Learning from Failures:

Optimal Contracts for Experimentation and Production[*]

Fahad Khalil, Jacques Lawarree
Alexander Rodivilov[**]

July 15, 2020

Abstract: Before embarking on a project, a principal must often rely on an agent to learn about its profitability. We model this learning as a two-armed bandit problem and highlight the interaction between learning (experimentation) and production. We derive the optimal contract for both experimentation and production when the agent has private information about the efficiency of experimentation. This private information in the experimentation stage generates asymmetric information in the production stage even though there was no disagreement about the profitability of the project at the outset. The degree of asymmetric information is endogenously determined by the length of the experimentation stage. An optimal contract uses the length of experimentation, the production scale, and the timing of payments to screen the agent. We find that *over*-experimentation and *over*-production can be used to reduce the agent's rent. An efficient type is rewarded early since he is more likely to succeed in experimenting, while an inefficient type is rewarded at the very end of the experimentation stage. This result is robust to the introduction of ex post moral hazard.

Keywords: Information gathering, optimal contracts, strategic experimentation.
JEL: D82, D83, D86.

[*] We are thankful for the helpful comments of Nageeb Ali, Dirk Bergemann, Renato Gomes, Marina Halac, Navin Kartik, Qingmin Liu, David Martimort, Dilip Mookherjee, Larry Samuelson, Edward Schlee, and Leeat Yariv. We also thank the editor, Alessandro Pavan, the associate editor, and two referees.
[**] khalil@uw.edu, lawarree@uw.edu, Department of Economics, University of Washington; arodivil@stevens.edu, School of Business, Stevens Institute of Technology

1. Introduction

Before embarking on a project, it is important to learn about its profitability to determine its optimal scale. Consider, for instance, shareholders (principal) who hire a manager (agent) to work on a new project.[1] To determine its profitability, the principal asks the agent to explore various ways to implement the project by experimenting with alternative technologies. Such experimentation might demonstrate the profitability of the project. A longer experimentation phase allows the agent to better determine the project's profitability but is also costly and delays production. Therefore, there is interdependence between the duration of the experimentation and the optimal scale of the project.

An additional complexity arises if the agent is privately informed about the efficiency of experimentation. If experimentation is not efficient, a poor result from experiments only provides weak evidence of the actual profitability of the project. However, if the principal is misled by the agent into believing that experimentation is highly efficient, she may form erroneous beliefs about the project's profitability. A trade-off appears for the principal. More experimentation may provide better information about the profitability of the project but can also increase asymmetric information about its expected profitability, which leads to information rent for the agent in the production stage.

In this paper, we derive the optimal contract for an agent who conducts both experimentation and production. We model the experimentation stage as a two-armed bandit problem.[2] At the outset, the principal and agent are symmetrically informed that production cost can be high or low. The contract determines the duration of the experimentation stage. Success in experimentation is assumed to take the form of uncovering "bad news", i.e., the agent finds out whether production cost is high.[3] After success, experimentation stops, and production occurs based on the knowledge that production cost is high.[4] If experimentation continues

[1] Applications include the testing of new drugs, medical specialists performing tests to diagnose and treat patients, the adoption of new technologies or products, the identification of new investment opportunities, consumer search, contract farming, etc. See Krähmer and Strausz (2011) and Manso (2011) for other relevant examples.
[2] The exponential bandit model has been widely used as a canonical model of learning: see Bolton and Harris (1999), Keller, Rady, and Cripps (2005), or Bergemann and Välimäki (2008).
[3] For example, stage 1 of a drug clinical trial looks for bad news by testing the safety of the drug. See for instance Halac and Kremer (2020) and references therein for experimentation models using bad news.
[4] We present our main insights by assuming that success in experimentation is publicly observed but show that our key results hold even if the agent could hide success. We also show that technically the case of success being good news mirrors that of bad news and our key insights continue to hold.

without success, the expected cost decreases, and both principal and agent become more optimistic about project profitability. We say that the experimentation stage fails if the agent never learns the true cost.

A key contribution of our model is to study how the asymmetric information generated during experimentation impacts production, and how production decisions affect experimentation.[5] At the end of the experimentation stage, there is a production decision, which generates information rent as it depends on what was learned during experimentation. Relative to the growing literature on incentives for experimentation, reviewed below, the novelty of our approach is to study the optimal contract for *both* experimentation and production. Focusing on incentives to experiment, the literature has equated project implementation with success in experimentation. In contrast, we study the impact of learning from failures on the optimal contract for production and experimentation. The production stage naturally introduces an option value of learning that affects optimal incentives for both experimentation and production. Thus, our analysis highlights the impact of endogenous asymmetric information on optimal decisions ex post, which is not present in a model without a production stage.

In our model, the agent privately knows his efficiency, the probability of success in any given period of the experimentation stage conditional on the true cost being high. When experimentation fails (i.e., no bad news is uncovered), an efficient agent pretending to be inefficient will have a lower expected cost of production compared to the principal. Mistakenly believing the agent is inefficient, the principal will then overcompensate him in the production stage. Therefore, an efficient agent must be paid a rent to prevent him from understating his efficiency. An important element of our setting is that the inefficient type may also get a rent. The reason is that the efficiency parameter directly enters the principal's objective function. As a result, we have what is called a common values problem in contract theory.[6] It is known that in such models both efficient and inefficient types can get rent in equilibrium due to a strong conflict between the principal's preference for efficiency and the screening role of contracts.

We summarize our main results next. First, in a model with experimentation and production, we show that *over*-experimentation relative to the first-best can be an optimal

[5] Intertemporal contractual externality across agency problems also plays an important role in Arve and Martimort (2016).

[6] See, e.g., Laffont and Martimort (2002).

screening strategy for the principal, whereas under-experimentation is the standard result in existing models of experimentation.[7] There are two main reasons the principal may ask the agent to over experiment. Since increasing the duration of experimentation helps raise the chance of success, the first reason to ask the agent to *over* experiment is that it makes it less likely for the agent to fail and exploit the asymmetry of information about expected cost. The second reason is that, even when experimentation fails, increasing the duration of experimentation may reduce the agent's rent. This is due to our finding that the difference in expected costs between the principal and the misreporting agent is non-monotonic in time. When an inefficient agent misreports, he faces a gamble: the benefit comes from the chance to collect the efficient agent's rent, but he faces a cost from the risk of being undercompensated at the production stage if experimentation fails since he is relatively less optimistic than the principal. As the difference in expected costs determine the benefit and cost of misreporting, we show that increasing the duration of experimentation can help both reduce the benefit as well as increase the cost of misreporting.

Second, we show that experimentation also influences the choice of output in the production stage. If experimentation succeeds, the output is at the first best level since there is no difference in beliefs regarding the true cost after success. However, if experimentation fails, the output is distorted to reduce the rent of the agent. Since the efficient agent always gets a rent, we expect, and indeed find, that the output of the inefficient agent is distorted downward. This is reminiscent of a standard adverse selection problem.

We find another effect: the output of the efficient agent is distorted *upward*. This is the case when the inefficient agent also commands a rent, which is a new result due to the interaction between the experimentation and production stages. To understand this result, recall that the inefficient type faces a cost of misreporting because he is under-compensated at the production stage when experimentation fails. The principal can increase this cost of misreporting by asking the efficient type to produce more. A higher output for the efficient agent makes it costlier for the inefficient agent who must produce more output with higher expected cost.

[7] To the best of our knowledge, ours is the first paper in the literature that predicts over-experimentation. The reason is that over-experimentation might reduce the rent in the production stage, non-existent in standard models of experimentation.

Third, to screen the agents, the principal distributes the information rent as rewards to the agent at different points in time. When both types obtain a rent, each type's comparative advantage on obtaining successes or failures determines a unique optimal contract. Each type is rewarded for events which are relatively more likely for him. It is optimal to reward the efficient agent *at the beginning* and the inefficient agent *at the very end* of the experimentation stage. Interestingly, the inefficient agent is rewarded after failure if the experimentation stage is relatively short and after success in the last period otherwise.[8] Our result suggests that the principal is more likely to tolerate failures in industries where the cost of an experiment is relatively high; for example, this is the case in oil drilling. In contrast, if the cost of experimentation is low (like on-line advertising) the principal will rely on rewarding the agent after success.

While we study a model of pure adverse selection, it is clear that most real-world situations will encompass a mix of adverse selection and moral hazard.[9] In an extension section, we introduce ex post moral hazard by assuming that success is privately observed by the agent. Since the agent is rewarded for late success or failure, a moral hazard rent has to be provided in every period to reveal success, in addition to the previously derived asymmetric information rent. It remains optimal to provide exaggerated rewards for the efficient type at the beginning and for the inefficient type at the end of experimentation even under ex post moral hazard.

