to work in any period  $t \in \hat{\Gamma}^L$ , given work in all future periods and a belief  $\beta_t^L$  in period  $t$ , is

$$-\beta_t^L \lambda^L \sum_{\tau \in \hat{\Gamma}^L, \tau \geq t} \delta^{\tau-t} \left[ \prod_{s \in \hat{\Gamma}^L, t \leq s \leq \tau} (1-\lambda^L) \right] l_\tau^L + \beta_t^L \lambda^L \sum_{\tau \in \hat{\Gamma}^L, \tau > t} \delta^{\tau-t} \left[ \prod_{s \in \hat{\Gamma}^L, t < s \leq \tau} (1-\lambda^L) \right] c - c.$$

By the definition of  $\hat{\Gamma}^L$  and  $\hat{l}_t^L$  above, this expression can be rewritten as

$$\begin{aligned} & -\beta_t^L \lambda^L \sum_{\tau \in \Gamma^L, t \leq \tau < \hat{T}} \delta^{\tau-t} \left[ \prod_{s \in \hat{\Gamma}^L, t \leq s \leq \tau} (1-\lambda^L) \right] l_\tau^L \\ & -\beta_t^L \lambda^L \delta^{\hat{T}-t} \left[ \prod_{s \in \Gamma^L, t \leq s \leq \hat{T}} (1-\lambda^L) \right] \left( l_{\hat{T}}^L + \delta^{T-\hat{T}}(1-\lambda^L)l_{\hat{T}}^L - \delta^{T-\hat{T}}c \right) \\ & + \beta_t^L \lambda^L \sum_{\tau \in \Gamma^L, t < \tau \leq \hat{T}} \delta^{\tau-t} \left[ \prod_{s \in \Gamma^L, t < s \leq \tau} (1-\lambda^L) \right] c - c \\ & = -\beta_t^L \lambda^L \sum_{\tau \in \Gamma^L, t \leq \tau \leq T} \delta^{\tau-t} \left[ \prod_{s \in \hat{\Gamma}^L, t \leq s \leq \tau} (1-\lambda^L) \right] l_\tau^L + \beta_t^L \lambda^L \sum_{\tau \in \Gamma^L, t < \tau \leq T} \delta^{\tau-t} \left[ \prod_{s \in \Gamma^L, t < s \leq \tau} (1-\lambda^L) \right] c - c, \end{aligned}$$

which is equal to expression (A.6) above. Hence, type  $L$  is willing to work in all periods  $t \in \hat{\Gamma}^L$  under contract  $\hat{\mathbf{C}}^L$ .

**Step 3b: Contract  $\hat{\mathbf{C}}^L$  weakly reduces type  $H$ 's information rent**

Since  $|\Gamma^L| > t^L$  and contract  $\hat{\mathbf{C}}^L$  induces type  $L$  to work for  $|\Gamma^L|-1$  periods, it is immediate that  $\hat{\mathbf{C}}^L$  strictly increases surplus from type  $L$  relative to  $\mathbf{C}^L$ . To show that  $\hat{\mathbf{C}}^L$  increases the principal's objective, it is thus sufficient to show that  $\hat{\mathbf{C}}^L$  weakly reduces type  $H$ 's information rent relative to  $\mathbf{C}^L$ .

Let  $\hat{\mathbf{a}}^{HL} \in \alpha^H(\hat{\mathbf{C}}^L)$  be an optimal action plan for type  $H$  under contract  $\hat{\mathbf{C}}^L$ ,  $\hat{\mathbf{a}}^{HL} = (\hat{a}_t^{HL})_{t \in \hat{\Gamma}^L}$ . Define an action plan  $\hat{\mathbf{a}}^{HL(1)}$  for type  $H$  under contract  $\mathbf{C}^L$  as follows:  $\hat{a}_t^{HL(1)} = \hat{a}_t^{HL}$  for  $t \in \hat{\Gamma}^L$  and  $\hat{a}_T^{HL(1)} = 1$ . Note that since  $\mathbf{a}^{HL} \in \alpha^H(\mathbf{C}^L)$ ,  $U_0^H(\mathbf{C}^L, \mathbf{a}^{HL}) \geq U_0^H(\mathbf{C}^L, \hat{\mathbf{a}}^{HL(1)})$ , and hence

$$R(\mathbf{C}^L, \mathbf{a}^{HL}) \geq R(\mathbf{C}^L, \hat{\mathbf{a}}^{HL(1)}). \quad (\text{A.7})$$

Now consider type  $H$ 's information rent under  $\hat{\mathbf{C}}^L$  given optimal action plan  $\hat{\mathbf{a}}^{HL} \in \alpha^H(\hat{\mathbf{C}}^L)$ :

$$\begin{aligned} R(\hat{\mathbf{C}}^L, \hat{\mathbf{a}}^{HL}) &= \beta_0 \sum_{t \in \hat{\Gamma}^L} \delta^t l_t^L \left[ \prod_{s \in \hat{\Gamma}^L, s \leq t} (1-\hat{a}_s^{HL} \lambda^H) - \prod_{s \in \hat{\Gamma}^L, s \leq t} (1-\lambda^L) \right] \\ & - \beta_0 c \sum_{t \in \hat{\Gamma}^L} \delta^t \hat{a}_t^{HL} \left[ \prod_{s \in \hat{\Gamma}^L, s < t} (1-\hat{a}_s^{HL} \lambda^H) - \prod_{s \in \hat{\Gamma}^L, s < t} (1-\lambda^L) \right] \\ & + c \sum_{t \in \hat{\Gamma}^L} \delta^t (1-\hat{a}_t^{HL})(1-\beta_0 + \beta_0 \prod_{s \in \hat{\Gamma}^L, s < t} (1-\lambda^L)). \end{aligned}$$



Using the definition of  $\widehat{\mathbf{C}}^L$ , this can be rewritten as

$$\begin{aligned} R(\widehat{\mathbf{C}}^L, \widehat{\mathbf{a}}^{HL}) = & \beta_0 \sum_{t \in \widehat{\Gamma}^L, t < \widehat{T}} \delta^t l_t^L \left[ \prod_{s \in \widehat{\Gamma}^L, s \leq t} (1 - \widehat{a}_s^{HL} \lambda^H) - \prod_{s \in \widehat{\Gamma}^L, s \leq t} (1 - \lambda^L) \right] \\ & + \beta_0 \delta^{\widehat{T}} \underbrace{\left( l_{\widehat{T}}^L + \delta^{T-\widehat{T}} (1 - \lambda^L) l_{\widehat{T}}^L - \delta^{T-\widehat{T}} c \right)}_{\widehat{l}_{\widehat{T}}^L} \left[ \prod_{s \in \widehat{\Gamma}^L} (1 - \widehat{a}_s^{HL} \lambda^H) - \prod_{s \in \widehat{\Gamma}^L} (1 - \lambda^L) \right] \\ & - \beta_0 c \prod_{t \in \widehat{\Gamma}^L} \delta^t \widehat{a}_t^{HL} \left[ \prod_{s \in \widehat{\Gamma}^L, s < t} (1 - \widehat{a}_s^{HL} \lambda^H) - \prod_{s \in \widehat{\Gamma}^L, s < t} (1 - \lambda^L) \right] \\ & + c \sum_{t \in \widehat{\Gamma}^L} \delta^t (1 - \widehat{a}_t^{HL}) (1 - \beta_0 + \beta_0 \prod_{s \in \widehat{\Gamma}^L, s < t} (1 - \lambda^L)). \end{aligned}$$

Simple algebraic manipulations yield

$$R(\widehat{\mathbf{C}}^L, \widehat{\mathbf{a}}^{HL}) = R(\mathbf{C}^L, \widehat{\mathbf{a}}^{HL(1)}) + \beta_0 \delta^T \prod_{s \in \widehat{\Gamma}^L, s \leq \widehat{T}} (1 - \widehat{a}_s^{HL} \lambda^H) (\lambda^H - \lambda^L) l_{\widehat{T}}^L. \quad (\text{A.8})$$

Note that since type  $L$  is willing to work in period  $T$  under contract  $\mathbf{C}^L$  (by Step 1), it holds that  $l_T^L < 0$ , and thus (A.8) yields

$$R(\widehat{\mathbf{C}}^L, \widehat{\mathbf{a}}^{HL}) - R(\mathbf{C}^L, \widehat{\mathbf{a}}^{HL(1)}) < 0.$$

Using (A.7), this implies  $R(\widehat{\mathbf{C}}^L, \widehat{\mathbf{a}}^{HL}) < R(\mathbf{C}^L, \mathbf{a}^{HL})$ .

#### A.4. Step 4: Efficient experimentation by the high type

The objective in [P2] involving the high type's contract is social surplus from the high type. Furthermore, when  $\mathbf{a}^H = (1, \dots, 1)$ , with the sequence having arbitrary finite length, there is obviously a sequence of (sufficiently severe) penalties  $l^H$  to ensure that  $(\text{IC}_a^H)$  is satisfied. It follows that we can take  $\mathbf{a}^H = (1, \dots, 1)$  in an optimal contract  $\mathbf{C}^H$ , where the number of periods of work is  $t^H$ .

## B. PROOF OF THEOREM 3

We remind the reader that Subsection 5.2 provides an outline and intuition for this proof. Without loss by Proposition 1, we focus on penalty contracts throughout the proof. In this Appendix, we will introduce programs and constraints that have analogies with those used in Appendix A. Accordingly, we often use the same labels for equations as before, but the reader should bear in mind that all references in this Appendix to such equations are to those defined in this Appendix.

### B.1. Step 1: The principal's program

By Step 1 and Step 2 in the proof of Theorem 2, we work with the principal's program [P1]. Recall that in this program, without loss, type  $L$  works in all periods  $t \in \Gamma^L$  and constraints  $(\text{IC}^{LH})$  and  $(\text{IR}^H)$  of program [P] are ignored. In this step, we relax the principal's program by considering a weak version of  $(\text{IC}^{HL})$  in which type  $H$  is assumed to exert effort in all periods  $t \in \Gamma^L$  if he chooses  $\mathbf{C}^L$ . The relaxed program, [RP1], is therefore:

$$\max_{(\mathbf{C}^H \in \mathcal{C}, \mathbf{C}^L \in \mathcal{C}, \mathbf{a}^H)} \mu_0 \Pi_0^H(\mathbf{C}^H, \mathbf{a}^H) + (1 - \mu_0) \Pi_0^L(\mathbf{C}^L, \mathbf{1}) \quad (\text{RP1})$$

subject to

$$\mathbf{1} \in \boldsymbol{\alpha}^L(\mathbf{C}^L), \quad (\text{IC}_a^L)$$

$$\mathbf{a}^H \in \boldsymbol{\alpha}^H(\mathbf{C}^H), \quad (\text{IC}_a^H)$$

$$U_0^L(\mathbf{C}^L, \mathbf{1}) \geq 0, \quad (\text{IR}^L)$$

$$U_0^H(\mathbf{C}^H, \mathbf{a}^H) \geq U_0^H(\mathbf{C}^L, \mathbf{1}). \quad (\text{Weak-IC}^{HL})$$

By the same arguments as in Step 2 in the proof of Theorem 2, it is clear that in any solution to program [RP1],  $(\text{IR}^L)$  and  $(\text{Weak-IC}^{HL})$  must be binding. Using these two binding constraints and substituting in the formulae from equations

(2.1) and (2.2), we can rewrite the objective function (RP1) as the sum of expected total surplus less type  $H$ 's "information rent", obtaining the following explicit version of the relaxed program which we call [RP2]:

$$\max_{\substack{\mathbf{C}^H \in \mathcal{C} \\ \mathbf{C}^L \in \mathcal{C} \\ \mathbf{a}^H}} \left\{ \begin{aligned} & \mu_0 \left\{ \beta_0 \sum_{t \in \Gamma^H} \delta^t \left[ \prod_{\substack{s \in \Gamma^H \\ s \leq t-1}} (1 - a_s^H \lambda^H) \right] a_t^H (\lambda^H - c) - (1 - \beta_0) \sum_{t \in \Gamma^H} \delta^t a_t^H c \right\} \\ & + (1 - \mu_0) \left\{ \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] (\lambda^L - c) - (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t c \right\} \\ & - \mu_0 \left\{ \begin{aligned} & \beta_0 \sum_{t \in \Gamma^L} \delta^t l_t^L \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t}} (1 - \lambda^H) - \prod_{\substack{s \in \Gamma^L \\ s \leq t}} (1 - \lambda^L) \right] \\ & - \beta_0 \sum_{t \in \Gamma^L} \delta^t c \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^H) - \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] \end{aligned} \right\} \end{aligned} \right\} \quad \text{(RP2)}$$

Information rent of type  $H$

subject to

$$\mathbf{1} \in \arg \max_{(a_t)_{t \in \Gamma^L}} \left\{ \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - a_s \lambda^L) \right] [(1 - a_t \lambda^L) l_t^L - a_t c] + (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t (l_t^L - a_t c) + W_0^L \right\}, \quad \text{(IC}_a^L)$$

$$\mathbf{a}^H \in \arg \max_{(a_t)_{t \in \Gamma^H}} \left\{ \beta_0 \sum_{t \in \Gamma^H} \delta^t \left[ \prod_{\substack{s \in \Gamma^H \\ s \leq t-1}} (1 - a_s \lambda^H) \right] [(1 - a_t \lambda^H) l_t^H - a_t c] + (1 - \beta_0) \sum_{t \in \Gamma^H} \delta^t (l_t^H - a_t c) + W_0^H \right\}. \quad \text{(IC}_a^H)$$

Program [RP2] is separable, *i.e.* it can be solved by maximizing (RP2) with respect to  $\mathbf{C}^L$  subject to (IC<sub>a</sub><sup>L</sup>) and separately maximizing (RP2) with respect to  $(\mathbf{C}^H, \mathbf{a}^H)$  subject to (IC<sub>a</sub><sup>H</sup>).

### B.2. Step 2: Connected contracts for the low type

We claim that in program [RP2], it is without loss to consider solutions in which the low type's contract is a connected penalty contract, *i.e.* solutions  $\mathbf{C}^L$  in which  $\Gamma^L = \{1, \dots, T^L\}$  for some  $T^L$ .