Related literature. Our paper builds on two strands of the literature. First, it is related to the literature on principal-agent contracts with endogenous information gathering before production.[10] It is typical in this literature to consider static models, where an agent exerts effort to gather information relevant to production. By modeling this effort as experimentation, we introduce a dynamic learning aspect, and especially the possibility of asymmetric learning by different agents. We contribute to this literature by characterizing the structure of incentive schemes in a dynamic learning stage. Importantly, in our model, the principal can determine the

[8] In an insightful paper, Manso (2011), argues that golden parachutes and managerial entrenchment, which seem to reward or tolerate failure, can be effective for encouraging corporate innovation (see also, Ederer and Manso (2013), and Sadler (2020)). A combination of stock options with long vesting periods and option repricing are evidence of rewarding late success. Our analysis suggests that such practices have screening properties in situations where innovators differ in expertise.

[9] By suppressing moral hazard, our framework allows us to highlight the screening properties of the optimal contract that deals with both experimentation and production in a tractable model.

[10] Early papers are Crémer and Khalil (1992), Lewis and Sappington (1997), and Crémer, Khalil, and Rochet (1998), while Krähmer and Strausz (2011) contains recent citations.

degree of asymmetric information by choosing the length of the experimentation stage, and over-experimentation can be optimal.

To model information gathering, we rely on the growing literature on contracting for experimentation following Bergemann and Hege (1998, 2005). Most of that literature has a different focus and characterizes incentive schemes for addressing moral hazard during experimentation but does not consider adverse selection.[11] Recent exceptions that introduce adverse selection are Gomes, Gottlieb and Maestri (2016) and Halac, Kartik and Liu (2016).[12] In Gomes, Gottlieb and Maestri, there is two-dimensional hidden information, where the agent is privately informed about the quality of the project as well as a private cost of effort for experimentation. They find conditions under which the second hidden information problem can be ignored. Halac, Kartik and Liu (2016) have both moral hazard and hidden information. They extend the moral hazard-based literature by introducing hidden information about expertise in the experimentation stage to study how asymmetric learning by the efficient and inefficient agents affects the bonus that needs to be paid to induce the agent to work.[13]

We add to the literature by showing that asymmetric information created during experimentation affects production, which in turn introduces novel aspects to the incentive scheme for experimentation. Unlike the rest of the literature, we find that over-experimentation relative to the first best, and rewarding an agent after failure can be optimal to screen the agent.

The rest of the paper is organized as follows. In Section 2, we present the base bad-news model under adverse selection with public success. In Section 3, we present the optimal contract and our main results. In Section 4, we study ex post moral hazard where the agent can hide success, and the case where success is good news.

[11] See also Horner and Samuelson (2013).

[12] See also Gerardi and Maestri (2012) for another model where the agent is privately informed about the ex ante quality of the project.

[13] They show that, without the moral hazard constraint, the first best can be reached. In our model, we impose a limited liability constraint instead of a moral hazard constraint.

2. The Model (Learning bad news)

A principal hires an agent to implement a project of a variable size. Both the principal and agent are risk neutral and have a common discount factor $\delta \in (0,1]$. It is common knowledge that the marginal cost of production can be low or high, i.e., $c \in \{\underline{c}, \overline{c}\}$, with $0 < \underline{c} < \overline{c}$. The probability that $c = \overline{c}$ is denoted by $\beta_0 \in (0,1)$. Before the *production stage*, the agent can gather information regarding the production cost. We call this the *experimentation stage*.

The experimentation stage

During the experimentation stage, the agent gathers information about the cost of the project. The experimentation stage takes place over time, $t \in \{1,2,3, \dots .T\}$, where T is the maximum length of the experimentation stage and is determined by the principal.[14] In each period t , experimentation costs $\gamma > 0$, and we assume that this cost γ is paid by the principal at the end of each period. We assume that it is optimal to experiment at least once under full information.[15]

In the main part of the paper, information gathering takes the form of uncovering bad news (see Section 4.2 for the case of good news). If the cost is high, the agent learns it with probability $\lambda \in (0,1)$ in each period $t \leq T$.[16] If the agent learns that the cost is high (*bad news*) in a period t, we say that the experimentation was successful. To focus on the screening features of the optimal contract, we assume for now that the agent cannot hide evidence of the cost being high. In Section 4.1, we will revisit this assumption and study a model with both adverse selection and ex post moral hazard.

We say that experimentation has failed if the agent fails to uncover high cost in all T periods. Even if the experimentation stage results in failure, the expected cost is updated, so there is much to learn from failure. We turn to this next.

We assume that the agent is privately informed about his experimentation efficiency represented by λ. Therefore, the principal faces an adverse selection problem even though all parties assess the same expected cost at the outset. The principal and agent may update their

---

[14] Modeling time as discrete is convenient in our setting as we will see that each type receives rent only once at the beginning or the end of the experimentation phase (Section 3.2.1).

[15] When deriving the optimal contract under asymmetric information, we allow the principal to choose zero experimentation for either type.

[16] If $\lambda^\theta = 1$, the first failure would be a perfect signal regarding the project quality.

beliefs differently during the experimentation stage. The efficiency parameter λ determines the agent's type, and we will refer to an agent with high or low efficiency as a high or low-type agent. With probability ν, the agent is a high type, $\theta = H$. With probability $(1-\nu)$, he is a low type, $\theta = L$. Thus, we define the learning parameter with the type superscript: $\lambda^\theta = Pr(type\ \theta\ learns\ c = \overline{c} | c = \overline{c})$, where $0 < \lambda^L < \lambda^H < 1$.

If experimentation fails to reveal high cost in a period, agents with different types form different beliefs about the expected cost of the project. We denote by β_t^θ the updated belief of a θ-type agent that the cost is actually high at the beginning of period t given $t-1$ failures. For period $t > 1$, we have $\beta_t^\theta = \frac{\beta_{t-1}^\theta(1-\lambda^\theta)}{\beta_{t-1}^\theta(1-\lambda^\theta)+(1-\beta_{t-1}^\theta)}$, which in terms of β_0 is $\beta_t^\theta = \frac{\beta_0(1-\lambda^\theta)^{t-1}}{\beta_0(1-\lambda^\theta)^{t-1}+(1-\beta_0)}$. The θ-type agent's expected cost at the beginning of period t is then given by: $c_t^\theta = \beta_t^\theta \overline{c} + \left(1-\beta_t^\theta\right) \underline{c}$.

Three aspects of learning are worth noting. First, after each period of failure during experimentation, there is more *optimism* that the true cost is low, i.e., β_t^θ falls. The expected cost c_t^θ decreases and converges to $\underline{c}$. Second, for the same number of failures during experimentation, the expected cost is higher for the low type, i.e., $c_t^L > c_t^H$. An example of how the expected cost c_t^θ converges to $\underline{c}$ for each type is presented in Figure 1 below.

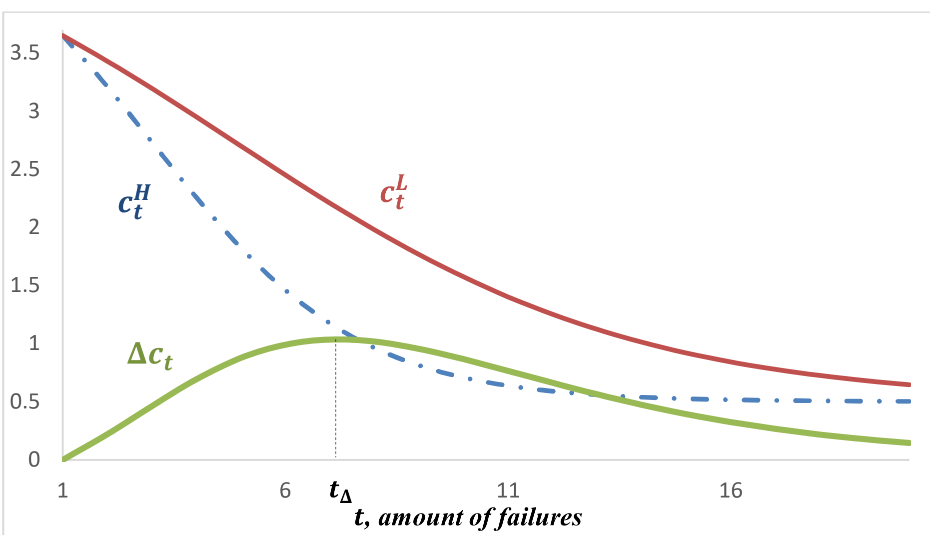

Figure 1. Expected cost with $\lambda^H = 0.35$, $\lambda^L = 0.2$, $\beta_0 = 0.7$, $\underline{c} = 0.5$, $\overline{c} = 5$.