To prove this, observe that the optimal  $\mathbf{C}^L$  is a solution of

$$\max_{\mathbf{C}^L} \left\{ \begin{aligned} & (1 - \mu_0) \left\{ \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] (\lambda^L - c) - (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t c \right\} \\ & - \mu_0 \beta_0 \left\{ \sum_{t \in \Gamma^L} \delta^t l_t^L \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t}} (1 - \lambda^H) - \prod_{\substack{s \in \Gamma^L \\ s \leq t}} (1 - \lambda^L) \right] - \sum_{t \in \Gamma^L} \delta^t c \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^H) - \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] \right\} \end{aligned} \right\} \quad \text{(B.1)}$$

subject to (IC<sub>a</sub><sup>L</sup>) and

$$\mathbf{1} \in \arg \max_{(a_t)_{t \in \Gamma^L}} \left\{ \beta_0 \sum_{t \in \Gamma^L} \delta^t \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - a_s \lambda^L) \right] [(1 - a_t \lambda^L) l_t^L - a_t c] + (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t (l_t^L - a_t c) + W_0^L \right\}. \quad \text{(B.2)}$$

To avoid trivialities, consider any optimal  $\mathbf{C}^L$  with  $\Gamma^L \neq \emptyset$ . First consider the possibility that  $1 \notin \Gamma^L$ . In this case, construct a new penalty contract  $\widehat{\mathbf{C}}^L$  that is "shifted up by one period":

$$\begin{aligned} \widehat{\Gamma}^L &= \{s : s + 1 \in \Gamma^L\}, \\ \widehat{l}_s^L &= l_{s+1}^L \text{ for all } s \in \widehat{\Gamma}^L, \\ \widehat{W}_0^L &= W_0^L. \end{aligned}$$

Clearly it remains optimal for the agent to work in every period in  $\widehat{\Gamma}^L$ , and since the value of (B.1) must have been weakly positive under  $\mathbf{C}^L$ , it is now weakly higher since the modification has just multiplied it by  $\delta^{-1} > 1$ . This procedure can

be repeated for all lockout periods at the beginning of the contract, so that without loss, we hereafter assume that  $1 \in \Gamma^L$ . We are of course done if  $\Gamma^L$  is now connected, so also assume that  $\Gamma^L$  is not connected.

Let  $t^\circ$  be the earliest lockout period in  $\mathbf{C}^L$ , i.e.  $t^\circ = \min\{t: t \notin \Gamma^L \text{ and } t-1 \in \Gamma^L\}$ . (Such a  $t^\circ > 1$  exists given the preceding discussion.) We will argue that one of two possible modifications preserves the agent's incentive to work in all periods in the modified contract and weakly improves the principal's payoff. This suffices because the procedure can then be applied iteratively to produce a connected contract.

Modification 1: Consider first a modified penalty contract  $\widehat{\mathbf{C}}^L$  that removes the lockout period  $t^\circ$  and shortens the contract by one period as follows:

$$\begin{aligned}\widehat{\Gamma}^L &= \{1, \dots, t^\circ - 1\} \cup \{s: s \geq t^\circ \text{ and } s+1 \in \Gamma^L\}, \\ \widehat{l}_s^L &= \begin{cases} l_s^L & \text{if } s < t^\circ - 1, \\ l_s^L + \Delta_1 & \text{if } s = t^\circ - 1, \\ l_{s+1}^L & \text{if } s \geq t^\circ \text{ and } s \in \widehat{\Gamma}^L, \end{cases} \\ \widehat{W}_0^L &= W_0^L.\end{aligned}$$

Note that in the above construction,  $\Delta_1$  is a free parameter. We will find conditions on  $\Delta_1$  such that type  $L$ 's incentives for effort are unchanged and the principal is weakly better off.

For an arbitrary  $t$ , define

$$\begin{aligned}S(t) &= (\lambda^L - c) \prod_{s \in \Gamma^L, s \leq t-1} (1 - \lambda^L), \\ R(t) &= \prod_{s \in \Gamma^L, s \leq t} (1 - \lambda^H) - \prod_{s \in \Gamma^L, s \leq t} (1 - \lambda^L).\end{aligned}$$

The value of (B.1) under  $\mathbf{C}^L$  is

$$V(\mathbf{C}^L) = (1 - \mu_0) \left[ \beta_0 \sum_{t \in \Gamma^L} \delta^t S(t) - (1 - \beta_0) \sum_{t \in \Gamma^L} \delta^t c \right] - \mu_0 \beta_0 \left[ \sum_{t \in \Gamma^L} \delta^t l_t^L R(t) - \sum_{t \in \Gamma^L} \delta^t c R(t-1) \right]. \quad (\text{B.3})$$

The value of (B.1) after the modification to  $\widehat{\mathbf{C}}^L$  is

$$\begin{aligned}V(\widehat{\mathbf{C}}^L) &= (1 - \mu_0) \left[ \beta_0 \left( \sum_{\substack{t \in \Gamma^L \\ t < t^\circ}} \delta^t S(t) + \delta^{-1} \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t S(t) \right) - (1 - \beta_0) \left( \sum_{\substack{t \in \Gamma^L \\ t < t^\circ}} \delta^t c + \delta^{-1} \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t c \right) \right] \\ &\quad - \mu_0 \beta_0 \left[ \sum_{\substack{t \in \Gamma^L \\ t < t^\circ - 1}} \delta^t l_t^L R(t) + \delta^{-1} \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t l_t^L R(t) + \delta^{t^\circ - 1} (l_{t^\circ - 1}^L + \Delta_1) R(t^\circ - 1) \right] \\ &\quad - \sum_{\substack{t \in \Gamma^L \\ t < t^\circ}} \delta^t c R(t-1) - \delta^{-1} \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t c R(t-1).\end{aligned}$$

Therefore, the modification benefits the principal if and only if

$$\begin{aligned}0 \leq V(\widehat{\mathbf{C}}^L) - V(\mathbf{C}^L) &= (1 - \mu_0) \left[ \beta_0 (\delta^{-1} - 1) \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t S(t) - (1 - \beta_0) (\delta^{-1} - 1) \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t c \right] \\ &\quad - \mu_0 \beta_0 \left[ (\delta^{-1} - 1) \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t l_t^L R(t) + \delta^{t^\circ - 1} \Delta_1 R(t^\circ - 1) - (\delta^{-1} - 1) \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t c R(t-1) \right].\end{aligned}$$

The above inequality is satisfied for any  $\Delta_1$  if  $\delta = 1$ , and if  $\delta < 1$ , then after rearranging terms, the above inequality is equivalent to

$$\begin{aligned} & (1 - \mu_0) \left[ \beta_0 \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t S(t) - (1 - \beta_0) \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t c \right] \\ & \geq \mu_0 \beta_0 \left[ \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t l_t^L R(t) - \frac{\delta^{t^\circ - 1} \Delta_1 R(t^\circ - 1)}{1 - \delta^{-1}} - \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t c R(t - 1) \right]. \end{aligned} \tag{B.4}$$

Now turn to the incentives for effort for the agent of type  $L$ . Clearly, since  $C^L$  induces the agent to work in all periods, it remains optimal for the agent to work under  $\widehat{C}^L$  in all periods beginning with  $t^\circ$ . Consider the incentive constraint for effort in period  $t^\circ - 1$  under  $\widehat{C}^L$ . Using (B.2), this is given by:

$$-\bar{\beta}_{t^\circ - 1}^L \lambda^L \left\{ l_{t^\circ - 1}^L + \Delta_1 + \delta^{-1} \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^{t - (t^\circ - 1)} \left[ \prod_{\substack{s \in \Gamma^L \\ t^\circ - 1 < s \leq t - 1}} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - c] \right\} \geq c. \tag{B.5}$$

Analogously, the incentive constraint in period  $t^\circ - 1$  under the original contract  $C^L$  is:

$$-\bar{\beta}_{t^\circ - 1}^L \lambda^L \left\{ l_{t^\circ - 1}^L + \sum_{\substack{t \in \Gamma^L \\ t > t^\circ - 1}} \delta^{t - (t^\circ - 1)} \left[ \prod_{\substack{s \in \Gamma^L \\ t^\circ - 1 < s \leq t - 1}} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - c] \right\} \geq c. \tag{B.6}$$

If we choose  $\Delta_1$  such that the left-hand side of (B.5) is equal to the left-hand side of (B.6), then since it is optimal to work under the original contract in period  $t^\circ - 1$ , it will also be optimal to work under the new contract in period  $t^\circ - 1$ . Accordingly, we choose  $\Delta_1$  such that:

$$\begin{aligned} \Delta_1 &= \sum_{t \in \Gamma^L, t > t^\circ - 1} \delta^{t - (t^\circ - 1)} \left[ \prod_{s \in \Gamma^L, t^\circ - 1 < s \leq t - 1} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - c] \\ &\quad - \delta^{-1} \sum_{t \in \Gamma^L, t > t^\circ} \delta^{t - (t^\circ - 1)} \left[ \prod_{s \in \Gamma^L, t^\circ - 1 < s \leq t - 1} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - c] \\ &= (1 - \delta^{-1}) \sum_{t \in \Gamma^L, t > t^\circ - 1} \delta^{t - (t^\circ - 1)} \left[ \prod_{s \in \Gamma^L, t^\circ - 1 < s \leq t - 1} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - c], \end{aligned} \tag{B.7}$$

where the second equality is because  $\{t : t \in \Gamma^L, t > t^\circ - 1\} = \{t : t \in \Gamma^L, t > t^\circ\}$ , since  $t^\circ \notin \Gamma^L$ . Note that (B.7) implies  $\Delta_1 = 0$  if  $\delta = 1$ .

Now consider the incentive constraint for effort in any period  $\tau < t^\circ - 1$ . We will show that because  $\Delta_1$  is such that the left-hand side of (B.5) is equal to the left-hand side of (B.6), the fact that it was optimal to work in period  $\tau$  under contract  $C^L$  implies that it is optimal to work in period  $\tau$  under contract  $\widehat{C}^L$ . Formally, the incentive constraint for effort in period  $\tau$  under  $C^L$  is

$$-\bar{\beta}_\tau^L \lambda^L \left\{ l_\tau^L + \sum_{t \in \Gamma^L, t > \tau} \delta^{t - \tau} \left[ \prod_{s \in \Gamma^L, \tau < s \leq t - 1} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - c] \right\} \geq c, \tag{B.8}$$

which is satisfied since  $C^L$  induces the agent to work in all periods. Analogously, the incentive constraint for effort in period  $\tau$  under  $\widehat{C}^L$  can be written as

$$-\bar{\beta}_\tau^L \lambda^L \left\{ \widehat{l}_\tau^L + \sum_{t \in \widehat{\Gamma}^L, t > \tau} \delta^{t - \tau} \left[ \prod_{s \in \widehat{\Gamma}^L, \tau < s \leq t - 1} (1 - \lambda^L) \right] [(1 - \lambda^L) \widehat{l}_t^L - c] \right\} \geq c.$$

Algebraic simplification using the definition of  $\widehat{C}^L$  and equation (B.7) shows that this constraint is identical to (B.8), and hence is satisfied.

Thus, if  $\delta = 1$ , this modification with  $\Delta_1 = 0$  weakly benefits the principal while preserving the agent's incentives, and we are done. So hereafter assume  $\delta < 1$ , which requires us to also consider another modification.

Modification 2: Now we consider a modified contract  $\tilde{\mathbf{C}}^L$  that eliminates all periods after  $t^\circ$ , defined as follows:

$$\tilde{W}_0^L = W_0^L, \quad \tilde{\Gamma}^L = \{1, \dots, t^\circ - 1\}, \quad \tilde{l}_s^L = \begin{cases} l_s^L & \text{if } s < t^\circ - 1, \\ l_s^L + \Delta_2 & \text{if } s = t^\circ - 1. \end{cases}$$

Again,  $\Delta_2$  is a free parameter above. We now find conditions on  $\Delta_2$  such that type  $L$ 's incentives are unchanged and the principal is weakly better off.

The value of (B.1) under the modification  $\tilde{\mathbf{C}}^L$  is

$$V(\tilde{\mathbf{C}}^L) = (1 - \mu_0) \left[ \beta_0 \sum_{t \in \Gamma^L, t < t^\circ} \delta^t S(t) - (1 - \beta_0) \sum_{t \in \Gamma^L, t < t^\circ} \delta^t c \right] - \mu_0 \beta_0 \left[ \sum_{t \in \Gamma^L, t < t^\circ - 1} \delta^t l_t^L R(t) + \delta^{t^\circ - 1} (l_{t^\circ - 1}^L + \Delta_2) R(t^\circ - 1) - \sum_{t \in \Gamma^L, t < t^\circ} \delta^t c R(t - 1) \right].$$

Therefore, recalling (B.3), this modification benefits the principal if and only if

$$0 \leq V(\tilde{\mathbf{C}}^L) - V(\mathbf{C}^L) = -(1 - \mu_0) \left[ \beta_0 \sum_{t \in \Gamma^L, t > t^\circ} \delta^t S(t) - (1 - \beta_0) \sum_{t \in \Gamma^L, t > t^\circ} \delta^t c \right] - \mu_0 \beta_0 \left[ - \sum_{t \in \Gamma^L, t > t^\circ} \delta^t l_t^L R(t) + \delta^{t^\circ - 1} \Delta_2 R(t^\circ - 1) + \sum_{t \in \Gamma^L, t > t^\circ} \delta^t c R(t - 1) \right],$$

or equivalently after rearranging terms, if and only if

$$(1 - \mu_0) \left[ \beta_0 \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t S(t) - (1 - \beta_0) \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t c \right] \leq \mu_0 \beta_0 \left[ \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t l_t^L R(t) - \delta^{t^\circ - 1} \Delta_2 R(t^\circ - 1) - \sum_{\substack{t \in \Gamma^L \\ t > t^\circ}} \delta^t c R(t - 1) \right]. \tag{B.9}$$

As with the previous modification, the only incentive constraint for effort that needs to be verified in  $\tilde{\mathbf{C}}^L$  is that of period  $t^\circ - 1$ , which since it is the last period of the contract is simply:

$$-\bar{\beta}_{t^\circ - 1}^L \lambda^L (l_{t^\circ - 1}^L + \Delta_2) \geq c. \tag{B.10}$$

We choose  $\Delta_2$  so that the left-hand side of (B.10) is equal to the left-hand side of (B.6):

$$\Delta_2 = \sum_{t \in \Gamma^L, t > t^\circ - 1} \delta^{t - (t^\circ - 1)} \left[ \prod_{s \in \Gamma^L, t^\circ - 1 < s \leq t - 1} (1 - \lambda^L) \right] [(1 - \lambda^L) l_t^L - c] = \frac{\Delta_1}{1 - \delta^{-1}}, \tag{B.11}$$

where the second equality follows from (B.7). But now, observe that (B.11) implies that either (B.4) or (B.9) is guaranteed to hold, and hence either the modification to  $\tilde{\mathbf{C}}^L$  or to  $\tilde{\mathbf{C}}^L$  weakly benefits the principal while preserving the agent's effort incentives.

**Remark 3.** Given  $\delta < 1$ , the choice of  $\Delta_2$  in (B.11) implies that if inequality (B.4) holds with equality then so does inequality (B.9), and vice-versa. In other words, if neither of the modifications strictly benefits the principal (while preserving the agent's effort incentives), then it must be that both modifications leave the principal's payoff unchanged (while preserving the agent's effort incentives).

### B.3. Step 3: Defining the critical contract for the low type

Take any connected penalty contract  $\mathbf{C}^L = (T^L, W_0^L, l^L)$  that induces effort from the low type in each period  $t \in \{1, \dots, T^L\}$ . We claim that the low type's incentive constraint for effort binds at all periods if and only if  $l^L = \bar{l}^L(T^L)$ , where  $\bar{l}^L(T^L)$  is defined as follows:

$$\bar{l}_t^L = \begin{cases} -(1 - \delta) \frac{c}{\bar{\beta}_t^L \lambda^L} & \text{if } t < T^L, \\ -\frac{c}{\bar{\beta}_{T^L}^L \lambda^L} & \text{if } t = T^L. \end{cases} \tag{B.12}$$

The proof of this claim is via three sub-steps; for the remainder of this step, since  $T^L$  is given and held fixed, we ease notation by just writing  $\bar{l}^L$  instead of  $\bar{l}^L(T^L)$ .

**Step 3a:** First, we argue that with the above penalty sequence, the low type is indifferent between working and shirking in each period  $t \in \{1, \dots, T^L\}$  given that he has worked in all prior periods and will do in all subsequent periods no matter his action at period  $t$ . In other words, we need to show that for all  $t \in \{1, \dots, T^L\}$ :

$$-\bar{\beta}_t^L \lambda^L \left\{ \bar{l}_t^L + \sum_{s=t+1}^{T^L} \delta^{s-t} (1-\lambda^L)^{s-(t+1)} \left[ (1-\lambda^L) \bar{l}_s^L - c \right] \right\} = c. \tag{B.13}$$

We prove that (5.13) is indeed satisfied for all  $t$  by induction. First, it is immediate from (B.12) that (5.13) holds for  $t = T^L$ . Next, for any  $t < T^L$ , assume (5.13) holds for  $t + 1$ . This is equivalent to

$$\sum_{s=t+2}^{T^L} \delta^{s-(t+1)} (1-\lambda^L)^{s-(t+2)} \left[ (1-\lambda^L) \bar{l}_s^L - c \right] = -\frac{c}{\bar{\beta}_{t+1}^L \lambda^L} - \bar{l}_{t+1}^L. \tag{B.14}$$

To show that (5.13) holds for  $t$ , it suffices to show that

$$-\bar{\beta}_t^L \lambda^L \left\{ \bar{l}_t^L + \delta \left[ (1-\lambda^L) \bar{l}_{t+1}^L - c \right] + \delta (1-\lambda^L) \sum_{s=t+2}^{T^L} \delta^{s-(t+1)} (1-\lambda^L)^{s-(t+2)} \left[ (1-\lambda^L) \bar{l}_s^L - c \right] \right\} = c.$$

Using (B.14), the above equality is equivalent to

$$-\bar{\beta}_t^L \lambda^L \left\{ \bar{l}_t^L + \delta \left[ (1-\lambda^L) \bar{l}_{t+1}^L - c \right] + \delta (1-\lambda^L) \left[ -\frac{c}{\bar{\beta}_{t+1}^L \lambda^L} - \bar{l}_{t+1}^L \right] \right\} = c,$$

which simplifies to

$$\bar{l}_t^L = -\frac{c}{\bar{\beta}_t^L \lambda^L} + \delta c + \delta (1-\lambda^L) \frac{c}{\bar{\beta}_{t+1}^L \lambda^L}. \tag{B.15}$$

Since  $\bar{\beta}_{t+1}^L = \frac{\bar{\beta}_t^L (1-\lambda^L)}{1-\bar{\beta}_t^L \lambda^L}$ , (B.15) is in turn equivalent to  $\bar{l}_t^L = -(1-\delta) \frac{c}{\bar{\beta}_t^L \lambda^L}$ , which is true by the definition of  $\bar{l}^L$  in (B.12).

**Step 3b:** Next, we show that given the sequence  $\bar{l}^L$ , it would be optimal for the low type to work in any period no matter the prior history of effort. Consider first the last period,  $T^L$ . No matter the history of prior effort, the current belief is some  $\beta_{T^L}^L \geq \bar{\beta}_{T^L}^L$ , hence  $-\beta_{T^L}^L \lambda^L \bar{l}_t^L \geq \bar{\beta}_{T^L}^L \lambda^L \bar{l}_t^L = c$  (where the equality is by definition), so that it is optimal to work in  $T^L$ .

Now assume inductively that the assertion is true for period  $t + 1 \leq T^L$ , and consider period  $t < T^L$  after any history of prior effort, with current belief  $\beta_t^L$ . Since we already showed that equation (5.13) holds, it follows from  $\beta_t^L \geq \bar{\beta}_t^L$  that

$$-\beta_t^L \lambda^L \left\{ \bar{l}_t^L + \sum_{s=t+1}^{T^L} \delta^{s-t} (1-\lambda^L)^{s-(t+1)} \left[ (1-\lambda^L) \bar{l}_s^L - c \right] \right\} \geq c,$$

and hence it is optimal for the agent to work in period  $t$ .

**Step 3c:** Finally, we argue that any profile of penalties,  $l^L$ , that makes the low type's incentive constraint for effort bind at every period  $t \in \{1, \dots, T^L\}$  must coincide with  $\bar{l}^L$ , given that the penalty contract must induce work from the low type in each period up to  $T^L$ . Again, we use induction. Since  $\bar{l}_{T^L}^L$  is the unique penalty that makes the agent indifferent

53. To derive this equality, observe that under the hypotheses, the payoff for type  $L$  from working at  $t$  is

$$-c + (1-\bar{\beta}_t^L) \bar{l}_t^L + (1-\bar{\beta}_t^L) \sum_{s=t+1}^{T^L} \delta^{s-t} (\bar{l}_s^L - c) + \bar{\beta}_t^L \left\{ (1-\lambda^L) \bar{l}_t^L + \sum_{s=t+1}^{T^L} \delta^{s-t} (1-\lambda^L)^{s-t} \left[ (1-\lambda^L) \bar{l}_s^L - c \right] \right\},$$

while the payoff from shirking at time  $t$  is

$$\bar{l}_t^L + (1-\bar{\beta}_t^L) \sum_{s=t+1}^{T^L} \delta^{s-t} (\bar{l}_s^L - c) + \bar{\beta}_t^L \left\{ \sum_{s=t+1}^{T^L} \delta^{s-t} (1-\lambda^L)^{s-(t+1)} \left[ (1-\lambda^L) \bar{l}_s^L - c \right] \right\}.$$

Setting these payoffs from working and shirking equal to each other and manipulating terms yields (5.13).

between working and shirking at period  $T^L$  given that he has worked in all prior periods, it follows that  $l_{T^L}^L = \bar{l}_{T^L}^L$ . Note from Step 3b that it would remain optimal for the agent to work in period  $T^L$  given any profile of effort in prior periods.

For the inductive step, pick some period  $t < T^L$  and assume that in every period  $x \in \{t, \dots, T^L\}$ , the agent is indifferent between working and shirking given that he has worked in all prior periods, and would also find it optimal to work at  $x$  following any other profile of effort prior to  $x$ . Under these hypotheses, the indifference at period  $t+1$  implies that

$$-\bar{\beta}_{t+1}^L \lambda^L \left\{ l_{t+1}^L + \sum_{s=t+2}^{T^L} \delta^{s-(t+1)} (1-\lambda^L)^{s-(t+2)} [(1-\lambda^L)l_s^L - c] \right\} = c. \quad (\text{B.16})$$

Given the inductive hypothesis, the incentive constraint for effort at period  $t$  is

$$-\bar{\beta}_t^L \lambda^L \left\{ l_t^L + \sum_{s=t+1}^{T^L} \delta^{s-t} (1-\lambda^L)^{s-(t+1)} [(1-\lambda^L)l_s^L - c] \right\} \geq c,$$

which, when set to bind, can be written as

$$-\bar{\beta}_t^L \lambda^L \left\{ l_t^L + \delta [(1-\lambda^L)l_{t+1}^L - c] + \delta (1-\lambda^L) \sum_{s=t+2}^{T^L} \delta^{s-(t+1)} (1-\lambda^L)^{s-(t+2)} [(1-\lambda^L)l_s^L - c] \right\} = c. \quad (\text{B.17})$$

Substituting (B.16) into (B.17), using the fact that  $\bar{\beta}_{t+1}^L = \frac{\bar{\beta}_t^L (1-\lambda)}{\bar{\beta}_t^L (1-\lambda) + 1 - \bar{\beta}_t^L}$ , and performing some algebra shows that  $l_t^L = \bar{l}_t^L$ . Moreover, by the reasoning in Step 3b, this also ensures that the agent would find it optimal to work in period  $t$  for any other history of actions prior to period  $t$ .

#### B.4. Step 4: The critical contract is optimal

By Step 2, we can restrict attention in solving program [RP2] to connected penalty contracts for the low type. For any  $T^L$ , Step 3 identified a particular sequence of penalties,  $\bar{l}^L(T^L)$ . We now show that any connected penalty contract for the low type that solves [RP2] must have precisely this penalty structure.

The proof involves two sub-steps; throughout, we hold an arbitrary  $T^L$  fixed and, to ease notation, drop the dependence of  $\bar{l}^L(\cdot)$  on  $T^L$ .

**Step 4a:** We first show that any connected penalty contract for the low type of length  $T^L$  that satisfies  $(IC_a^L)$  and has  $l_t^L > \bar{l}_t^L$  in some period  $t \leq T^L$  is not optimal. To prove this, consider any such connected penalty contract. Define

$$\hat{t} = \max \{ t : t \leq T^L \text{ and } l_t^L > \bar{l}_t^L \}.$$

Observe that we must have  $\hat{t} < T^L$  because otherwise  $(IC_a^L)$  would be violated in period  $T^L$ . Furthermore, by definition of  $\hat{t}$ ,  $l_t^L \leq \bar{l}_t^L$  for all  $T^L \geq t > \hat{t}$ . We will prove that we can change the penalty structure by lowering  $l_{\hat{t}}^L$  and raising some subsequent  $l_s^L$  for  $s \in \{\hat{t}+1, \dots, T^L\}$  in a way that keeps type  $L$ 's incentives for effort unchanged, and yet increase the value of the objective function (RP2).

*Claim:* There exists  $\tilde{t} \in \{\hat{t}+1, \dots, T^L\}$  such that  $(IC_a^L)$  at  $\tilde{t}$  is slack and  $l_{\tilde{t}}^L < \bar{l}_{\tilde{t}}^L$ .

*Proof:* Suppose not, then for each  $T^L \geq t > \hat{t}$ , either  $l_t^L = \bar{l}_t^L$ , or  $l_t^L < \bar{l}_t^L$  and  $(IC_a^L)$  binds. Then since whenever  $l_t^L < \bar{l}_t^L$ ,  $(IC_a^L)$  binds by supposition, it must be that in all  $t > \hat{t}$ ,  $(IC_a^L)$  binds (this follows from Step 3). But then  $(IC_a^L)$  at  $\hat{t}$  is violated since  $l_{\hat{t}}^L > \bar{l}_{\hat{t}}^L$ .  $\parallel$

*Claim:* There exists  $\bar{t} \in \{\hat{t}+1, \dots, T^L\}$  such that  $l_{\bar{t}}^L < \bar{l}_{\bar{t}}^L$  and for any  $t \in \{\hat{t}+1, \dots, \bar{t}\}$ ,  $(IC_a^L)$  at  $t$  is slack. In particular, we can take  $\bar{t}$  to be the first such period after  $\hat{t}$ .

*Proof:* Fix  $\bar{t}$  in the previous claim. Note that  $(IC_a^L)$  at  $\hat{t}+1$  must be slack because otherwise  $(IC_a^L)$  at  $\hat{t}$  is violated by  $l_{\hat{t}}^L > \bar{l}_{\hat{t}}^L$  and Step 3. There are two cases. (1)  $l_{\hat{t}+1}^L < \bar{l}_{\hat{t}+1}^L$ ; then  $\hat{t}+1$  is the  $\bar{t}$  we want. (2)  $l_{\hat{t}+1}^L = \bar{l}_{\hat{t}+1}^L$ ; in this case, since  $(IC_a^L)$  is slack at  $\hat{t}+1$ , it must be that  $(IC_a^L)$  at  $\hat{t}+2$  is slack (otherwise, the claim in Step 3 is violated); now if  $l_{\hat{t}+2}^L < \bar{l}_{\hat{t}+2}^L$ , we are done because  $\hat{t}+2$  is the  $\bar{t}$  we are looking for; if  $l_{\hat{t}+2}^L = \bar{l}_{\hat{t}+2}^L$ , then we continue to  $\hat{t}+3$  and so on until we reach  $\bar{t}$  which we know gives us a slack  $(IC_a^L)$ ,  $l_{\bar{t}}^L < \bar{l}_{\bar{t}}^L$ , and we are sure that  $(IC_a^L)$  is slack in all periods of this process before reaching  $\bar{t}$ .  $\parallel$

Now we shall show that we can slightly reduce  $l_t^L > \bar{l}_t^L$  and slightly increase  $l_t^L < \bar{l}_t^L$  and meanwhile keep the incentives for effort of type  $L$  satisfied for all periods. By the same reasoning as used in Step 2, the incentive constraint for effort in period  $\hat{t}$  (given that the agent will work in all subsequent periods no matter his behaviour at period  $t$ ) can be written as

$$-\beta_t^L \lambda^L \left\{ l_t^L + \sum_{t>\hat{t}} \delta^{t-\hat{t}} (1-\lambda^L)^{t-(\hat{t}+1)} [(1-\lambda^L)l_t^L - c] \right\} \geq c. \tag{B.18}$$

Observe that if we reduce  $l_t^L$  by  $\Delta > 0$  and increase  $l_{\bar{t}}^L$  by  $\frac{\Delta}{\delta^{\bar{t}-\hat{t}}(1-\lambda^L)^{\bar{t}-(\hat{t}+1)}(1-\lambda^L)} = \frac{\Delta}{\delta^{\bar{t}-\hat{t}}(1-\lambda^L)^{\bar{t}-\hat{t}}}$ , then the left-hand side of (B.18) does not change. Moreover, it follows that incentives for effort at  $t < \bar{t}$  are also unchanged (see Step 2), and the incentive condition at  $\bar{t}$  will be satisfied if  $\Delta$  is small enough because the original  $(IC_a^L)$  at  $\bar{t}$  is slack.