Third, we also note the important property that the difference in expected costs, $\Delta c_t = c_t^L - c_t^H > 0$, is a *non-monotonic* function of time: initially increasing and then decreasing, reaching a maximum at time period t_Δ.[17] Intuitively, each type starts with the same expected cost $\beta_0\overline{c} + (1-\beta_0)\underline{c}$. The expected costs diverge as each type of agent updates differently, but they eventually have to converge to $\underline{c}$. As λ^H becomes relatively close to λ^L, the Δc_t function becomes flatter, moving t_Δ to the right, which makes the increasing part of Δc_t relatively larger.

The production stage

After the experimentation stage ends, we assume that production takes place immediately. The principal's value of the project is $V(q)$, where $q > 0$ is the size of the project. The function $V(\cdot)$ is strictly increasing, strictly concave, twice differentiable on $(0, +\infty)$, and satisfies the Inada conditions.[18] The size of the project and the payment to the agent are determined in the contract offered by the principal before the experimentation stage takes place. If experimentation reveals that cost is high in a period $t \leq T$, experimentation stops, and production takes place based on $c = \overline{c}$.[19] If experimentation fails, i.e., there is no success in uncovering high cost during the experimentation stage, production occurs based on the expected cost.[20]

The contract

Before the experimentation stage takes place, the principal offers the agent a menu of dynamic contracts. Without loss of generality, we use a direct truthful mechanism, where the agent is asked to announce his type, denoted by $\hat{\theta}$. A contract specifies, for each type of agent, the length of the experimentation stage, the size of the project, and a transfer as a function of whether or not the agent succeeded while experimenting. We assume the agent cannot quit and must produce once he has accepted the contract.[21] In terms of notation, in the case of success we

---

[17] There exists a unique time period t_Δ such that Δc_t achieves the highest value at this time period, where

$$t_\Delta = arg \max_{1 \leq t \leq T} \frac{(1-\lambda^L)^{t-1} - (1-\lambda^H)^{t-1}}{(1-\beta_0+\beta_0(1-\lambda^H)^{t-1})(1-\beta_0+\beta_0(1-\lambda^L)^{t-1})}.$$

[18] Without the Inada conditions, it may be optimal to shut down the production of the high type after failure if expected cost is high enough. In such a case, neither type will get a rent.

[19] In this model, there is no reason for the principal to continue to experiment once she learns that cost is high.

[20] We assume that the agent will learn the exact cost later, but it is not contractible.

[21] There are many examples where there are penalties and legal restrictions on an agent prematurely terminating a contract. For instance, contracts often provide for penalties when one party breaches the contract and quits (see for

include $\overline{c}$ as an argument in the wage and output for each t. In the case of failure, we include the updated expected cost after failure, denoted by $c^{\hat{\theta}}_{T^{\hat{\theta}}+1}$.[22] A contract is defined formally by

$$\varpi^{\hat{\theta}} = \left(T^{\hat{\theta}}, \left\{w_t^{\hat{\theta}}(\overline{c}), q_t^{\hat{\theta}}(\overline{c})\right\}_{t=1}^{T^{\hat{\theta}}}, \left\{w^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}+1}\right), q^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}+1}\right)\right\}\right),$$

where $T^{\hat{\theta}}$ is the maximum duration of the experimentation stage for the announced type $\hat{\theta}$, $w_t^{\hat{\theta}}(\overline{c})$ and $q_t^{\hat{\theta}}(\overline{c})$ are the agent's wage and the output if he succeeded in period $t \leq T^{\hat{\theta}}$ and $w^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}+1}\right)$ and $q^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}+1}\right)$ are the agent's wage and the output if the agent fails $T^{\hat{\theta}}$ consecutive times. An agent of type θ, announcing his type as $\hat{\theta}$, receives expected utility $U^{\theta}\left(\varpi^{\hat{\theta}}\right)$ at time zero from a contract $\varpi^{\hat{\theta}}$:

$$U^{\theta}\left(\varpi^{\hat{\theta}}\right) = \beta_0 \sum_{t=1}^{T^{\hat{\theta}}} \delta^t \left(1-\lambda^{\theta}\right)^{t-1} \lambda^{\theta} \left(w_t^{\hat{\theta}}(\overline{c}) - \overline{c} q_t^{\hat{\theta}}(\overline{c})\right)$$

$$+\delta^{T^{\hat{\theta}}}\left(1-\beta_0+\beta_0\left(1-\lambda^{\theta}\right)^{T^{\hat{\theta}}}\right)\left(w^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}+1}\right) - c^{\theta}_{T^{\hat{\theta}}+1} q^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}+1}\right)\right).$$

Conditional on the actual cost being high, which happens with probability β_0, the probability of succeeding for the first time in period $t \leq T^{\hat{\theta}}$ is given by $\left(1-\lambda^{\theta}\right)^{t-1}\lambda^{\theta}$. Experimentation fails if either the cost is low ($c = \underline{c}$), which happens with probability $1-\beta_0$, or, if the agent fails $T^{\hat{\theta}}$ times despite $c = \overline{c}$, which happens with probability $\beta_0(1-\lambda^{\theta})^{T^{\hat{\theta}}}$.

To summarize, the timing is as follows:

(1) The agent learns his type θ.

(2) The principal offers a contract to the agent. In case the agent rejects the contract, the game is over, and both parties get payoffs normalized to zero; if the agent accepts the contract, the game proceeds to the experimentation stage with duration as specified in the contract.

(3) The experimentation stage begins.

instance U.S. Uniform Civil Code §2-713: Buyer's Damages for Non-delivery or Repudiation). Because of such penalties, there is a cost for the agent to quit after the experimentation phase. Our assumption is that the cost is high enough to deter the agent from quitting. In our model, we will see that, since the contract covers expected cost in equilibrium, only a lying agent would want to quit.

[22] Since the principal pays for the experimentation cost, the agent is not paid if he does not succeed in any $t < T^{\hat{\theta}}$.

(4) If the agent learns that $c = \overline{c}$, the experimentation stage stops, and production occurs with output and transfers as specified in the contract.

In case the experimentation stage ends in failure, production occurs with output and transfers as specified in the contract.

Before solving for the optimal contract, we discuss three key features of our model. First, our pure adverse selection model assumes that there is limited scope for moral hazard during learning. For instance, the availability of low-cost monitoring technologies, such as cameras, make effort easy to observe and limits the scope of moral hazard. Another example is when the learning phase is based on set protocols and legal requirements that must be followed. Consider, for instance, the case of medical specialists such as surgeons who diagnose and treat injuries or illnesses. Patients often go through a series of tests (experimentation) before the treatment (production) begins. Specialists such as surgeons must follow protocols and regulations for healthcare activities required by the health insurance company, Medicare or HMO (principal). In addition, they are required by law to record patient medical histories and to retain detailed case histories. There is also little room for skipping tests or altering results since this behavior might be simply illegal and a surgeon might be subject to prosecution. Such behavior would also violate the Hippocratic Oath. [23]

Second, there is an alternative interpretation of the adverse selection problem, where the efficiency parameter, λ, is tied to a *project* rather than the agent. Our analytical framework would remain unchanged. An example comes from the literature on medical contracts where third-party payers must take into account that physicians have private information about their patients. As Dranove (1987) explains, based on characteristics observable upon admission, physicians can sometimes know whether a patient will be easy or difficult to diagnose (experimentation phase). Part of this literature is focused on misallocating patients to diagnostic related groups (DRG) based on financial incentives. An important example is the rise in C-Section deliveries associated with the introduction of incentive payments. While studying this case, Currie-MacLeod (2017) also stress that physicians have private information when

[23] Similar protocols and legal requirements also exist for prosecuting attorneys evaluating evidence before deciding on charges, and pharmaceutical companies testing new drugs before commercializing them. For instance, "Crime Scene Investigation: A Guide for Law Enforcement" published by the U.S. Department of Justice in 2013 provides a detailed description of steps and procedures an enforcement official must follow. The FDA dictates how many patients to test, age/gender/blood type distributions, and how to document the results.

diagnosing a patient (decision making about the appropriate treatment) because they privately observe some patient characteristics. In our model, this would be consistent with a learning parameter purely based on a given patient's characteristics (instead of the agent's). Similarly, a prosecutor could know if a specific case is easy or hard to solve and prosecute, or whether he is well matched for the type of case. A drug company manager could know the prospects of establishing the efficacy of a specific drug under consideration. Thus, in those examples, the probability of success in experimentation (i.e., the type λ) is tied to the project rather than the agent.

Third, to highlight the role of asymmetric information on decisions ex post, we model a production stage ex post that is performed by the same agent who experiments. This is common in a wide range of applications such as surgeons/medical specialists diagnosing patients before deciding on a treatment, prosecuting attorneys evaluating evidence before deciding on charges, and pharmaceutical companies testing new drugs before commercializing them. As already noted by Laffont and Tirole (1988), in the presence of cost uncertainty and risk aversion, separating the two tasks may not be optimal. Moreover, hiring one agent for experimentation and another one for production might lead to an informed principal problem. For example, in case the first agent provides negative evidence about the project's profitability, the principal may benefit from hiding this information from the second agent to keep him more optimistic about the project.