Finally, we show that the modification above leads to a reduction of the rent of type  $H$  in (RP2), i.e. raises the value of the objective. The rent is given by

$$\mu_0 \beta_0 \left\{ \sum_{t=1}^{T^L} \delta^t l_t^L [(1-\lambda^H)^t - (1-\lambda^L)^t] - \sum_{t=1}^{T^L} \delta^t c [(1-\lambda^H)^{t-1} - (1-\lambda^L)^{t-1}] \right\}. \tag{B.19}$$

Hence, the change in the rent from reducing  $l_t^L$  by  $\Delta$  and increasing  $l_{\bar{t}}^L$  by  $\frac{\Delta}{\delta^{\bar{t}-\hat{t}}(1-\lambda^L)^{\bar{t}-\hat{t}}}$  is

$$\begin{aligned} & \mu_0 \beta_0 \delta^{\hat{t}} \Delta \left\{ - \left[ (1-\lambda^H)^{\hat{t}} - (1-\lambda^L)^{\hat{t}} \right] + \frac{1}{(1-\lambda^L)^{\bar{t}-\hat{t}}} \left[ (1-\lambda^H)^{\bar{t}} - (1-\lambda^L)^{\bar{t}} \right] \right\} \\ & = \mu_0 \beta_0 \delta^{\hat{t}} \Delta \frac{(1-\lambda^H)^{\hat{t}}}{(1-\lambda^L)^{\bar{t}-\hat{t}}} \left[ (1-\lambda^H)^{\bar{t}-\hat{t}} - (1-\lambda^L)^{\bar{t}-\hat{t}} \right] < 0, \end{aligned}$$

where the inequality is because  $\bar{t} > \hat{t}$  and  $1-\lambda^H < 1-\lambda^L$ .

**Step 4b:** By Step 4a, we can restrict attention to penalty sequences  $l^L$  such that  $l_t^L \leq \bar{l}_t^L$  for all  $t \leq T^L$ . Now we show that unless  $l^L(\cdot) = \bar{l}^L(\cdot)$ , the value of the objective (RP2) can be improved while satisfying the incentive constraint for effort,  $(IC_a^L)$ . Recall that by Step 4a,  $(IC_a^L)$  is satisfied in all periods  $t = 1, \dots, T^L$  whenever  $l_t^L = \bar{l}_t^L$ . Thus, if  $l_t^L < \bar{l}_t^L$  for any period, we can replace  $l_t^L$  by  $\bar{l}_t^L$  without affecting the effort incentives for type  $L$ . Moreover, by doing this we reduce the rent of type  $H$ , given by (B.19) above, and thus raise the value of (RP2).

### B.5. Step 5: Generic uniqueness of the optimal contract for the low type

By Step 4, an optimal contract for the low type that solves program [RP2] can be found by optimizing over  $T^L$ , i.e. the length of connected penalty contracts with the penalty structure  $\bar{l}^L(T^L)$ . By Theorem 2,  $T^L \leq t^L$ . In this step, we establish generic uniqueness of the optimal contract for the low type. We proceed in two sub-steps.

**Step 5a:** First, we show that the optimal length  $T^L$  of connected penalty contracts with the penalty structure  $\bar{l}^L(T^L)$  is generically unique. The portion of the objective (RP2) that involves  $T^L$  is

$$\begin{aligned} V(T^L) := & (1-\mu_0) \left[ \beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1} (\lambda^L - c) - (1-\beta_0) \sum_{t=1}^{T^L} \delta^t c \right] \\ & - \mu_0 \beta_0 \left\{ \sum_{t=1}^{T^L} \delta^t \bar{l}_t^L(T^L) [(1-\lambda^H)^t - (1-\lambda^L)^t] - \sum_{t=1}^{T^L} \delta^t c [(1-\lambda^H)^{t-1} - (1-\lambda^L)^{t-1}] \right\}, \tag{B.20} \end{aligned}$$

where we have used the desired penalty sequence. Note that by Theorem 2  $T^L \leq t^L$  in any optimal contract for the low type and hence there is a finite number of maximizers of  $V(T^L)$ . It follows that if we perturb  $\mu_0$  locally, the set of maximizers will not change. Now suppose that the maximizer of  $V(T^L)$  is not unique. Without loss, pick any two maximizers  $\tilde{T}^L$  and  $\hat{T}^L$ . We must have  $V(\tilde{T}^L) = V(\hat{T}^L)$  and (again by Theorem 2)  $t^L \geq \max\{\tilde{T}^L, \hat{T}^L\}$ . Without loss, assume  $\hat{T}^L > \tilde{T}^L$ . Note that the first term in square brackets in (B.20) is social surplus from the low type and hence it is strictly increasing in  $T^L$  for  $T^L < t^L$ . Therefore, both the first and second terms in  $V(\hat{T}^L)$  must be larger than the first and second terms in  $V(\tilde{T}^L)$  respectively. But then it is immediate that perturbing  $\mu_0$  within an arbitrarily small neighborhood will change the ranking of  $V(\tilde{T}^L)$  and  $V(\hat{T}^L)$ , which implies that the assumed multiplicity is non-generic.



It follows that there is generically a unique  $T^L$  that maximizes  $V(T^L)$ ; hereafter we denote this solution  $\bar{T}^L$ . In the non-generic cases in which multiple maximizers exist, we select the largest one.

**Step 5b:** In Step 5a we showed that among connected penalty contracts, there is generically a unique contract for type  $L$  that solves [RP2]. We now claim that there generically cannot be any other penalty contract for type  $L$  that solves [RP2]. Suppose, to contradiction, that this is false: there is an optimal non-connected penalty contract  $\mathbf{C}^L = (\Gamma^L, W_0^L, t^L)$  in which  $\mathbf{1} \in \alpha^L(\mathbf{C}^L)$ . Let  $t^\circ < \max \Gamma^L$  be the earliest lockout period in  $\mathbf{C}^L$ . Without loss, owing to genericity, we take  $\delta < 1$ . Following the arguments of Step 2, in particular Remark 3, the optimality of  $\mathbf{C}^L$  implies that there are two *connected* penalty contracts that are also optimal:  $\hat{\mathbf{C}}^L = (\hat{T}^L, \hat{W}_0^L, \hat{T}^L)$  obtained from  $\mathbf{C}^L$  by applying Modification 1 of Step 2 as many times as needed to eliminate all lockout periods, and  $\tilde{\mathbf{C}}^L = (\tilde{T}^L, \tilde{W}_0^L, \tilde{T}^L)$  obtained from  $\mathbf{C}^L$  by applying Modification 2 of Step 2 to shorten the contract by just eliminating all periods from  $t^\circ$  on. Note that the modifications ensure that  $\mathbf{1} \in \alpha^L(\hat{\mathbf{C}}^L)$  and  $\mathbf{1} \in \alpha^L(\tilde{\mathbf{C}}^L)$ . But now, the fact that  $\hat{T}^L > \tilde{T}^L$  contradicts the generic uniqueness of connected penalty contracts for the low type that solve [RP2].

### B.6. Step 6: Back to the original program

We have shown so far that there is a solution to program [RP2] in which the low type's contract is a connected penalty contract of length  $\bar{T}^L \leq t^L$  and in which the penalty sequence is given by  $\bar{T}^L(\bar{T}^L)$ . In terms of optimizing over the high type's contract, note that, as shown in Theorem 2, we can take the solution as inducing the high type to work in each period up to  $t^H$  and no longer: this follows from the fact that the portion of the objective in (RP2) involving the high type's contract is social surplus from the high type.

Recall that solutions to [RP2] produce solutions to [RP1] by choosing  $W_0^L$  to make  $(\text{IR}^L)$  bind and  $W_0^H$  to make (Weak- $\text{IC}^{HL}$ ) bind, which can always be done. Accordingly, let  $\bar{\mathbf{C}}^L = (\bar{t}^L, \bar{W}_0^L, \bar{T}^L(\bar{t}^L))$  be the connected penalty contract where  $\bar{W}_0^L$  is set to make  $(\text{IR}^L)$  bind. Recall that [RP1] differs from program [P1] in that it imposes (Weak- $\text{IC}^{HL}$ ) rather than ( $\text{IC}^{HL}$ ). We will argue that any solution to [RP1] using  $\bar{\mathbf{C}}^L$  satisfies ( $\text{IC}^{HL}$ ) and hence is also a solution to program [P1]. As shown in Step 2 of the proof of Theorem 2, contract  $\bar{\mathbf{C}}^L$  can then be combined with a suitable onetime-penalty contract for type  $H$  to produce a solution to the principal's original program [P].

We show that given any connected penalty contract of length  $T^L \leq t^L$  with penalty sequence  $\bar{T}^L(T^L)$ , it would be optimal for type  $H$  to work in every period  $1, \dots, T^L$ , no matter the history of prior effort. Fix any  $T^L \leq t^L$  and write  $\bar{T}^L \equiv \bar{T}^L(T^L)$ . The argument is by induction. Consider the last period,  $T^L$ . Since  $-\bar{\beta}_{T^L}^L \lambda^L \bar{T}_{T^L}^L = c$ , it follows from the fact that  $t^H > t^L$  (hence  $\bar{\beta}_t^H \lambda^H > \bar{\beta}_t^L \lambda^L$  for all  $t < t^H$ ) that no matter the history of effort,  $-\beta_{T^L}^H \lambda^H \bar{T}_{T^L}^L \geq c$ , *i.e.* regardless of the history, type  $H$  will work in period  $T^L$ . Now assume that it is optimal for type  $H$  to work in period  $t+1 \leq T^L$  no matter the history of effort, and consider period  $t$  with belief  $\beta_t^H$ . This inductive hypothesis implies that

$$-\beta_{t+1}^H \lambda^H \left\{ \bar{T}_{t+1}^L + \sum_{s=t+2}^{T^L} \delta^{s-(t+1)} (1-\lambda^H)^{s-(t+2)} \left[ (1-\lambda^H) \bar{T}_s^L - c \right] \right\} \geq c,$$

or equivalently,

$$\sum_{s=t+2}^{T^L} \delta^{s-(t+1)} (1-\lambda^H)^{s-(t+2)} \left[ (1-\lambda^H) \bar{T}_s^L - c \right] \leq -\frac{c}{\beta_{t+1}^H \lambda^H} - \bar{T}_{t+1}^L. \quad (\text{B.21})$$

Therefore, at period  $t < T^L$ :

$$\begin{aligned} & -\beta_t^H \lambda^H \left\{ \bar{T}_t^L + \delta \left[ (1-\lambda^H) \bar{T}_{t+1}^L - c \right] + \delta (1-\lambda^H) \sum_{s=t+2}^{T^L} \delta^{s-(t+1)} (1-\lambda^H)^{s-(t+2)} \left[ (1-\lambda^H) \bar{T}_s^L - c \right] \right\} \\ & \geq -\beta_t^H \lambda^H \left\{ \bar{T}_t^L + \delta \left[ (1-\lambda^H) \bar{T}_{t+1}^L - c \right] + \delta (1-\lambda^H) \left( -\frac{c}{\beta_{t+1}^H \lambda^H} - \bar{T}_{t+1}^L \right) \right\} \\ & = -\beta_t^H \lambda^H \left( \bar{T}_t^L - \delta c \right) + \delta (1-\lambda^H) \frac{\beta_t^H c}{\beta_{t+1}^H} = -\beta_t^H \lambda^H \bar{T}_t^L + \delta c \geq -\bar{\beta}_t^L \lambda^L \bar{T}_t^L + \delta c = c, \end{aligned}$$

where the first inequality uses (B.21), the second equality uses  $\beta_{t+1}^H = \frac{\beta_t^H(1-\lambda^H)}{1-\beta_t^H+\beta_t^H(1-\lambda^H)}$ , and the final equality uses the fact that  $\tilde{t}_t^L = -\frac{(1-\delta)c}{\beta_t^L\lambda^L}$ .

### C. PROOF OF THEOREM 5

We assume throughout this Appendix that  $\delta=1$ . Without loss of optimality by Proposition 1, we focus on menus of penalty contracts. In this Appendix, we will introduce programs and constraints that have analogies with those used in Appendix B for the case of  $t^H > t^L$ . Accordingly, we often use the same labels for equations as before, but the reader should bear in mind that all references in this Appendix to such equations are to those defined in this Appendix.

**Outline.** Since this is a long proof, let us outline the components. We begin in Step 1 by taking program [P2] from the proof of Theorem 2 for the case of  $\delta=1$ ; we continue to call this program [P2]. Note that a critical difference here relative to the relaxed program [RP2] in the proof of Theorem 3 is that the current program [P2] does not constrain what the high type must do when taking the low type’s contract.

In Step 2, we show that there is an optimal penalty contract for type  $L$  that is connected. In Step 3, we develop three lemmas pertaining to properties of the set  $\alpha^H(\mathbf{C}^L)$  in any  $\mathbf{C}^L$  that is an optimal contract for type  $L$ . We then use these lemmas in Step 4 to show that in solving [P2], we can restrict attention to connected penalty contracts  $\mathbf{C}^L$  for type  $L$  such that  $\alpha^H(\mathbf{C}^L)$  includes a *stopping strategy with the most work property*, i.e. an action plan that involves consecutive work for some number of periods followed by shirking thereafter, and where the number of work periods is larger than in any action plan in  $\alpha^H(\mathbf{C}^L)$ . Building on the restriction to stopping strategies, we then show in Step 5 that there is always an optimal contract for type  $L$  that is a onetime-penalty contract.

The last step, Step 6, is relegated to the Supplementary Appendix. For an arbitrary time  $T^L$ , this step first defines a particular last-period penalty  $l_{T^L}^L(T^L)$  and an associated time  $T^{HL}(T^L) \leq T^L$ , and then establishes that if  $T^L$  is the optimal length of experimentation for type  $L$ , there is an optimal onetime-penalty contract for type  $L$  with penalty  $l_{T^L}^L(T^L)$  and in which type  $H$ ’s most-work optimal stopping strategy involves  $T^{HL}(T^L)$  periods of work. Hence, using  $l_{T^L}^L(T^L)$  and  $T^{HL}(T^L)$ , an optimal contract for type  $L$  that solves [P2] can be found by optimizing over the length  $T^L$ . By Theorem 2, the optimal length,  $\tilde{t}^L$ , is no larger than the first-best stopping time,  $t^L$ .

#### C.1. Step 1: The principal’s program

By Step 1 and Step 2 in the proof of Theorem 2, we work with the principal’s program [P2]. Here we restate this program given  $\delta=1$ :

$$\max_{\substack{\mathbf{C}^H \in \mathcal{C} \\ \mathbf{C}^L \in \mathcal{C} \\ \mathbf{a}^H \\ \mathbf{a}^{HL} \in \alpha^H(\mathbf{C}^L)}} \left\{ \begin{aligned} & \mu_0 \left\{ \beta_0 \sum_{t \in \Gamma^H} \left[ \prod_{\substack{s \in \Gamma^H \\ s \leq t-1}} (1 - a_s^H \lambda^H) \right] a_t^H (\lambda^H - c) - (1 - \beta_0) \sum_{t \in \Gamma^H} a_t^H c \right\} \\ & + (1 - \mu_0) \left\{ \beta_0 \sum_{t \in \Gamma^L} \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] (\lambda^L - c) - (1 - \beta_0) \sum_{t \in \Gamma^L} c \right\} \\ & - \mu_0 \left\{ \begin{aligned} & \beta_0 \sum_{t \in \Gamma^L} l_t^L \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t}} (1 - a_s^{HL} \lambda^H) - \prod_{\substack{s \in \Gamma^L \\ s \leq t}} (1 - \lambda^L) \right] \\ & - \beta_0 c \sum_{t \in \Gamma^L} a_t^{HL} \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - a_s^{HL} \lambda^H) - \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] \\ & + c \sum_{t \in \Gamma^L} (1 - a_t^{HL}) \left[ (1 - \beta_0) + \beta_0 \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - \lambda^L) \right] \end{aligned} \right\} \end{aligned} \right. \quad (\text{P2})$$

Information rent of type  $H$

subject to

$$\mathbf{1} \in \arg \max_{(a_t)_{t \in \Gamma^L}} \left\{ \beta_0 \sum_{t \in \Gamma^L} \left[ \prod_{\substack{s \in \Gamma^L \\ s \leq t-1}} (1 - a_s \lambda^L) \right] [(1 - a_t \lambda^L) l_t^L - a_t c] + (1 - \beta_0) \sum_{t \in \Gamma^L} (l_t^L - a_t c) + W_0^L \right\}, \quad (\text{IC}_a^L)$$

$$\mathbf{a}^H \in \arg \max_{(a_t)_{t \in \Gamma^H}} \left\{ \beta_0 \sum_{t \in \Gamma^H} \left[ \prod_{\substack{s \in \Gamma^H \\ s \leq t-1}} (1 - a_s \lambda^H) \right] [(1 - a_t \lambda^H) l_t^H - a_t c] + (1 - \beta_0) \sum_{t \in \Gamma^H} (l_t^H - a_t c) + W_0^H \right\}. \quad (\text{IC}_a^H)$$

As in the second step in the proof of Theorem 2, the information rent of type  $H$  when he takes action plan  $\mathbf{a}$  under type  $L$ 's contract  $\mathbf{C}^L$  is given by  $R(\mathbf{C}^L, \mathbf{a}) = U_0^H(\mathbf{C}^L, \mathbf{a}) - U_0^L(\mathbf{C}^L, \mathbf{1})$ , and  $R(\mathbf{C}^L, \mathbf{a}) = R(\mathbf{C}^L, \mathbf{a}')$  whenever  $\mathbf{a}, \mathbf{a}' \in \alpha^H(\mathbf{C}^L)$ . The difference in information rents under contracts  $\widehat{\mathbf{C}}^L$  and  $\mathbf{C}^L$  and corresponding action plans  $\widehat{\mathbf{a}}$  and  $\mathbf{a}$  is:

$$R(\widehat{\mathbf{C}}^L, \widehat{\mathbf{a}}) - R(\mathbf{C}^L, \mathbf{a}) = (U_0^H(\widehat{\mathbf{C}}^L, \widehat{\mathbf{a}}) - U_0^H(\mathbf{C}^L, \mathbf{a})) - (U_0^L(\widehat{\mathbf{C}}^L, \mathbf{1}) - U_0^L(\mathbf{C}^L, \mathbf{1})). \quad (\text{C.1})$$

When the action plan does not change across contracts (*i.e.*  $\mathbf{a} = \widehat{\mathbf{a}}$  above), (C.1) specializes to

$$R(\widehat{\mathbf{C}}^L, \mathbf{a}) - R(\mathbf{C}^L, \mathbf{a}) = \beta_0 \sum_{t \in \Gamma^L} (\widehat{l}_t^L - l_t^L) \left[ \prod_{s \in \Gamma^L, s \leq t} (1 - a_s \lambda^H) - \prod_{s \in \Gamma^L, s \leq t} (1 - \lambda^L) \right]. \quad (\text{C.2})$$

### C.2. Step 2: Connected contracts for the low type

We now claim that in program [P2], it is without loss to consider solutions in which the low type's contract is a connected penalty contract, *i.e.* solutions  $\mathbf{C}^L$  in which  $\Gamma^L = \{1, \dots, T^L\}$  for some  $T^L$ .

To avoid trivialities, consider any optimal non-connected  $\mathbf{C}^L$  with  $\Gamma^L \neq \emptyset$ . Let  $t^\circ$  be the earliest lockout period in  $\mathbf{C}^L$ , *i.e.*  $t^\circ = \min\{t: t > 0, t \notin \Gamma^L\}$ . Consider a modified penalty contract  $\widehat{\mathbf{C}}^L$  that removes the lockout period  $t^\circ$  and shortens the contract by one period as follows:

$$\widehat{W}_0^L = W_0^L, \quad \widehat{\Gamma}^L = \{1, \dots, t^\circ - 1\} \cup \{s: s \geq t^\circ \text{ and } s+1 \in \Gamma^L\}, \quad \widehat{l}_s^L = \begin{cases} l_s^L & \text{if } s \leq t^\circ - 1, \\ l_{s+1}^L & \text{if } s \geq t^\circ \text{ and } s \in \widehat{\Gamma}^L. \end{cases}$$

Given  $\delta = 1$ , it is straightforward that it remains optimal for type  $L$  to work in every period in  $\widehat{\Gamma}^L$ , and given any optimal action plan for type  $H$  under the original contract,  $\mathbf{a}^{HL} \in \alpha^H(\mathbf{C}^L)$ , the action plan

$$\widehat{\mathbf{a}}^{HL} = (\widehat{a}_s^{HL})_{s \in \widehat{\Gamma}^L} = \begin{cases} a_s^{HL} & \text{if } s \leq t^\circ - 1, \\ a_{s+1}^{HL} & \text{if } s \geq t^\circ \text{ and } s \in \widehat{\Gamma}^L, \end{cases}$$

is optimal for type  $H$  under the modified contract, *i.e.*  $\widehat{\mathbf{a}}^{HL} \in \alpha^H(\widehat{\mathbf{C}}^L)$ . Given no discounting, it is also immediate that the surplus generated by type  $L$  is unchanged by the modification. It thus follows that the value of (P2) is unchanged by the modification. This procedure can be applied iteratively to all lockout periods to produce a connected contract.

### C.3. Step 3: Optimal deviation action plans for the high type

By the previous steps, we can restrict our attention to connected penalty contracts  $\mathbf{C}^L = (T^L, W_0^L, l^L)$  that induce effort from the low type in each period  $t \in \{1, \dots, T^L\}$ . We now describe properties of an optimal connected penalty contract for the low type (Step 3a) and an optimal action plan for the high type when taking the low type's contract (Step 3b).

**Step 3a:** Consider an optimal connected penalty contract for type  $L$ ,  $\mathbf{C}^L = (T^L, W_0^L, l^L)$ . The next two lemmas describe properties of such a contract.

**Lemma 1.** Suppose that  $\mathbf{C}^L = (T^L, W_0^L, I^L)$  is an optimal contract for type  $L$ . Then for any  $t = 1, \dots, T^L$ , there exists an optimal action plan  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$  such that  $a_t = 1$ .

*Proof* Suppose to the contrary that for some  $\tau \in \{1, \dots, T^L\}$ ,  $a_\tau = 0$  for all  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$ . For any  $\varepsilon > 0$ , define a contract  $\mathbf{C}^L(\varepsilon) = (T^L, W_0^L, I^L(\varepsilon))$  modified from  $\mathbf{C}^L = (T^L, W_0^L, I^L)$  as follows: (1)  $I_\tau^L(\varepsilon) = I_\tau^L - \varepsilon$ ; (2)  $I_{\tau-1}^L(\varepsilon) = I_{\tau-1}^L + \varepsilon(1 - \lambda^L)$ ; and (3)  $I_t^L(\varepsilon) = I_t^L$  if  $t \notin \{\tau - 1, \tau\}$ . We derive a contradiction by showing that for small enough  $\varepsilon > 0$ ,  $\mathbf{C}^L(\varepsilon)$  together with an original optimal contract for type  $H$ ,  $\mathbf{C}^H$ , is feasible in [P2] and strictly improves the objective. Note that by construction,  $(\mathbf{C}^L(\varepsilon), \mathbf{C}^H)$  satisfy  $(IC_a^L)$  and  $(IC_a^H)$ . To evaluate how the objective changes when  $\mathbf{C}^L(\varepsilon)$  is used instead of  $\mathbf{C}^L$ , we thus only need to consider the difference in the information rents associated with these contracts,  $R(\mathbf{C}^L(\varepsilon)) - R(\mathbf{C}^L)$ .

We first claim that  $\alpha^H(\mathbf{C}^L(\varepsilon)) \subseteq \alpha^H(\mathbf{C}^L)$  when  $\varepsilon$  is small enough. To see this, fix any  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$ . Since the set of action plans is discrete, the optimality of  $\mathbf{a}$  implies that there is some  $\eta > 0$  such that  $U_0^H(\mathbf{C}^L, \mathbf{a}) > U_0^H(\mathbf{C}^L, \mathbf{a}') + \eta$  for any  $\mathbf{a}' \notin \alpha^H(\mathbf{C}^L)$ . Since  $U_0^H(\mathbf{C}^L(\varepsilon), \mathbf{a}')$  is continuous in  $\varepsilon$ , it follows immediately that for all  $\varepsilon$  small enough and all  $\mathbf{a}' \notin \alpha^H(\mathbf{C}^L)$ :  $U_0^H(\mathbf{C}^L(\varepsilon), \mathbf{a}') > U_0^H(\mathbf{C}^L(\varepsilon), \mathbf{a}) + \eta$ . Thus,  $\mathbf{a}' \notin \alpha^H(\mathbf{C}^L(\varepsilon))$ . It follows that  $\alpha^H(\mathbf{C}^L(\varepsilon)) \subseteq \alpha^H(\mathbf{C}^L)$ .

Next, for small enough  $\varepsilon$ , take  $\mathbf{a} \in \alpha^H(\mathbf{C}^L(\varepsilon)) \subseteq \alpha^H(\mathbf{C}^L)$ . Since  $a_\tau = 0$  by assumption, (C.2) implies

$$\begin{aligned} R(\mathbf{C}^L(\varepsilon), \mathbf{a}) - R(\mathbf{C}^L, \mathbf{a}) &= \beta_0 \varepsilon (1 - \lambda^L) \left[ \prod_{s=1}^{\tau-1} (1 - a_s \lambda^H) - (1 - \lambda^L)^{\tau-1} \right] - \beta_0 \varepsilon \left[ \prod_{s=1}^{\tau} (1 - a_s \lambda^H) - (1 - \lambda^L)^\tau \right] \\ &= -\lambda^L \beta_0 \varepsilon \prod_{s=1}^{\tau-1} (1 - a_s \lambda^H) < 0. \end{aligned}$$

Hence,  $\mathbf{C}^L(\varepsilon)$  strictly improves the objective relative to  $\mathbf{C}^L$ .  $\parallel$

**Lemma 2.** Suppose that  $\mathbf{C}^L = (T^L, W_0^L, I^L)$  is an optimal contract for type  $L$  and there is some  $\tau \in \{1, \dots, T^L\}$  such that  $a_\tau = 1$  for all  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$ . Then  $(IC_a^L)$  binds at  $\tau$ .

*Proof* Recall from  $(IC_a^L)$  that  $\mathbf{a}^L = \mathbf{1}$ . Suppose to the contrary that  $(IC_a^L)$  is not binding at some  $\tau$  but  $a_\tau = 1$  for all  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$ . For any  $\varepsilon > 0$ , define a contract  $\mathbf{C}^L(\varepsilon) = (T^L, W_0^L, I^L(\varepsilon))$  modified from  $\mathbf{C}^L = (T^L, W_0^L, I^L)$  as follows: (1)  $I_\tau^L(\varepsilon) = I_\tau^L + \varepsilon$ ; (2)  $I_{\tau-1}^L(\varepsilon) = I_{\tau-1}^L - \varepsilon(1 - \lambda^L)$ ; and (3)  $I_t^L(\varepsilon) = I_t^L$  if  $t \notin \{\tau - 1, \tau\}$ . We derive a contradiction by showing that for small enough  $\varepsilon > 0$ ,  $\mathbf{C}^L(\varepsilon)$  together with an original optimal contract for type  $H$ ,  $\mathbf{C}^H$ , is feasible in [P2] and strictly improves the objective. Note that by construction  $(IC_a^L)$  is still satisfied under  $\mathbf{C}^L(\varepsilon)$  at  $t = 1, \dots, \tau - 1, \tau + 1, \dots, T^L$ . Moreover, since  $(IC_a^L)$  is slack at  $\tau$  under contract  $\mathbf{C}^L$ , it continues to be slack at  $\tau$  under  $\mathbf{C}^L(\varepsilon)$  for  $\varepsilon$  small enough.

For small enough  $\varepsilon$ , take any  $\mathbf{a} \in \alpha^H(\mathbf{C}^L(\varepsilon)) \subseteq \alpha^H(\mathbf{C}^L)$ , where the subset inequality follows from the arguments in the proof of Lemma 1. Recall that by assumption  $a_\tau = 1$ . Using (C.2),

$$\begin{aligned} R(\mathbf{C}^L(\varepsilon), \mathbf{a}) - R(\mathbf{C}^L, \mathbf{a}) &= -\beta_0 \varepsilon (1 - \lambda^L) \left[ \prod_{s=1}^{\tau-1} (1 - a_s \lambda^H) - (1 - \lambda^L)^{\tau-1} \right] + \beta_0 \varepsilon \left[ \prod_{s=1}^{\tau} (1 - a_s \lambda^H) - (1 - \lambda^L)^\tau \right] \\ &= \beta_0 \varepsilon \left\{ \prod_{s=1}^{\tau-1} (1 - a_s \lambda^H) [-(1 - \lambda^L) + (1 - \lambda^H)] \right\} \\ &= -(\lambda^H - \lambda^L) \beta_0 \varepsilon \prod_{s=1}^{\tau-1} (1 - a_s \lambda^H) \\ &< 0. \end{aligned}$$

Hence,  $\mathbf{C}^L(\varepsilon)$  strictly improves the objective relative to  $\mathbf{C}^L$ .  $\parallel$

**Step 3b:** For any  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$  and  $s < t$ , define

$$D(s, t, \mathbf{a}) := \sum_{\tau=s}^{t-1} I_\tau^L (1 - \lambda^H)^{\sum_{n=s+1}^{\tau} a_n}.$$

The next lemma describes properties of any action plan  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$ .

**Lemma 3.** Suppose  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$  and  $s < t$ .

- (1) If  $D(s, t, \mathbf{a}) > 0$  and  $a_s = 1$ , then  $a_t = 1$ .
- (2) If  $D(s, t, \mathbf{a}) < 0$  and  $a_s = 0$ , then  $a_t = 0$ .
- (3) If  $D(s, t, \mathbf{a}) = 0$ , then  $\mathbf{a}' \in \alpha^H(\mathbf{C}^L)$  where  $a'_s = a_t$ ,  $a'_t = a_s$ , and  $a'_\tau = a_\tau$  if  $\tau \neq s, t$ .