The First Best Benchmark

Suppose the agent's type θ is common knowledge *before* the principal offers the contract. The first-best termination dates and outputs are found by maximizing the principal's profit:

$$\beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta \left(V\left(q_t^\theta(\overline{c})\right) - \overline{c} q_t^\theta(\overline{c}) - \Gamma_t\right)$$

$$+\delta^{T^\theta}\left(1-\beta_0+\beta_0\left(1-\lambda^\theta\right)^{T^\theta}\right)\left(V\left(q^\theta\left(c^\theta_{T^\theta+1}\right)\right) - c^\theta_{T^\theta+1} q^\theta\left(c^\theta_{T^\theta+1}\right) - \Gamma_{T^\theta}\right),$$

where the cost of experimentation borne by the principal is $\Gamma_t = \frac{\sum_{s=1}^{t} \delta^s \gamma}{\delta^t}$.

If the agent succeeds, the efficient output will be produced such that $V'\left(q^\theta_{t^\theta}(\overline{c})\right) = \overline{c}$ for any t^θ. In case the agent fails, the efficient output is based on the *expected* cost at the end of experimentation, such that $V'\left(q^\theta\left(c^\theta_{T^\theta+1}\right)\right) = c^\theta_{T^\theta+1}$.

The first-best termination date T_{FB}^{θ} is given by the highest t^{θ} such that the following condition holds:

$$\beta_{t^{\theta}}^{\theta}\lambda^{\theta}\left[V\left(q_{t^{\theta}}^{\theta}(\overline{c})\right)-\overline{c}q_{t^{\theta}}^{\theta}(\overline{c})\right]+\left(1-\beta_{t^{\theta}}^{\theta}\lambda^{\theta}\right)\left[V\left(q^{\theta}\left(c_{t^{\theta}+1}^{\theta}\right)\right)-c_{t^{\theta}+1}^{\theta}q^{\theta}\left(c_{t^{\theta}+1}^{\theta}\right)\right]$$
$$\geq \gamma+\left[V\left(q^{\theta}\left(c_{t^{\theta}}^{\theta}\right)\right)-c_{t^{\theta}}^{\theta}q^{\theta}\left(c_{t^{\theta}}^{\theta}\right)\right].$$

The benefit of experimentation comes from lowering expected cost as long as experimentation continues to fail, but there is a risk of uncovering that cost is high. Extending the experimentation stage by one additional period costs γ, and a θ-type agent can learn that $c=\overline{c}$ with probability $\beta_{t^{\theta}}^{\theta}\lambda^{\theta}$. Note that the first-best termination date of the experimentation stage T_{FB}^{θ} is a *non-monotonic* function of the agent's type: there exists a unique value of λ^{θ} called $\hat{\lambda}$, such that:[24]

$$\frac{dT_{FB}^{\theta}}{d\lambda^{\theta}}>0 \text{ for } \lambda^{\theta}<\hat{\lambda} \text{ and } \frac{dT_{FB}^{\theta}}{d\lambda^{\theta}}\leq 0 \text{ for } \lambda^{\theta}\geq\hat{\lambda}.$$

This non-monotonicity in the first-best termination dates is a result of two countervailing forces.[25] Since the high type is relatively more efficient at uncovering bad news (conditional on the actual cost being high), it suggests that the principal should allow the high type to experiment longer. However, with repeated failures, the high type agent also becomes relatively more optimistic about cost being low and the expected cost does not decrease as much after failure as with the low type. This can be seen by looking at the probability of success conditional on reaching period t, given by $\beta_0\left(1-\lambda^{\theta}\right)^{t-1}\lambda^{\theta}$. In Figure 2, we see that this conditional probability of success for the high type becomes smaller than that for the low type at some period t. We will use later the important property that the relative likelihood of success $\left(\frac{\beta_0\left(1-\lambda^{H}\right)^{t-1}\lambda^{H}}{\beta_0(1-\lambda^{L})^{t-1}\lambda^{L}}\right)$ is decreasing over time.

Given these two countervailing forces, the first-best termination date for the high type agent can be shorter or longer than that of the low type depending on the parameters of the problem.[26] The first-best termination date is increasing in the agent's type for small values of λ^{θ}

[24] See Supplementary Appendix F.
[25] A similar intuition can be found in Halac, Kartik and Liu (2016) in a model without production.
[26] For example, if $\lambda^{L}=0.26$, $\lambda^{H}=0.32$, $\underline{c}=0.5$, $\overline{c}=5$, $\beta_0=0.25$, $\delta=0.9$, $\gamma=1$, and $V=15\ln q$, then the first-best termination date for the high type agent is $T_{FB}^{H}=6$, whereas it is optimal to allow the low type agent to experiment for five periods, $T_{FB}^{L}=5<T_{FB}^{H}=6$. However, if we now change λ^{H} to 0.6, the low type agent is allowed to experiment less, that is, $T_{FB}^{H}=4>T_{FB}^{L}=5$.

when the first force (relative efficiency) dominates, but becomes decreasing for larger values when the second force (relative optimism) becomes dominant.

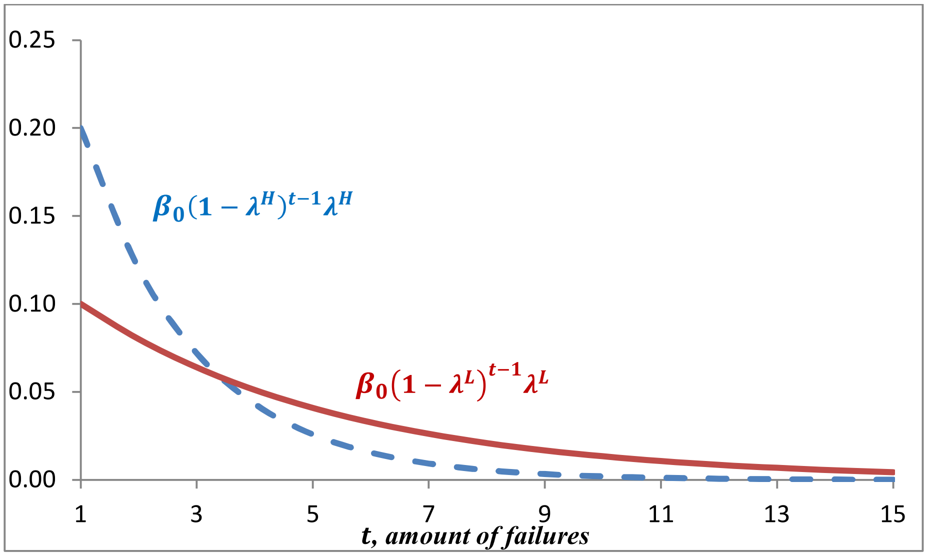

Figure 2. Probability of success with $\lambda^H = 0.4$, $\lambda^L = 0.2$, $\beta_0 = 0.5$.

3. The optimal Contract under Asymmetric information

We now return to the main model where asymmetric information arises because the two types learn asymmetrically in the experimentation stage. The optimal contract will have to satisfy the following incentive compatibility constraints for all θ and $\hat{\theta}$:

$$(IC) \qquad U^\theta(\varpi^\theta) \geq U^\theta(\varpi^{\hat{\theta}}).$$

To simplify the exposition, we define by y_t^θ the wage net of cost to the θ type who *succeeds* in period t, and by x^θ the wage net of the expected cost to the θ type who *failed* during the entire experimentation stage:

$$y_t^\theta \equiv w_t^\theta(\overline{c}) - \overline{c} q_t^\theta(\overline{c}) \text{ for } 1 \leq t \leq T^\theta,$$

$$x^\theta \equiv w^\theta(c_{T^\theta+1}^\theta) - c_{T^\theta+1}^\theta q^\theta(c_{T^\theta+1}^\theta).$$

We also denote with P_T^θ the probability that a θ-type agent fails during the T periods of the experimentation stage:

$$P_T^\theta = 1 - \beta_0 + \beta_0\left(1-\lambda^\theta\right)^T.$$

Using this notation, we can rewrite the two incentive constraints as:

$(IC^{L,H})$ $$\beta_0 \sum_{t=1}^{T^L} \delta^t \left(1-\lambda^L\right)^{t-1}\lambda^L y_t^L + \delta^{T^L} P_{T^L}^L x^L$$
$$\geq \beta_0 \sum_{t=1}^{T^H} \delta^t \left(1-\lambda^L\right)^{t-1}\lambda^L y_t^H + \delta^{T^H} P_{T^H}^L\left[x^H - \Delta c_{T^H+1} q^H\left(c_{T^H+1}^H\right)\right],$$

$(IC^{H,L})$ $$\beta_0 \sum_{t=1}^{T^H} \delta^t \left(1-\lambda^H\right)^{t-1}\lambda^H y_t^H + \delta^{T^H} P_{T^H}^H x^H$$
$$\geq \beta_0 \sum_{t=1}^{T^L} \delta^t \left(1-\lambda^H\right)^{t-1}\lambda^H y_t^L + \delta^{T^L} P_{T^L}^H\left[x^L + \Delta c_{T^L+1} q^L\left(c_{T^L+1}^L\right)\right].$$

We also assume that the agent must be paid his expected production costs whether experimentation succeeds or fails.[27] Therefore, we introduce the following limited liability constraints:

$\left(LLS_t^\theta\right)$ $y_t^\theta \geq 0$ for $t \leq T^\theta$,

$\left(LLF_{T^\theta}^\theta\right)$ $x^\theta \geq 0$,

where the S and F denote success and failure, respectively.