*Proof* Consider the first case of  $D(s, t, \mathbf{a}) > 0$ . Suppose to the contrary that for some optimal action plan  $\mathbf{a}$  and two periods  $s < t$ , we have  $D(s, t, \mathbf{a}) > 0$  and  $a_s = 1$  but  $a_t = 0$ . Consider an action plan  $\mathbf{a}'$  such that  $\mathbf{a}'$  and  $\mathbf{a}$  agree except that  $a'_s = 0$  and  $a'_t = 1$ . That is,

$$\mathbf{a} = (\dots, \underbrace{1}_{\text{period } s}, \dots, \underbrace{0}_{\text{period } t}, \dots),$$

$$\mathbf{a}' = (\dots, \underbrace{0}_{\text{period } s}, \dots, \underbrace{1}_{\text{period } t}, \dots).$$

Let  $U_s^H(\mathbf{C}^L, \mathbf{a})$  be type  $H$ 's payoff evaluated at the beginning of period  $s$ . Then,

$$U_s^H(\mathbf{C}^L, \mathbf{a}) - U_s^H(\mathbf{C}^L, \mathbf{a}') = -\beta_s^H \lambda^H \sum_{\tau=s}^{t-1} l_\tau^L (1 - \lambda^H)^{\sum_{n=s+1}^{\tau} a_n} = -\beta_s^H \lambda^H D(s, t, \mathbf{a}).$$

The intuition for this expression is as follows. Since action plans  $\mathbf{a}$  and  $\mathbf{a}'$  have the same number of working periods, the assumption of no discounting implies that neither the effort costs nor the penalty sequence matters for the difference in utilities conditional on the bad state. Conditional on the good state, the effort costs again do not affect the difference in utilities; however, the probability with which the agent receives  $l_\tau^L$  for any  $\tau \in \{s, \dots, t-1\}$  is "shifted up" in  $\mathbf{a}'$  as compared to  $\mathbf{a}$ .

Therefore,  $U_s^H(\mathbf{C}^L, \mathbf{a}) - U_s^H(\mathbf{C}^L, \mathbf{a}') < 0$  if  $D(s, t, \mathbf{a}) > 0$ . But this contradicts the assumption that  $\mathbf{a}$  is optimal; hence, the claim in part (1) follows. The proof of part (2) is analogous.

Finally, consider part (3). The claim is trivial if  $a_s = a_t$ . If  $a_s = 1$  and  $a_t = 0$ , then  $U_s^H(\mathbf{C}^L, \mathbf{a}) - U_s^H(\mathbf{C}^L, \mathbf{a}') = 0$  from the argument above; hence, both  $\mathbf{a}$  and  $\mathbf{a}'$  are optimal. The case of  $a_s = 0$  and  $a_t = 1$  is analogous.  $\parallel$

#### C.4. Step 4: Stopping strategies for the high type

We use the following concepts to characterize the solution to [P2]:

**Definition 1.** An action plan  $\mathbf{a}$  is a stopping strategy (that stops at  $t$ ) if there exists  $t \geq 1$  such that  $a_s = 1$  for  $s \leq t$  and  $a_s = 0$  for  $s > t$ .

**Definition 2.** An optimal action plan for type  $\theta$  under contract  $\mathbf{C}$ ,  $\mathbf{a} \in \alpha^\theta(\mathbf{C})$ , has the most-work property (or is a most-work optimal strategy) if no other optimal action plan under the contract has more work periods; that is, for all  $\mathbf{a}' \in \alpha^\theta(\mathbf{C})$ ,  $\#\{n: a_n = 1\} \geq \#\{n: a'_n = 1\}$ .

Step 3 described properties of optimal contracts for the low type and optimal action plans for the high type under the low type's contract. We now use these properties to show that in solving program [P2], we can restrict attention to connected penalty contracts for the low type  $\mathbf{C}^L = (T^L, W_0^L, l^L)$  such that there is an optimal action plan for the high type under the contract  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$  that is a stopping strategy with the most work property.

Let  $N = \min_{\mathbf{a} \in \alpha^H(\mathbf{C}^L)} \#\{n: a_n = 0\}$ . That is, among all action plans that are optimal for type  $H$  under contract  $\mathbf{C}^L$ , the action plan in which type  $H$  works the largest number of periods involves type  $H$  shirking in  $N$  periods. Let  $A_N$  be the set of optimal action plans that involve type  $H$  shirking in  $N$  periods. Let

$$A_{N,k} = \{\mathbf{a} \in A_N : a_t = 0 \text{ for all } t > T^L - k\},$$

i.e. any  $\mathbf{a} \in A_{N,k}$  contains a total of  $N$  shirking periods, (at least)  $k$  of which are in the tail.

Our goal is to establish the following:

$$\text{for any } k < N: A_{N,k} \neq \emptyset \implies \bigcup_{n=k+1}^N A_{N,n} \neq \emptyset. \quad (\text{C.3})$$

In other words, whenever  $A_N$  contains an action plan that has  $k < N$  shirks in the tail,  $A_N$  must contain an action plan that has at least  $k+1$  shirks in the tail. By induction, this implies  $A_{N,N} \neq \emptyset$ , which is equivalent to the existence of an optimal action plan that is a stopping strategy with the most work property.

Suppose to contradiction that (C.3) is not true; *i.e.* there is some  $k < N$  such that  $A_{N,k} \neq \emptyset$  and yet  $\bigcup_{n=k+1}^N A_{N,n} = \emptyset$ . Then there exists

$$\hat{t} = \min \{t: \mathbf{a} \in A_{N,k}, a_t = 0, t < T^L - k, a_s = 1 \text{ for each } s = t + 1, \dots, T^L - k\}. \quad (\text{C.4})$$

In words,  $\hat{t}$  is the smallest shirking period preceding a working period such that there is an optimal action plan  $\mathbf{a} \in A_{N,k}$  with  $k + 1$  shirking periods from (including)  $\hat{t}$ . Now take  $\hat{t}_0 = \hat{t}$ . For  $n = 0, 1, \dots$ , whenever  $\{t: a_t = 0, \mathbf{a} \in A_{N,k}, t < \hat{t}_n\} \neq \emptyset$ , define

$$\hat{t}_{n+1} = \min \{t: a_t = 0, \mathbf{a} \in A_{N,k}, t < \hat{t}_n, a_s = 1 \text{ for each } s = t + 1, \dots, \hat{t}_n - 1\}. \quad (\text{C.5})$$

The sequence  $\{\hat{t}_n\}$  uniquely pins down an action profile  $\hat{\mathbf{a}} \in A_{N,k}$ . In words, among all effort profiles in  $A_{N,k}$ ,  $\hat{\mathbf{a}}$  has the earliest  $n$ -th shirk for each  $n = 1, \dots, N$ . Note that  $\hat{\mathbf{a}}$  takes the following form:

period:	$\hat{t}$	$\hat{t} + 1$	...	$T^L - k$	$T^L - k + 1$	...
$\hat{\mathbf{a}}$ :	0	1	...	1	0	...

We will prove that  $\bigcup_{n=k+1}^N A_{N,n} \neq \emptyset$  (contradicting the hypothesis above) by showing that we can “move” the shirking in period  $\hat{t}$  of  $\hat{\mathbf{a}}$  to the end. This is done via three lemmas.

**Lemma 4.** *Suppose  $A_{N,k} \neq \emptyset$  and  $\bigcup_{n=k+1}^N A_{N,n} = \emptyset$ . Then  $l_t^L = 0$  for any  $t = \hat{t} + 1, \dots, T^L - k - 1$ .*

*Proof.* We proceed by induction. Take any  $t \in \{\hat{t} + 1, \dots, T^L - k - 1\}$  and assume that  $l_s^L = 0$  for  $s = t + 1, \dots, T^L - k - 1$ . We show that  $l_t^L = 0$ .

Step 1:  $l_t^L \geq 0$ .

Proof of Step 1: Suppose not, *i.e.*  $l_t^L < 0$ . Then the fact that  $(IC_a^L)$  is satisfied at period  $t + 1$  and the hypothesis that  $l_t^L < 0$  imply that  $(IC_a^L)$  is slack at period  $t$ .<sup>54</sup> Hence, by Lemma 2, there exists an action plan  $\mathbf{a}' \in \alpha^H(\mathbf{C}^L)$  such that  $a'_t = 0$ . Now, by the assumption that  $l_t^L < 0$  together with the induction hypothesis, we obtain  $\sum_{s=t}^m l_s^L < 0$  for  $m \in \{t, \dots, T^L - k - 1\}$ . By Lemma 3, part (2),  $a'_s = 0$  for any  $s = t, \dots, T^L - k$ . Thus,  $\mathbf{a}' \in \alpha^H(\mathbf{C}^L)$  is as follows:

period:	$\hat{t}$	$\hat{t} + 1$	...	$t$	$t + 1$	...	$T^L - k - 1$	$T^L - k$	$T^L - k + 1$	...
$\hat{\mathbf{a}}$ :	0	1	...	1	1	...	1	1	0	...
$\mathbf{a}'$ :				0	0	...	0	0		

*Claim 1:* There exists  $s^* > T^L - k$  such that  $a'_{s^*} = 1$ .

*Proof:* Suppose not. Then  $a'_s = 0$  for all  $s \geq T^L - k$  (recall  $a'_{T^L - k} = 0$ ). We claim this implies  $\#\{n: a'_n = 0\} > N$ . To see this, note that  $\#\{n: a'_n = 0\} \geq N$  by assumption. If  $\#\{n: a'_n = 0\} = N$ , then  $\mathbf{a}' \in A_N$ , and since  $\mathbf{a}'$  contains  $k + 1$  shirking periods in its tail, it follows that  $\mathbf{a}' \in A_{N,k+1}$ , contradicting the assumption that  $\bigcup_{n=k+1}^N A_{N,n} = \emptyset$ . Given that  $\#\{n: a'_n = 0\} > N$  and  $a'_s = 0$  for  $s \geq T^L - k$ , it follows that  $\beta_{T^L - k}^H(\mathbf{a}') \geq \beta_{T^L - k}^H(\hat{\mathbf{a}})$  and taking  $a'_{T^L - k} = 1$  is optimal, a contradiction.  $\parallel$

Now let  $s^*$  be the first such working period after  $T^L - k$ . Then,

period:	$\hat{t}$	$\hat{t} + 1$	...	$t$	$t + 1$	...	$T^L - k - 1$	$T^L - k$	$T^L - k + 1$	...	$s^*$	...
$\hat{\mathbf{a}}$ :	0	1	...	1	1	...	1	1	0	...	0	...
$\mathbf{a}'$ :				0	0	...	0	0	0	...	1	

Applying parts (1) and (2) of Lemma 3 to  $\hat{\mathbf{a}}$  and  $\mathbf{a}'$ , we obtain  $\sum_{s=T^L - k}^{s^* - 1} l_s^L = 0$ . Now applying part (3) of Lemma 3, we obtain that the agent is indifferent between  $\mathbf{a}'$  and  $\mathbf{a}''$  where  $\mathbf{a}''$  differs from  $\mathbf{a}'$  only by switching the actions in period  $T^L - k$  and period  $s^*$ . But since  $\sum_{s=t}^{T^L - k - 1} l_s^L < 0$ , the optimality of  $a'_t = 0, a'_{T^L - k} = 1$  contradicts part (2) of Lemma 3.

Step 2:  $l_t^L \leq 0$ .

Proof of Step 2: Assume to the contrary that  $l_t^L > 0$ . We have two cases to consider.

*Case 1:*  $l_{T^L - k}^L \geq 0$ .

By the induction hypothesis and the assumption that  $l_t^L > 0$ , we have  $\sum_{s=t}^{T^L - k} l_s^L (1 - \lambda^H)^{\sum_{n=t+1}^s \hat{a}_n} > 0$ . Therefore, by part (1) of Lemma 3,  $\hat{a}_{T^L - k + 1} = 1$ . But this contradicts the definition of  $\hat{\mathbf{a}}$ .

*Case 2:*  $l_{T^L - k}^L < 0$ .

54. This can be proved along very similar lines to part (2) of Lemma 3.

In this case,  $(IC_a^L)$  must be slack in period  $T^L - k$  (since it is satisfied in the next period and  $l_{T^L - k}^L < 0$ ). Hence by Lemma 2, there exists  $\tilde{\mathbf{a}}$  such that  $\tilde{a}_{T^L - k} = 0$ .

period:	$\hat{t}$	$\hat{t}+1$	...	$t$	$t+1$	...	$T^L - k - 1$	$T^L - k$	...
$\hat{\mathbf{a}}$ :	0	1	...	1	1	...	1	1	...
$\tilde{\mathbf{a}}$ :							1	0	

Claim 2:  $\tilde{a}_s = 0$  for any  $s > T^L - k$ .

Proof: Suppose the claim is not true. Then define

$$\tau := \min \{s : s > T^L - k \text{ and } \tilde{a}_s = 1\}.$$

This is shown in the following table:

period:	$\hat{t}$	$\hat{t}+1$	...	$t$	$t+1$	...	$T^L - k - 1$	$T^L - k$	...	$\tau - 1$	$\tau$	...
$\hat{\mathbf{a}}$ :	0	1	...	1	1	...	1	1	...	0	0	...
$\tilde{\mathbf{a}}$ :								0	...	0	1	

Applying parts (1) and (2) of Lemma 3 to  $\hat{\mathbf{a}}$  and  $\tilde{\mathbf{a}}$  respectively, we obtain

$$\sum_{s=T^L - k}^{\tau - 1} l_s^L = 0. \tag{C.6}$$

But then, by the induction hypothesis and the assumption that  $l_t^L > 0$ , we obtain

$$\sum_{s=t}^{T^L - k - 1} l_s^L (1 - \lambda^H)^{\sum_{n=t+1}^s \hat{a}_n} > 0. \tag{C.7}$$

Notice that  $\hat{a}_s = 0$  for  $s > T^L - k$  by definition. Hence, (C.6) and (C.7) imply  $\sum_{s=t}^{\tau - 1} l_s^L (1 - \lambda^H)^{\sum_{n=t+1}^s \hat{a}_n} > 0$ . Now applying part (1) of Lemma 3 to  $\hat{\mathbf{a}}$ , we reach the conclusion that  $\hat{a}_\tau = 1$ , a contradiction.  $\parallel$

Hence, we have established the claim that  $\tilde{a}_s = 0$  for all  $s > T^L - k$ , as depicted below:

period:	$\hat{t}$	$\hat{t}+1$	...	$t$	$t+1$	...	$T^L - k - 1$	$T^L - k$	...	$\tau - 1$	$\tau$	...
$\hat{\mathbf{a}}$ :	0	1	...	1	1	...	1	1	...	0	0	...
$\tilde{\mathbf{a}}$ :								0	...	0	0	...

Claim 3:  $\#\{n : \tilde{a}_n = 0\} = N + 1$  and  $\beta_{T^L - k}^H(\tilde{\mathbf{a}}) = \beta_{T^L - k}^H(\hat{\mathbf{a}})$ .