Now we can state the principal's problem. The principal maximizes the following objective function

$$E_\theta\left[\beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1}\lambda^\theta \left[V\left(q_t^\theta(\overline{c})\right) - \overline{c} q_t^\theta(\overline{c}) - \Gamma_t - y_t^\theta\right]\right.$$
$$\left. + \delta^{T^\theta} P_{T^\theta}^\theta \left[V\left(q^\theta\left(c_{T^\theta+1}^\theta\right)\right) - c_{T^\theta+1}^\theta q^\theta\left(c_{T^\theta+1}^\theta\right) - \Gamma_{T^\theta} - x^\theta\right]\right]$$

subject to (LLS_t^L), $\left(LLF_{T^L}^L\right)$, (LLS_t^H), $\left(LLF_{T^H}^H\right)$, $(IC^{L,H})$, and $(IC^{H,L})$, where the cost of experimentation borne by the principal is $\Gamma_t = \frac{\sum_{s=1}^t \delta^s \gamma}{\delta^t}$.

---

[27] Examples of legal restrictions on transfers that exemplify limited liability in contracts are ubiquitous (bankruptcy laws, minimum wage laws). See, e.g., Krähmer and Strausz (2015) for more examples. This constraint is not meant to represent a no-wealth setting. In particular, the agent is not protected off the equilibrium path. Technically, without limited liability, the principal can receive first best profit since success during experimentation is a random event correlated with the agent's type (Crémer-McLean (1985)). For simplicity, we require the transfers to cover expected cost, which means that the contract is analogous to the well-known cost-plus contracts in the procurement literature.

Both (IC) may be binding

We now explain why both (IC) constraints can be binding. The high type's $(IC^{H,L})$ constraint is binding for the standard reason in adverse selection models. A high type has an incentive to claim to be a low type in order to collect the higher transfer given to the low type to cover the higher expected cost following failure.[28] That is, the RHS of $(IC^{H,L})$ is strictly positive since $\Delta c_{T^L+1} = c^L_{T^L+1} - c^H_{T^L+1} > 0$.

In addition, the low-type's $(IC^{L,H})$ constraint can also be binding, which is atypical in adverse selection models.[29] Experimentation introduces two key features to the standard adverse selection problem. First, a common value problem emerges since the agent's type λ^θ directly enters the principal's objective function by determining the probability of success and failure. Second, the difference in expected costs Δc_T, that determines the incentive to misreport, depends (non-monotonically) on the termination date T, which is endogenous in our model. In common-value settings, there can be a strong conflict between the principal's desire for one type to experiment longer for pure efficiency reasons and the monotonicity condition imposed by asymmetric information leading to both (IC) binding (see, e.g., Laffont and Martimort (2002), page 53). We explain in more detail next.

To highlight the role of these two new features due to experimentation, we first briefly outline a benchmark case without experimentation. Consider a standard second-best problem where the private information parameter is the *expected* marginal cost (e.g., Baron and Myerson (1982)). To facilitate comparison with our model with experimentation, we assume that a high type has a lower expected cost, i.e., $c^H < c^L$, and that production occurs based on expected cost. Denoting $\Delta c = c^L - c^H > 0$, the wage to the agent by $w(c^\theta)$ and the output by $q(c^\theta)$, the two incentive constraints can be written in equilibrium as:

$$\left(IC_b^{H,L}\right) \qquad w(c^H) - c^H q(c^H) = \Delta c\, q(c^L)$$

$$\left(IC_b^{L,H}\right) \qquad w(c^L) - c^L q(c^L) = 0 > \Delta c\, q(c^L) - \Delta c\, q(c^H),$$

[28] We prove this result as Claim 1 in Appendix A.

[29] There are other examples of low type (IC) binding in contract theory, but it is due to different reasons. For instance, in the case of countervailing incentives, a low type may have incentive to pretend to be a high type due to outside options.

where the subscript "b" refers to this benchmark without experimentation.[30]

In this benchmark, the low type has no incentive to misreport his type. When the low type lies, he collects the rent of the high type, $\Delta c\, q(c^L)$, as part of the transfer $w(c^H)$. However, he then must produce the higher output $q(c^H)$ while being undercompensated relative to the high type as his true expected cost c^L exceeds that of the high type: $c^L > c^H$. Therefore, the $\left(IC_b^{L,H}\right)$ is never binding. We want to highlight two main differences with our experimentation model. First Δc, the high type's cost advantage, is exogenous and is also identical to the low type's cost disadvantage when he has to produce $q(c^H)$. In our experimentation model, the cost advantage is endogenous. Second the agent's type does not explicitly appear in the principal's objective function (see footnote 30) while it does in the experimentation model.

We now return to the main model with experimentation. First, note that the low type's benefit from misreporting his type is a *gamble*, with a positive and a negative part. The positive part again stems from a chance to claim the rent of the high type. This part is positively related to Δc_{T^L+1} adjusted by the output and the corresponding probability of collecting the high type's rent. There is also a negative part. It arises from the risk of failing in experimentation and having to produce while being undercompensated. Indeed, when experimentation fails, the low type is paid as a high type whose expected cost is lower. The negative part of the gamble is positively related to Δc_{T^H+1} adjusted by the output and the probability of having to produce after failure. The $(IC^{L,H})$ is binding when the gamble becomes positive, i.e., when its positive part dominates the negative one. Unlike the benchmark model, the positive part can dominate the negative one because the termination dates are endogenous.

The termination dates T^L and T^H play a key role in the sign of the gamble since they determine Δc_{T^L+1} and Δc_{T^H+1}. We prove that, when the duration of the experimentation stage is identical for both types ($T^H = T^L$), the gamble is negative, and only $(IC^{H,L})$ is binding. Intuitively, the magnitudes of Δc_{T^L+1} and Δc_{T^H+1} are the same just as in the benchmark without

---

[30] Denoting again the probability of a high type by ν, the principal maximizes $\upsilon\left[V\left(q(c^H)\right) - w(c^H)\right] + (1-\upsilon)\left[V\left(q(c^L)\right) - w(c^L)\right]$, such that, for $\theta, \hat{\theta} \in \{L, H\}$, $w\left(c^\theta\right) - c^\theta q\left(c^\theta\right) \geq 0$ to induce participation, and $w\left(c^\theta\right) - c^\theta q\left(c^\theta\right) \geq w\left(c^{\hat{\theta}}\right) - c^\theta q\left(c^{\hat{\theta}}\right)$ to induce truth telling. The solution to this problem is well known, where only the low type's output is distorted downwards and only the high type gets a positive informational rent.

experimentation, where only $(IC_b^{H,L})$ is binding.[31] However, having the same duration for both types is suboptimal.

There are two reasons why the principal may want a different duration for each type. First, as explained in Section 2, first-best efficiency can require either $T_{FB}^H > T_{FB}^L$ or $T_{FB}^H < T_{FB}^L$ depending on the values of λ^θ. Second, the incentive constraints indicate that the principal can lower the rent of both types by altering the termination dates T^L and T^H. In this common-value setting, there can be a conflict if, for instance, rent minimization necessitates the low type to experiment less than the high type, but first best efficiency requires the opposite. This leads to both upward and downward ICs to be binding. In our model, the conflict is complicated by the fact that the Δc_t is not monotonic and that termination dates and quantities are both screening variables. However, the key intuition can be provided focusing on the termination dates.

To illustrate, assume that $(IC^{L,H})$ is not binding, i.e., the low-type's gamble is negative. Then, standard arguments would suggest there is no reason to distort the first-best stopping time $T^H = T_{FB}^H$ for the high type, but distorting T^L from its first-best level T_{FB}^L can lower the rent of the high type. First-best efficiency can conflict with the order of termination dates required by the two IC constraints. For example, if efficiency requires $T_{FB}^L > T_{FB}^H$ and, in addition, Δc_t is increasing in t (see Figure 1), we have $\Delta c_{T_{FB}^L+1} > \Delta c_{T_{FB}^H+1}$.[32] If the value of T^L under asymmetric information remains close to its first best value, then $\Delta c_{T^L+1} > \Delta c_{T_{FB}^H+1}$, and the gamble becomes positive. That would violate $(IC^{L,H})$, and thus both (IC) are binding.[33]

Therefore, the principal trades off first-best efficiency in experimentation with the rent in the production stage and this may result in both types getting positive rent. This trade-off is absent in models of experimentation without an ex post production stage.

3.2.1. The timing of the payments: rewarding failure or early/late success?

Having established that both types may receive rent, we now study the principal's choice of timing of rewards to each type: should the principal reward early or late success in the experimentation stage? Should she reward failure? We will see that the relative likelihood of success for a high type at a specific period t plays a critical role in screening.

[31] See Supplementary Appendix E.
[32] This is the case when the λs are large but relatively close to each other.
[33] We provide in Supplementary Appendix B sufficient conditions for $(IC^{L,H})$ to be binding.

There are two cases to consider. First, when $(IC^{L,H})$ is not binding, $y_t^L = x^L = 0$, the optimal contract is not unique, and the principal can use any combination of y_t^H and x^H to satisfy the binding $(IC^{H,L})$: there is no restriction on when and how the principal pays the rent to the high type as long as $\beta_0 \sum_{t=1}^{T^H} \delta^t (1-\lambda^H)^{t-1} \lambda^H y_t^H + \delta^{T^H} P_{T^H}^H x^H = \delta^{T^L} P_{T^L}^H \Delta c_{T^L+1} q^L(c_{T^L+1}^L)$. Therefore, in this case, the principal can reward either early or late success, or even failure.[34] Second, when $(IC^{L,H})$ is binding, the optimal contract is unique. The high type's rent is paid in the very first period while the low type's rent is paid at the end. Whether it is paid after success or failure depends on the length of the experimentation stage, which depends on the cost of experimentation. Both cases are described in the following Proposition.

Proposition 1. *Optimal timing of payments.*

When only the high type's IC is binding

The low type gets no rent. There is no restriction on when to reward the high type.

When both types' IC are binding

The high type agent is rewarded for early success (in the very first period)

$$y_1^H > 0 = x^H = y_t^H \text{ for all } t > 1.$$

The low type agent is rewarded

(i) after failure if the cost of experimentation is large $(\gamma > \gamma^*)$*:*

$$x^L > 0 = y_t^L \text{ for all } t \leq T^L, \text{ and}$$

(ii) after success in the last period if the cost of experimentation is small $(\gamma < \gamma^*)$*:*

$$y_{T^L}^L > 0 = x^L = y_t^L \text{ for all } t \leq T^L.$$

Proof: See Appendix A.

We start by analyzing the case where the principal rewards the low type agent after success and then explain that it is optimal to do so when experimentation cost is small. We first show in Appendix A that, if the principal rewards success, it will be in at most one period.[35] Since the relative likelihood ratio of success, $\frac{\beta_0(1-\lambda^H)^{t-1}\lambda^H}{\beta_0(1-\lambda^L)^{t-1}\lambda^L}$, is strictly decreasing in t, the principal chooses to postpone rewarding the low type until the very last period, T^L, to minimize the high type's incentive to misreport. Thus, we have $y_t^L = 0$ for all $t < T^L$, while $y_{T^L}^L \geq 0$.

---

[34] See Case A in Appendix A.

[35] See in Lemmas 2 and B.2.2 in Appendix A.

To see why the principal may want to reward the low type agent after failure at T^L, we need to compare the relative likelihood of ratios of success $\left(\frac{\beta_0(1-\lambda^H)^{t-1}\lambda^H}{\beta_0(1-\lambda^L)^{t-1}\lambda^L}\right)$ and failure $\left(\frac{P^H_{T^L}}{P^L_{T^L}}\right)$ for a lying high type. We show in Appendix A that there is a unique period $\hat{T}^L$ such that the two relative probabilities are equal:[36]

$$\frac{(1-\lambda^H)^{\hat{T}^L-1}\lambda^H}{(1-\lambda^L)^{\hat{T}^L-1}\lambda^L} \equiv \frac{P^H_{T^L}}{P^L_{T^L}}.$$

In any period $t < \hat{T}^L$, depicted in Figure 3 below, the high type is relatively more likely to succeed than fail compared to the low type. For $t > \hat{T}^L$, the opposite is true. Thus, if the experimentation stage is short, $T^L < \hat{T}^L$, the principal will pay the rent to the low type by rewarding failure since the high type is relatively more likely to succeed during the experimentation stage. Otherwise, the principal rewards the low type for success in the last period.

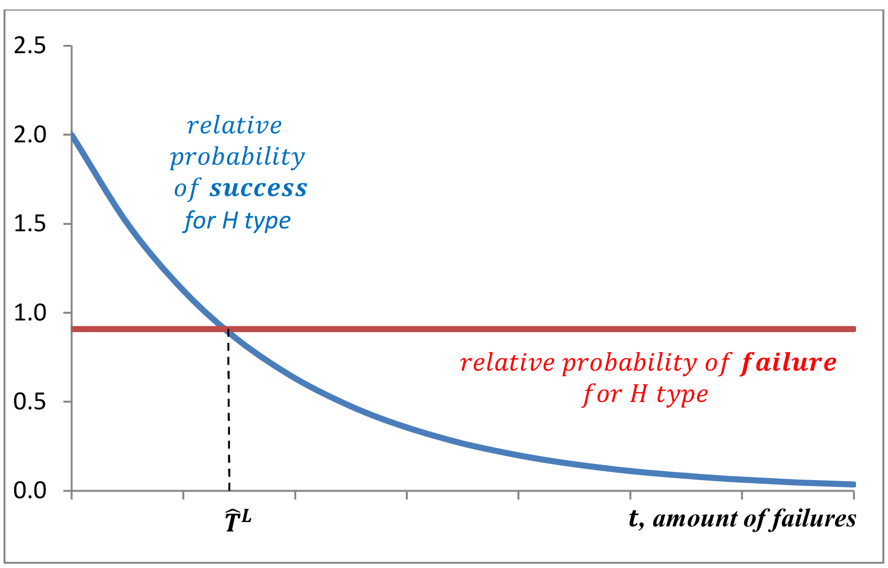

Figure 3. Relative probability of success/failure with $\lambda^H = 0.4$, $\lambda^L = 0.2$, $\beta_0 = 0.5$.

The optimal value of T^L is inversely related to the cost of experimentation γ. In Appendix A, we prove in Lemma 6 that there exists a unique value of γ^* such that $T^L < \hat{T}^L$ for any $\gamma > \gamma^*$. Therefore, when the cost of experimentation is high ($\gamma > \gamma^*$), the length of experimentation will be short, and it will be optimal for the principal to reward the low type after

[36] See Lemma 1 in Appendix A for the proof.

failure. Intuitively, failure is a better instrument to screen the high type when experimentation cost is high. So, it is the adverse selection concern that makes it optimal to reward failure.

Finally, we show in Appendix A, that the principal will reward the high type for success in the first period only. This is the period when success is most likely to come from a high type than a low type.

3.2.2. The length of the experimentation period: optimality of *over*-experimentation

While the standard result in the experimentation literature is under-experimentation, we find that *over*-experimentation can occur when there is a production stage following experimentation. The reason why over-experimentation may be optimal is that it may reduce the rent in the production stage, non-existent in standard models of experimentation.[37] We explain this next in details.

There are two main reasons why the principal may ask the agent to over experiment. First, because λ directly enters the principal's objective function, over-experimentation increases the chances of success. Since the agent collects rent in our model due to a possibility of failure in the experimentation stage, the principal lowers the chances of paying rent to the agent with over-experimentation.

Second, even if the agent fails, increasing the duration of experimentation can help reduce the impact of asymmetric information and thus the agent's rent in the production stage. Over-experimentation can reduce the benefit of lying for both types through Δc_{T^L+1} and raise the cost of lying for the low type through Δc_{T^H+1}.

To find sufficient conditions for over-experimentation, we need to also consider the impact of the relative probabilities of failure and of the output on the rent. The following Proposition gives sufficient conditions for over-experimentation.

Proposition 2: *There exist* $\underline{\beta}_0 < 1$ *and* $\overline{\lambda}^H(\lambda^L) > \lambda^L$ *for any* $\lambda^L > \hat{\lambda}$*, such that there is over-experimentation in* T^H*, i.e.,* $T^H_{SB} \geq T^H_{FB}$*, when* $\beta_0 > \underline{\beta}_0$ *and* $\hat{\lambda} < \lambda^L < \lambda^H < \overline{\lambda}^H(\lambda^L)$.

Proof: See Appendix A.

[37] What is important is that a positive output is produced after failure even if the level is given exogenously ex ante. In a standard model of experimentation (see Halac, Kartik and Liu (2016) and references therein), the output after failure is zero, $q^\theta(c^\theta_{T^\theta+1}) \equiv 0$.