Proof: By definition of  $N$ ,  $\#\{n : \tilde{a}_n = 0\} \geq N$ . If  $\#\{n : \tilde{a}_n = 0\} = N$ , then  $\tilde{\mathbf{a}}$  contains  $k + 1$  shirking periods in its tail, contradicting the assumption that  $A_{N, k+1} = \emptyset$ . Moreover, if  $\#\{n : \tilde{a}_n = 0\} > N + 1$ , then  $\beta_{T^L - k}^H(\tilde{\mathbf{a}}) > \beta_{T^L - k}^H(\hat{\mathbf{a}})$ . But then since  $\hat{a}_{T^L - k} = 1$ , we should have  $\tilde{a}_{T^L - k} = 1$ , a contradiction. Therefore, it must be  $\#\{n : \tilde{a}_n = 0\} = N + 1$ .  $\parallel$

By Claim 3, we can choose  $\tilde{\mathbf{a}}$  such that  $\tilde{\mathbf{a}}$  differs from  $\hat{\mathbf{a}}$  only in period  $T^L - k$ . This is shown in the following table:

period:	$\hat{t}$	$\hat{t}+1$	...	$t$	$t+1$	...	$T^L - k - 1$	$T^L - k$	...	$\tau - 1$	$\tau$	...
$\hat{\mathbf{a}}$ :	0	1	...	1	1	...	1	1	...	0	0	...
$\tilde{\mathbf{a}}$ :	0	1	...	1	1	...	1	0	...	0	0	...

But by assumption  $l_t^L > 0$ , and by the induction hypothesis  $l_s^L = 0$  for  $s = t + 1, \dots, T^L - k - 1$ . Therefore,

$$\sum_{s=t}^{T^L - k - 1} l_s^L (1 - \lambda^H)^{\sum_{n=t+1}^s \tilde{a}_n} > 0.$$

Applying part (1) of Lemma 3, we must conclude that  $\tilde{a}_{T^L - k} = 1$ , a contradiction.  $\parallel$

**Lemma 5.** Suppose  $A_{N, k} \neq \emptyset$  and  $\bigcup_{m=k+1}^N A_{N, m} = \emptyset$ . Then  $l_t^L = 0$ .

Proof Step 1:  $l_t^L \geq 0$ .

Proof of Step 1: To the contrary, suppose  $l_t^L < 0$ . Lemma 4 implies  $\sum_{s=t}^{T^L - k - 1} l_s^L (1 - \lambda^H)^{\sum_{n=t+1}^s \hat{a}_n} < 0$ . Then part (2) of Lemma 3 implies that  $\hat{a}_{T^L - k} = 0$ , a contradiction.

Step 2:  $l_t^L \leq 0$ .

**Proof of Step 2:** Suppose to the contrary that  $l_t^L > 0$ . Note that by Lemma 1, there exists an action plan  $\mathbf{a}' \in \alpha^H(\mathbf{C}^L)$  such that  $a'_t = 1$ . Then since, by Lemma 4,  $l_t^L = 0$  for  $t = \hat{t} + 1, \dots, T^L - k - 1$ , it follows from part (1) of Lemma 3 that  $a'_s = 1$  for  $s = \hat{t} + 1, \dots, T^L - k$ . Hence, we obtain the following table:

period:	$\hat{t}$	$\hat{t} + 1$	$\dots$	$T^L - k - 1$	$T^L - k$	$T^L - k + 1$	$\dots$
$\hat{\mathbf{a}}$ :	0	1	$\dots$	1	1	0	$\dots$
$\mathbf{a}'$ :	1	1	$\dots$	1	1		

*Claim:* there exists  $\tilde{t} < \hat{t}$ , such that  $a'_{\tilde{t}} = 0$  and  $\hat{a}_{\tilde{t}} = 1$ .

*Proof:* since  $\#\{t : a'_t = 0\} \geq N = \#\{t : \hat{a}_t = 0\}$ ,  $a'_t = 1$ ,  $\hat{a}_t = 0$ , and  $\hat{a}_t = 0$  for all  $t > T^L - k$ , we have  $\#\{t : a'_t = 0, t < \hat{t}\} > \#\{t : \hat{a}_t = 0, t < \hat{t}\}$ . The claim follows immediately.

We can take  $\tilde{t}$  to be the largest period that satisfies the above claim. Hence  $\hat{\mathbf{a}}$  and  $\mathbf{a}'$  are as follows:

period:	$\tilde{t}$	$\dots$	$\hat{t}$	$\hat{t} + 1$	$\dots$	$T^L - k - 1$	$T^L - k$	$T^L - k + 1$	$\dots$
$\hat{\mathbf{a}}$ :	1		0	1	$\dots$	1	1	0	$\dots$
$\mathbf{a}'$ :	0		1	1	$\dots$	1	1		

There are two cases to consider.

*Case 1:*  $\hat{a}_t = a'_t$  for each  $t = \tilde{t} + 1, \dots, \hat{t} - 1$ .

Lemma 3 implies that  $\sum_{s=\tilde{t}}^{\hat{t}-1} l_s^L (1 - \lambda^H)^{\sum_{n=\tilde{t}+1}^s \hat{a}_n} = 0$  and the agent is indifferent between  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{a}}'$  where  $\hat{\mathbf{a}}'$  differs from  $\hat{\mathbf{a}}$  only in that the actions at periods  $\tilde{t}$  and  $\hat{t}$  are switched. But this contradicts the definition of  $\hat{t}$  (see (C.4)).

*Case 2:*  $\hat{a}_m = 0$  and  $a'_m = 1$  for some  $m \in \{\tilde{t} + 1, \dots, \hat{t} - 1\}$ .

First note that Case 1 and Case 2 are exhaustive because  $\tilde{t}$  is taken to be the largest period  $t < \hat{t}$  such that  $\hat{a}_t = 1$  and  $a'_t = 0$ . Without loss, we take  $m$  to be the smallest possible. Hence  $\hat{a}_t = a'_t$  for each  $t = \tilde{t} + 1, \dots, m - 1$ . Then  $\hat{\mathbf{a}}$  and  $\mathbf{a}'$  are as follows:

period:	$\tilde{t}$	$\dots$	$m$	$\dots$	$\hat{t}$	$\hat{t} + 1$	$\dots$	$T^L - k - 1$	$T^L - k$	$T^L - k + 1$	$\dots$
$\hat{\mathbf{a}}$ :	1		0		0	1	$\dots$	1	1	0	$\dots$
$\mathbf{a}'$ :	0		1		1	1	$\dots$	1	1		

But again, by Lemma 3, we can switch the actions at periods  $\tilde{t}$  and  $m$  in  $\hat{\mathbf{a}}$ , contradicting the definition of  $\hat{\mathbf{a}}$  (see (C.5)).  
 $\parallel$

**Lemma 6.** *If  $A_{N,k} \neq \emptyset$  then  $\bigcup_{n=k+1}^N A_{N,n} \neq \emptyset$ .*

*Proof* Suppose to the contrary that  $\bigcup_{n=k+1}^N A_{N,n} = \emptyset$ . Then  $l_t^L = 0$  for  $t = \hat{t}, \dots, T^L - k - 1$ , by Lemma 4 and Lemma 5. Therefore, by part (3) of Lemma 3, we can switch  $\hat{a}_{\hat{t}}$  with  $\hat{a}_{T^L - k}$  to obtain  $\hat{\mathbf{a}}'$ . However, since  $\#\{t : \hat{a}'_t = 0\} = \#\{t : \hat{a}_t = 0\} = N$ , it follows immediately that  $\hat{\mathbf{a}}' \in A_N$ . Since  $\hat{a}'_t = 0$  for all  $t > T^L - k - 1$ ,  $\hat{\mathbf{a}}' \in \bigcup_{n=k+1}^N A_{N,n}$ .  $\parallel$

**C.5. Step 5: Onetime-penalty contracts for the low type**

In Step 4, we showed that we can restrict attention in solving program [P2] to connected penalty contracts for the low type  $\mathbf{C}^L = (T^L, W_0^L, l^L)$  such that there is an optimal action plan for the high type  $\mathbf{a} \in \alpha^H(\mathbf{C}^L)$  that is a stopping strategy with the most work property. We now use this result to show that we can further restrict attention to onetime-penalty contracts for the low type,  $\mathbf{C}^L = (T^L, W_0^L, l_{T^L}^L)$ .

This result is proved via two lemmas.

**Lemma 7.** *Let  $\mathbf{C}^L = (T^L, W_0^L, l^L)$  be an optimal contract for the low type with a most-work optimal stopping strategy for the high type  $\hat{\mathbf{a}}$  that stops at  $\hat{t}$ , i.e.  $\hat{t} = \max\{t \in \{1, \dots, T^L\} : \hat{a}_t = 1\}$ . For each  $t > \hat{t}$ , there is an optimal action plan,  $\hat{\mathbf{a}} \in \alpha^H(\mathbf{C}^L)$ , such that for any  $s$ ,  $\hat{a}_s = \hat{a}_s \iff s \notin \{\hat{t}, t\}$ .*

*Proof* **Step 1:** First, we show that the Lemma's claim is true for some  $t > \hat{t}$  (rather than for all  $t > \hat{t}$ ). Suppose not, to contradiction. Then Lemma 3 implies that

$$\text{for any } n \in \{\hat{t}, \hat{t} + 1, \dots, T^L - 1\}, \sum_{s=i}^n l_s^L < 0. \tag{C.8}$$

Hence,  $(IC_a^L)$  is slack at  $\hat{t}$  (since it is satisfied in the next period and  $l_{\hat{t}}^L < 0$ ) and, by Lemma 2, there exists an optimal action plan,  $\mathbf{a}''$ , with  $a''_{\hat{t}} = 0$ .



*Claim 1:*  $a''_s = 0$  for all  $s > \hat{t}$ .

*Proof:* Suppose to contradiction that there exists  $\tau > \hat{t}$  such that  $a''_\tau = 1$ . Take the smallest such  $\tau$ . Then it follows from Lemma 3 applied to  $\hat{\mathbf{a}}$  and  $\mathbf{a}''$  that  $\sum_{s=\hat{t}}^{\tau-1} l^L_s = 0$ , contradicting (C.8).  $\parallel$

Hence, we obtain that  $a''_s = 0$  for all  $s \geq \hat{t}$ , and it follows from the optimality of  $\hat{a}_{\hat{t}} = 1$  and  $a''_{\hat{t}} = 0$  that  $\mathbf{a}''$  is a stopping strategy that stops at  $\hat{t} - 1$ :

period:	...	$\hat{t} - 2$	$\hat{t} - 1$	$\hat{t}$	$\hat{t} + 1$	$\hat{t} + 2$	...
$\hat{\mathbf{a}}$ :	...	1	1	1	0	0	...
$\mathbf{a}''$ :	...	1	1	0	0	0	...

Next, note that by Lemma 1, there is an optimal action plan,  $\mathbf{a}'$ , with  $a'_{T^L} = 1$ .

*Claim 2:*  $a'_t = 1$ .

*Proof:* Suppose  $a'_t = 0$ . Then by (C.8) and Lemma 3,  $a'_{t+1} = 0$ . But then again by (C.8) and Lemma 3,  $a'_{t+2} = 0$ , and using induction we arrive at the conclusion that  $a'_{T^L} = 0$ . Contradiction.  $\parallel$

Since  $a'_{T^L} = 1$  and  $a'_t = 1$ , by the most work property of  $\hat{\mathbf{a}}$ , there must exist a period  $m < \hat{t}$  such that  $a'_m = 0$ . Take the largest such period:

period:	...	$m$	$m + 1$	...	$\hat{t} - 1$	$\hat{t}$	$\hat{t} + 1$	$\hat{t} + 2$	...	$T^L$
$\hat{\mathbf{a}}$ :	...	1	1	...	1	1	0	0	...	0
$\mathbf{a}''$ :	...	1	1	...	1	0	0	0	...	0
$\mathbf{a}'$ :	...	0	1	...	1	1	1	1	...	1

Applying Lemma 3 to  $\mathbf{a}''$  and  $\mathbf{a}'$  yields  $\sum_{s=m}^{\hat{t}-1} l^L_s (1 - \lambda^H)^{\sum_{n=m+1}^s a'_n} = 0$ . Hence, there exists an optimal action plan  $\mathbf{a}'''$  obtained from  $\mathbf{a}'$  by switching  $a'_m$  and  $a'_t$ . But then the optimality of  $\mathbf{a}'''$  contradicts  $a'''_t = 0$ ,  $a'''_{T^L} = 1$ , (C.8), and Lemma 3.

Step 2: We now prove the Lemma's claim for  $\hat{t} + 1$ . That is, we show that there exists an optimal action plan, call it  $\hat{\mathbf{a}}^{\hat{t}+1}$ , such that for any  $s$ ,  $\hat{a}^{\hat{t}+1}_s = \hat{a}_s \iff s \notin \{\hat{t}, \hat{t} + 1\}$ . Suppose, to contradiction, that the claim is false. Then, by Lemma 3,  $l^L_{\hat{t}} < 0$ . Using Step 1, there is some  $\tau > \hat{t}$  that satisfies the Lemma's claim; let  $\mathbf{a}^\tau$  be the corresponding optimal action plan (which is identical to  $\hat{\mathbf{a}}$  in exactly all periods except from  $\hat{t}$  and  $\tau$ ). Since by Lemma 1 there exists an optimal action plan, call it  $\mathbf{a}'$ , with  $a'_{\hat{t}+1} = 1$ , Lemma 3 and  $l^L_{\hat{t}} < 0$  imply  $a'_t = 1$ . By the most work property of  $\hat{\mathbf{a}}$ , there must exist a period  $m < \hat{t}$  such that  $a'_m = 0$ . Take the largest such period:

period:	...	$m$	$m + 1$	...	$\hat{t} - 1$	$\hat{t}$	$\hat{t} + 1$	...	$\tau$	$\tau + 1$	...
$\hat{\mathbf{a}}$ :	...	1	1	...	1	1	0	...	0	0	...
$\mathbf{a}^\tau$ :	...	1	1	...	1	0	0	...	1	0	...
$\mathbf{a}'$ :	...	0	1	...	1	1	1	...	1	1	...

Applying Lemma 3 to  $\mathbf{a}^\tau$  and  $\mathbf{a}'$  yields  $\sum_{s=m}^{\hat{t}-1} l^L_s (1 - \lambda^H)^{\sum_{n=m+1}^s a'_n} = 0$ . Hence, there exists an optimal action plan  $\mathbf{a}''$  obtained from  $\mathbf{a}'$  by switching  $a'_m$  and  $a'_t$ . But then the optimality of  $\mathbf{a}''$  contradicts  $a''_t = 0$ ,  $a''_{\hat{t}+1} = 1$ ,  $l^L_{\hat{t}} < 0$ , and Lemma 3.

Step 3: Finally, we use induction to prove that the Lemma's claim is true for any  $s > \hat{t} + 1$ . (Note the claim is true for  $\hat{t} + 1$  by Step 2.) Take any  $t + 1 \in \{\hat{t} + 2, \dots, T^L\}$ . Assume the claim is true for  $s = \hat{t} + 2, \dots, t$ . We show that the claim is true for  $t + 1$ .