To understand the intuition behind these conditions, it is convenient to focus on the difference in expected costs at period T^H adjusted by the probability of failing in the experimentation stage after misreporting for the low type: $P^L_{T^H}\Delta c_{T^H+1}$. By increasing T^H, the principal can increase $P^L_{T^H}\Delta c_{T^H+1}$ which is proportional to the cost of lying for the low type (negative part of the gamble). Therefore, the principal may benefit from asking the high type to over experiment: $T^H \geq T^H_{FB}$.

Technically, since $P^L_{T^H}$ is decreasing in T^H and Δc_{T^H+1} is non-monotonic in T^H, for λ^L and λ^H relatively large (above $\hat{\lambda}$) but close to each other and a high value of β_0, the gamble is positive and the optimal termination date T^H is on the increasing part of the $P^L_t\Delta c_t$ function. Also, when β_0 is high, the monotonicity of $P^L_t\Delta c_t$ is determined by Δc_t, which is increasing if λ^L and λ^H are close to each other.[38]

For the intuition for over-experimentation in T^L, we can focus on the case where, at the optimum, T^L is on the decreasing part of Δc_t. Increasing T^L decreases Δc_{T^L+1} which is proportional to the benefit of lying for both the high type and the low type (positive part of the gamble). Therefore, the principal may benefit from asking the low type to over experiment: $T^L \geq T^L_{FB}$.[39]

3.2.3. The output: under- or over-production

When experimentation is successful, there is no asymmetric information and no reason to distort the output. Both types produce the first best output. When experimentation fails to reveal the cost, there is asymmetric information, and the principal will distort the output to limit the rent. This is a familiar result in contract theory. In a standard second-best contract à la Baron-Myerson, the type who receives rent produces the first best level of output while the type with no rent underproduces relative to the first best.

We find a similar result when only the high type's incentive constraint binds. The high type produces the first best output while the low type underproduces relative to the first best. To limit the rent of the high type, the low type is asked to produce a lower output.

---

[38] Recall that each type starts with the same expected cost $\beta_0\overline{c} + (1-\beta_0)\underline{c}$, which then diverge as each type of agent updates differently. With higher β_0, the expected costs continue to diverge for a longer period before eventually converging to $\underline{c}$.

[39] The exact conditions are given in Supplementary Appendix G. Note that there is a condition on the concavity of the function $V(\cdot)$. It is due to the presence of the output in determining whether the rent is increasing or decreasing, and it ensures that the rents are monotonic in T^L.

However, we find a new result when both IC are binding simultaneously. In this case, to limit the rent of the low type, the principal will *increase* the output of the high type and require *over*-production relative to the first best. To understand the intuition behind this result, recall that the rent of the low type mimicking the high type is a gamble with two components. The positive part is due to the rent promised to the high type after failure in the experimentation stage which is increasing in $q^L\left(c^L_{T^L+1}\right)$. Lowering this output decreases the positive component of the gamble. The negative part comes from the higher expected cost of producing the output required from the high type, and it is increasing in $q^H\left(c^H_{T^H+1}\right)$. Increasing the high-type's output after failure lowers the rent of the low type by increasing his cost of lying. We summarize the results in Proposition 3 below.

Proposition 3. *Optimal output.*

After success, each type produces at the first best level:

$$V'\left(q^\theta_t(\overline{c})\right) = \overline{c} \ \textit{for } t \leq T^\theta.$$

After failure, the low type underproduces relative to the first best output:

$$q^L_{SB}\left(c^L_{T^L+1}\right) < q^L_{FB}\left(c^L_{T^L+1}\right).$$

After failure, the high type overproduces:

$$q^H_{SB}\left(c^H_{T^H+1}\right) \geq q^H_{FB}\left(c^H_{T^H+1}\right).$$

Proof: See Appendix A.

4. Extensions

4.1. Success might be hidden: *ex post* moral hazard

Our base model without moral hazard allowed us to highlight the screening properties of the timing of rewards and show that delaying the reward or paying after failure is optimal. We now explore how the payment scheme could change in the presence of moral hazard. If there were moral hazard concerns in every period, we would expect rent in every period. As we noted before, modeling both hidden effort and adverse selection in experimentation is beyond the scope of this paper. However, we can introduce *ex post* moral hazard by relaxing our assumption that the outcome of experiments in each period is publicly observable. We show that our key insights remain.

Specifically, we assume that success is privately observed by the agent, and that an agent who finds success in some period j can choose to reveal it at any period $t \geq j$. Thus, we assume that success generates hard information that can be presented to the principal when desired, but it cannot be fabricated.

In addition to the (LL) and (IC) constraints, the optimal scheme must now satisfy the ex post moral hazard constraints described below. The $\left(EMH^{\theta}\right)$ constraint makes it unprofitable for the agent to hide success in the last period. The $\left(EMP_t^{\theta}\right)$ constraint makes it unprofitable to postpone revealing success in prior periods. The two together imply that the agent cannot gain by postponing or hiding success.

$\left(EMH^{\theta}\right)$ $\qquad y_{T^{\theta}}^{\theta} \geq x^{\theta} + \left(c_{T^{\theta}+1}^{\theta} - \overline{c}\right) q^{\theta}\left(c_{T^{\theta}+1}^{\theta}\right)$ for $\theta = H, L$, and

$\left(EMP_t^{\theta}\right)$ $\qquad y_t^{\theta} \geq \delta y_{t+1}^{\theta}$ for $t \leq T^{\theta} - 1$.

We provide intuition for our results here, while the formal proofs are in Supplementary Appendix C. First, both constraints are slack for a high type contract. By paying rent after the first success, as under public success, the principal can address both incentives of hiding and postponing.[40] The (EMH^H) is slack since there is no incentive to hide success, which is learning that cost is high $\overline{c}$. This is because the principal would only compensate for a lower cost, $c_{T^H+1}^{H} q^H\left(c_{T^H+1}^{H}\right)$, if experimentation fails. The (EMP_t^H) is also slack since there is no incentive to postpone revealing success if the principal only pays a rent after success in the first period.

Consider next the two moral hazard constraints for a low type contract. Recall that under public success the low type is rewarded either for the very last success ($y_{T^L}^{L} > 0$) or for failure ($x^L > 0$). Postponing the announcement of an early success makes the principal overly optimistic, $c_t^L < \overline{c}$, but allows the agent to collect rent at the end of experimentation. This introduces a moral hazard rent in every period. An increase of \$1 in x^L causes an increase of \$1 in $y_{T^L}^{L}$, which in turn causes an increase in all the previous y_t^L according to the discount factor. However, the benefit of delaying the reward or paying after failure for screening stems from the relative probabilities of success and failure between types, which are not affected by the ex post

[40] Recall that in the main model with public success, the high type was either rewarded for the very first success (Case B), or there were no restrictions on when to pay the high type (Case A). Therefore, it is without loss of generality to consider a contract where $y_1^H > 0 = x^H = y_t^H$ for $t > 1$.

moral hazard constraint above. When both $(IC^{H,L})$ and $(IC^{L,H})$ are binding, just as in Proposition 1, it is optimal to have exaggerated rewards at the two extremes of the experimentation phase, including reward after failure if the low type experiments for a relatively brief length of time.[41]

Having analyzed how the ex post moral hazard constraints affect the timing of payments, we next consider how those constraints interact with the other screening instruments. Since the agent's adverse selection rent is still determined by the difference in expected costs, which remains non-monotonic in time, we again find that over-experimentation and over-production can occur.

First consider the length of experimentation. Moral hazard makes it costlier to over-experiment. The longer the agent experiments, the costlier it is to deter hiding and postponing early success. Therefore, there is a tradeoff: by asking the agent to over experiment, the principal mitigates the adverse selection rent but increases the moral hazard rent. In Supplementary Appendix C, we show that over-experimentation remains optimal.

Second consider the impact of moral hazard on the output as a screening instrument. As can be seen from (EMH^{θ}), increasing the output $q^{\theta}\left(c^{\theta}_{T^{\theta}+1}\right)$ relaxes the moral hazard constraint. Therefore, by asking the agent to over produce after failure, the principal mitigates both the adverse selection and the moral hazard rents.

4.2. Learning good news

In this section, we show that technically the case of good news in experimentation mirrors the case of bad news. Thus, our main results survive if the object of experimentation is to seek good news, where success in an experiment means discovery of low cost $c = \underline{c}$. The parallel between bad news and good news is not difficult to explain. In both cases, the agent is looking for news. The types determine efficiency in obtaining this news. The contract gives incentives for each type of agent to reveal his type, not the actual news.