By Step 2 and the induction hypothesis, there exists an optimal action plan,  $\hat{\mathbf{a}}^t$ , such that for any  $s$ ,  $\hat{a}^t_s = \hat{a}_s \iff s \notin \{\hat{t}, t\}$ . We shall show that there exists an optimal action plan,  $\hat{\mathbf{a}}^{t+1}$ , such that for any  $s$ ,  $\hat{a}^{t+1}_s = \hat{a}_s \iff s \notin \{\hat{t}, t + 1\}$ . Suppose, to contradiction, that the claim is false. Note that Step 2, the induction hypothesis, and Lemma 3 imply  $l^L_s = 0$  for all  $s = \hat{t}, \dots, t - 1$ . It thus follows from Lemma 3 and the claim being false that  $l^L_t < 0$ . By Lemma 1 there exists an optimal action plan, call it  $\mathbf{a}'$ , with  $a'_{t+1} = 1$ . Then Lemma 3 and  $l^L_t < 0$  imply that  $a'_t = 1$ .

*Claim 3:*  $a'_s = 1$  for all  $s = \hat{t}, \dots, t - 1$ .

*Proof:* Suppose to contradiction that  $a'_{s^*} = 0$  for some  $s^* \in \{\hat{t}, \dots, t - 1\}$ . Then since  $l^L_s = 0$  for all  $s = \hat{t}, \dots, t - 1$ , by Lemma 3, there exists an optimal action plan,  $\mathbf{a}''$ , obtained from  $\mathbf{a}'$  by switching  $a'_{s^*}$  and  $a'_t$ . But then the optimality of  $\mathbf{a}''$  contradicts  $a''_t = 0$ ,  $a''_{t+1} = 1$ ,  $l^L_t < 0$ , and Lemma 3.  $\parallel$

Hence, we obtain  $a'_s = 1$  for all  $s = \hat{t}, \dots, t + 1$ , and by the most work property of  $\hat{\mathbf{a}}$ , there must exist a period  $m < \hat{t}$  such that  $a'_m = 0$ . Take the largest such period:

period:	...	$m$	$m + 1$	...	$\hat{t} - 1$	$\hat{t}$	$\hat{t} + 1$	...	$t$	$t + 1$	...
$\hat{\mathbf{a}}$ :	...	1	1	...	1	1	0	...	0	0	...
$\hat{\mathbf{a}}^t$ :	...	1	1	...	1	0	0	...	1	0	...
$\mathbf{a}'$ :	...	0	1	...	1	1	1	...	1	1	...

Applying Lemma 3 to  $\hat{\mathbf{a}}^t$  and  $\mathbf{a}'$  yields  $\sum_{s=m}^{\hat{t}-1} l_s^L (1-\lambda^H)^{\sum_{n=m+1}^s a_n'} = 0$ . Since  $l_s^L = 0$  for all  $s = \hat{t}, \dots, t-1$ , we obtain  $\sum_{s=m}^{t-1} l_s^L (1-\lambda^H)^{\sum_{n=m+1}^s a_n'} = 0$ . Hence, by Lemma 3, there exists an optimal action plan  $\mathbf{a}''$  obtained from  $\mathbf{a}'$  by switching  $a_m'$  and  $a_t'$ . But then the optimality of  $\mathbf{a}''$  contradicts  $a_t'' = 0$ ,  $a_{t+1}'' = 1$ ,  $l_t^L < 0$ , and Lemma 3.  $\parallel$

**Lemma 8.** *If  $C^L$  is an optimal contract for the low type with a most-work optimal stopping strategy for the high type, then  $C^L$  is a onetime-penalty contract.*

*Proof.* Fix  $C^L$  per the Lemma's assumptions. Let  $\hat{\mathbf{a}}$  and  $\hat{t}$  be as defined in the statement of Lemma 7. Then, it immediately follows from Lemma 7 and Lemma 3 that  $l_t^L = 0$  for all  $t \in \{\hat{t}, \hat{t}+1, \dots, T^L-1\}$ . We use induction to prove that  $l_t^L = 0$  for all  $t < \hat{t}$ .

Assume  $l_t^L = 0$  for all  $t \in \{m+1, m+2, \dots, T^L-1\}$  for  $m < \hat{t}$ . We will show that  $l_m^L = 0$ . First,  $l_m^L > 0$  is not possible because then  $\sum_{s=m}^{\hat{t}} l_s^L (1-\lambda^H)^{\sum_{n=m+1}^s \hat{a}_n} > 0$  (by Lemma 7 and the inductive assumption), contradicting the optimality of  $\hat{\mathbf{a}}$  and Lemma 3. Second, we claim  $l_m^L < 0$  is not possible. Suppose, to contradiction, that  $l_m^L < 0$ . Then  $(IC_a^L)$  is slack at  $m$  and, by Lemma 2, there exists an optimal plan  $\mathbf{a}'$  with  $a_m' = 0$ . Now by Lemma 3, Lemma 7, and the inductive assumption,  $a_s' = 0$  for all  $s \geq m$ . Hence,  $\beta_t^H(\mathbf{a}') > \beta_t^H(\hat{\mathbf{a}})$ , and thus the optimality of  $\hat{\mathbf{a}}$  implies that  $\mathbf{a}'$  is suboptimal at  $\hat{t}$ , a contradiction.  $\parallel$

*Acknowledgments.* We thank Andrea Attar, Patrick Bolton, Tilman Börgers, Pierre-André Chiappori, Bob Gibbons, Alex Frankel, Zhiguo He, Supreet Kaur, Alessandro Lizzeri, Suresh Naidu, Derek Neal, Alessandro Pavan, Andrea Prat, Canice Prendergast, Jonah Rockoff, Andy Skrzypacz, Lars Stole, Pierre Yared, various seminar and conference audiences, and anonymous referees and the Co-editor for helpful comments. We also thank Johannes Hörner and Gustavo Manso for valuable discussions of the paper. Sébastien Turban provided excellent research assistance. Kartik gratefully acknowledges the hospitality of and funding from the University of Chicago Booth School of Business during a portion of this research; he also thanks the Sloan Foundation for financial support through an Alfred P. Sloan Fellowship.

## Supplementary Data

Supplementary Materials are available at *Review of Economic Studies* online.

## REFERENCES

- ATKIN, D., CHAUDHRY, A., CHAUDRY, S., KHANDELWAL, A. K. and VERHOOGEN, E. (2015), "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan", Unpublished.
- BARON, D. P. and BESANKO, D. (1984), "Regulation and Information in a Continuing Relationship", *Information Economics and Policy*, **1**, 267–302.
- BARRETT, C., BACHKE, M., BELLEMARE, M., MICHELSON, H., NARAYANAN, S. and WALKER, T. (2012), "Smallholder Participation in Contract Farming: Comparative Evidence from Five Countries", *World Development*, **40**, 715–730.
- BATTAGLINI, M. (2005), "Long-Term Contracting with Markovian Consumers", *American Economic Review*, **95**, 637–658.
- BEAMAN, L., KARLAN, D., THUYSBAERT, B. and UDRY, C. (2015), "Selection into Credit Markets: Evidence from Agriculture in Mali", Unpublished.
- BERGEMANN, D. and HEGE, U. (1998), "Venture Capital Financing, Moral Hazard, and Learning", *Journal of Banking & Finance*, **22**, 703–735.
- BERGEMANN, D. and HEGE, U. (2005), "The Financing of Innovation: Learning and Stopping", *RAND Journal of Economics*, **36**, 719–752.
- BERGEMANN, D. and VÄLIMÄKI, J. (2008), "Bandit Problems", in Durlauf S. N. and Blume L. E., (eds), *The New Palgrave Dictionary of Economics* (Palgrave Macmillan, Basingstoke).
- BESLEY, T. and CASE, A. (1993), "Modeling Technology Adoption in Developing Countries", *American Economic Review, Papers and Proceedings*, **83**, 396–402.
- BHASKAR, V. (2012), "Dynamic Moral Hazard, Learning and Belief Manipulation", Unpublished.
- BHASKAR, V. (2014), "The Ratchet Effect Re-Examined: A Learning Perspective", Unpublished.
- BIAIS, B., MARIOTTI, T., ROCHET, J.-C. and VILLENEUVE, S. (2010), "Large Risks, Limited Liability, and Dynamic Moral Hazard", *Econometrica*, **78**, 73–118.
- BOBTSCHEFF, C. and LEVY, R. (2015), "More Haste, Less Speed: Signaling Through Investment Timing", Unpublished.
- BOLESZAVSKY, R. and SAID, M. (2013), "Progressive Screening: Long-Term Contracting with a Privately Known Stochastic Process", *Review of Economic Studies*, **80**, 1–34.

- BONATTI, A. and HÖRNER, J. (2011), "Collaborating", *American Economic Review*, **101**, 632–663.
- BONATTI, A. and HÖRNER, J. (2015), "Career Concerns with Exponential Learning", *Theoretical Economics*, forthcoming.
- BUNNIN, B. (1983), *Author Law & Strategies*. (Berkeley, CA: Nolo).
- CHASSANG, S. (2013), "Calibrated Incentive Contracts", *Econometrica*, **81**, 1935–1971.
- CONLEY, T. G. and UDRY, C. R. (2010), "Learning about a New Technology: Pineapple in Ghana", *American Economic Review*, **100**, 35–69.
- CREMER, J. and MCLEAN, R. P. (1985), "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent", *Econometrica*, **53**, 345–361.
- CREMER, J. and MCLEAN, R. P. (1988), "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions", *Econometrica*, **56**, 1247–1257.
- DEMARZO, P. and SANNIKOV, Y. (2011), "Learning in Dynamic Incentive Contracts", Unpublished.
- EDERER, F. (2013), "Incentives for Parallel Innovation", Unpublished.
- ESŐ, P. and SZENTES, B. (2015), "Dynamic Contracting: An Irrelevance Theorem", *Theoretical Economics*, forthcoming.
- FEDER, G., JUST, R. and ZILBERMAN, D. (1985), "Adoption of Agricultural Innovations in Developing Countries: A Survey", *Economic Development and Cultural Change*, **33**, 255–298.
- FOSTER, A. D. and ROSENZWEIG, M. R. (2010), "Microeconomics of Technology Adoption", *Annual Review of Economics*, **2**, 395–424.
- FOWLER, M. (1985), "The 'Satisfactory Manuscript' Clause in Book Publishing Contracts", *Columbia VLA Journal of Law & the Arts*, **10**, 119–152.
- FRICK, M. and ISHII, Y. (2015), "Innovation Adoption by Forward-Looking Social Learners", Unpublished.
- GERARDI, D. and MAESTRI, L. (2012), "A Principal-Agent Model of Sequential Testing", *Theoretical Economics*, **7**, 425–463.
- GERSHKOV, A. and PERRY, M. (2012), "Dynamic Contracts with Moral Hazard and Adverse Selection", *Review of Economic Studies*, **79**, 268–306.
- GOMES, R., GOTTLIEB, D. and MAESTRI, L. (2015), "Experimentation and Project Selection: Screening and Learning", Unpublished.
- GUO, Y. (2014), "Dynamic Delegation of Experimentation", Unpublished.
- HALAC, M., KARTIK, N. and LIU, Q. (2015), "Contests for Experimentation", Unpublished.
- HE, Z., WEI, B., YU, J. and GAO, F. (2014), "Optimal Long-Term Contracting with Learning", Unpublished.
- HOLMSTRÖM, B. and MILGROM, P. R. (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica*, **55**, 303–328.
- HÖRNER, J. and SAMUELSON, L. (2013), "Incentives for Experimenting Agents", *RAND Journal of Economics*, **44**, 632–663.
- JACK, B. K., OLIVA, P., SEVEREN, C., WALKER, E. and BELL, S. (2014), "Technology Adoption under Uncertainty: Take up and Subsequent Investment in Zambia", Unpublished.
- KELLER, G., RADY, S. and CRIPPS, M. (2005), "Strategic Experimentation with Exponential Bandits", *Econometrica*, **73**, 39–68.
- KELSEY, J. (2013), "Private Information and the Allocation of Land Use Subsidies in Malawi. *American Economic Journal: Applied Economics*, **5**, 113–135.
- KLEIN, N. (2012), "The Importance of Being Honest", Unpublished.
- KUZYK, R. (2006), "Inside Publishing", *Poets & Writers*, **34**, 71–78.
- KWON, S. (2013), "Dynamic Moral Hazard with Persistent States", Unpublished.
- LAFFONT, J.-J. and MARTIMORT, D. (2001), *The Theory of Incentives: The Principal-Agent Model* (Princeton, NJ: Princeton University Press).
- LAFFONT, J.-J. and TIROLE, J. (1988), "The Dynamics of Incentive Contracts", *Econometrica*, **56**, 1153–1175.
- LEWIS, T. R. (2011), "A Theory of Delegated Search for the Best Alternative", Unpublished.
- LEWIS, T. R. and OTTAVIANI, M. (2008), "Search Agency", Unpublished.
- MANSO, G. (2011), "Motivating Innovation", *Journal of Finance*, **66**, 1823–1860.
- MASON, R. and VÄLIMÄKI, J. (2011), "Dynamic Moral Hazard and Stopping", Unpublished.
- MINOT, N. (2007), "Contract Farming in Developing Countries: Patterns, Impact, and Policy Implications", *Case Study 6–3 of the Program, Food Policy for Developing Countries: the Role of Government in the Global Food System*.
- MIYATA, S., MINOT, N. and HU, D. (2009), "Impact of Contract Farming on Income: Linking Small Farmers, Packers, and Supermarkets in China. *World Development*, **37**, 1781–1790.
- MORONI, S. (2015), "Experimentation in Organizations", Unpublished.
- OBARA, I. (2008), "The Full Surplus Extraction Theorem with Hidden Actions", *The B.E. Journal of Theoretical Economics*, **8**, 1–28.
- OWEN, L. (2013). *Clark's Publishing Agreements: A Book of Precedents: Ninth Edition* (London: Bloomsbury Professional).
- PAVAN, A., SEGAL, I., and TOIKKA, J. (2014), "Dynamic Mechanism Design: A Myersonian Approach", *Econometrica*, **82**, 601–653.
- PRAT, J. and JOVANOVIĆ, B. (2014), "Dynamic Contracts when Agent's Quality is Unknown", *Theoretical Economics*, **9**, 865–914.

- RIORDAN, M. H. and SAPPINGTON, D. E. M. (1988), "Optimal Contracts with Public Ex Post Information", *Journal of Economic Theory*, **45**, 189–199.
- ROTHSCHILD, M. (1974), "A Two-Armed Bandit Theory of Market Pricing", *Journal of Economic Theory*, **9**, 185–202.
- SANNIKOV, Y. (2007), "Agency Problems, Screening and Increasing Credit Lines", Unpublished.
- SANNIKOV, Y. (2013), "Moral Hazard and Long-Run Incentives", Unpublished.
- SUDDATH, C. (2012), "Penguin Group Sues Writers Over Book Advances", *Bloomberg Businessweek*, September 27.
- SURI, T. (2011), "Selection and Comparative Advantage in Technology Adoption", *Econometrica*, **79**, 159–209.