If success is not achieved in a particular period, the principal and agent both become more pessimistic (instead of optimistic in a bad news model). Also, as time goes by without learning that the cost is low, the expected cost becomes higher. In addition, the difference in expected costs is now negative, $\Delta c_t = c^L_t - c^H_t < 0$ since the low type is relatively more

[41] We formally show in the Supplementary Appendix C that $(IC^{L,H})$ may be binding. Since the ex post moral hazard constraints imply that the low type will receive rent, this rent may be sufficient to satisfy the $(IC^{L,H})$ constraint.

optimistic after the same duration without success. However, Δc_t remains non-monotonic in time and the reasons for over-experimentation remain unchanged.

Under asymmetric information about the agent's type, the intuition behind the key incentive problem is similar to that under learning bad news. The optimization problem mirrors the case for bad news, and we find results similar to those in Propositions 1, 2, and 3. We present these results formally in Supplementary Appendix D.

Finally, unlike in the case of bad news, if the agent is rewarded for success, he might have incentive to hide success in the last period as he will be overcompensated in the production phase.

5. Conclusion

In this paper, we have studied the interaction between experimentation and production where the length of the experimentation stage determines the degree of asymmetric information at the production stage. While there has been much recent attention on studying incentives for experimentation in two-armed bandit settings, details of the optimal production decision are typically suppressed to focus on incentives for experimentation. Each stage may impact the other in interesting ways and our paper is a step towards studying this interaction.

When there is an optimal production decision after experimentation, we find a new result that over-experimentation is a useful screening device. Likewise, over-production is also useful to mitigate the agent's information rent. By analyzing the stochastic structure of the dynamic problem, we clarify how the principal can rely on the relative probabilities of success and failure of the two types to screen them. The rent to a high type should come after early success and to the low type for late success. If the experimentation stage is relatively short, the principal has no recourse but to pay the low type's rent after failure, which is another novel result.

While our main section relies on publicly observed success, we show that our key insights survive if the agent can hide success. Then, there is ex post moral hazard, which implies that the agent is paid a rent in every period, but the screening properties of the optimal contract remain intact. Finally, we prove that our key insights do hold in both bad and good-news models.

References

Arve M. and Martimort D. (2016). Dynamic procurement under uncertainty: Optimal design and implications for incomplete contracts. *American Economic Review* 106 (11), 3238-3274.

Bergemann D. and Hege U. (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance* 22 (6-8), 703-735.

Bergemann D. and Hege U. (2005). The Financing of Innovation: Learning and Stopping. *Rand Journal of Economics* 36 (4), 719-752.

Bergemann D. and Välimäki J. (2008). Bandit Problems. In Durlauf, S., Blume, L. (Eds.), The New Palgrave Dictionary of Economics. 2nd edition. Palgrave Macmillan Ltd, Basingstoke, New York.

Bolton P. and Harris C. (1999). Strategic Experimentation. *Econometrica* 67(2), 349–374.

Currie J. and McLeod W. B. (2017). Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians. *Journal of Labor Economics*, 35(1), 1-43.

Crémer J. and Khalil F. (1992). Gathering Information before Signing a Contract. *American Economic Review* 82(3), 566-578.

Crémer J., Khalil F. and Rochet J.-C. (1997). Contracts and productive information gathering. *Games and Economic Behavior* 25(2), 174-193.

Dranove D. (1987). Rate-setting by diagnosis related groups and hospital specialization. *RAND Journal of Economics* 18(3), 417-427.

Ederer F. and Manso G. (2013). Is Pay-for-Performance Detrimental to Innovation? *Management Science*, 59 (7), 1496-1513.

Gerardi D. and Maestri L. (2012). A Principal-Agent Model of Sequential Testing. *Theoretical Economics* 7, 425–463.

Gomes R., Gottlieb D. and Maestri L. (2016). Experimentation and Project Selection: Screening and Learning. *Games and Economic Behavior* 96, 145-169.

Halac M., Kartik N. and Liu Q. (2016). Optimal contracts for experimentation. *Review of Economic Studies* 83(3), 1040-1091.

Halac M. and Kremer I. (2020). Experimenting with Career Concerns. *American Economic Journal: Microeconomics* 12(1), 260-288.

Horner J. and Samuelson L. (2013). Incentives for Experimenting Agents. *RAND Journal of Economics* 44(4), 632-663.

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

Krähmer D. and Strausz R. (2011). Optimal Procurement Contracts with Pre-Project Planning. *Review of Economic Studies* 78, 1015-1041.

Krähmer D. and Strausz R. (2015). Optimal Sales Contracts with Withdrawal Rights. *Review of Economic Studies* 82, 762–790.

Laffont J.-J. and Martimort D. (2002). *Theory of Incentives*. Princeton University Press.

Laffont J.-J. and Tirole J. (1988). The Dynamics of Incentive Contracts. *Econometrica* 56(5), 1153-1175.

Lewis T. and Sappington D. (1997). Information Management in Incentive Problems. *Journal of Political Economy* 105(4), 796-821.

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

Rosenberg D., Salomon A. and Vieille N. (2013). On games of strategic experimentation. *Games and Economic Behavior* 82, 31-51.

Sadler E. (2020). Dead Ends. *Working Paper*, March 2020.

Appendix A (Proofs of Claim 1, and Propositions 1, 2 and 3)

The Principal's Maximization Problem and Claim 1 with proof

We first characterize the optimal payment structure x^L, $\{y_t^L\}_{t=1}^{T^L}$, x^H and $\{y_t^H\}_{t=1}^{T^H}$ (Proposition 1) given the lengths of experimentation and the output levels. Then, we characterize the optimal length of experimentation, T^L and T^H (Proposition 2), and finally the optimal outputs $\{q_t^H(\overline{c})\}_{t=1}^{T^H}$, $q^H\left(c_{T^H}^H\right)$, $\{q_t^L(\overline{c})\}_{t=1}^{T^L}$ and $q^L\left(c_{T^L}^L\right)$ (Proposition 3).

Denote the expected surplus net of costs for $\theta = H, L$ by

$$\Omega^\theta\left(\varpi^\theta\right) = \beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta \left[V\left(q_t^\theta(\overline{c})\right) - \overline{c} q_t^\theta(\overline{c}) - \Gamma_t\right] +$$
$$\delta^{T^\theta} P_{T^\theta}^\theta \left[V\left(q^\theta\left(c_{T^\theta+1}^\theta\right)\right) - c_{T^\theta+1}^\theta q^\theta\left(c_{T^\theta+1}^\theta\right) - \Gamma_{T^\theta}\right].$$

The principal's optimization problem then is to choose contracts ϖ^H and ϖ^L to maximize the expected net surplus minus rent of the agent, subject to the respective IC and LL constraints given below:

$$Max\ E_\theta \left\{\Omega^\theta\left(\varpi^\theta\right) - \beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta y_t^\theta - \delta^{T^\theta} P_{T^\theta}^\theta x^\theta\right\} \text{ subject to:}$$

$(IC^{H,L})$ $\beta_0 \sum_{t=1}^{T^H} \delta^t (1-\lambda^H)^{t-1} \lambda^H y_t^H + \delta^{T^H} P_{T^H}^H x^H$

$\geq \beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^H)^{t-1} \lambda^H y_t^L + \delta^{T^L} P_{T^L}^H \left[x^L + \Delta c_{T^L+1} q^L\left(c_{T^L+1}^L\right)\right]$,

$(IC^{L,H})$ $\beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1} \lambda^L y_t^L + \delta^{T^L} P_{T^L}^L x^L$

$\geq \beta_0 \sum_{t=1}^{T^H} \delta^t (1-\lambda^L)^{t-1} \lambda^L y_t^H + \delta^{T^H} P_{T^H}^L \left[x^H - \Delta c_{T^H+1} q^H\left(c_{T^H+1}^H\right)\right]$,

(LLS_t^H) $y_t^H \geq 0$ for $t \leq T^H$,

(LLS_t^L) $y_t^L \geq 0$ for $t \leq T^L$,

$\left(LLF_{T^H}^H\right)$ $x^H \geq 0$,

$\left(LLF_{T^L}^L\right)$ $x^L \geq 0$.

In the rest of the proof, we now focus on the case where $T^L > 0$ and $T^H > 0$. If both T^L and T^H are equal to zero, there is no experimentation and production takes place under symmetric information based on expected costs. If $T^L = 0$ and $T^H > 0$, there is no rent for the high type or the low type. If $T^L > 0$ and $T^H = 0$, only $(IC^{H,L})$ is binding and we are in *Case A*, analyzed below. We begin to solve the problem by first proving the following claim.

Claim 1: The constraint $(IC^{H,L})$ is binding and the high type obtains a strictly positive rent.

Proof: If the $(IC^{H,L})$ constraint was not binding, it would be possible to decrease the payment to the high type unless (LLS_t^H) and (LLF_t^H) are binding, but in that case $(IC^{H,L})$ would be violated since $\Delta c_{T^L+1} q^L\left(c_{T^L+1}^L\right) > 0$. *Q.E.D.*

I. Optimal payment structure **(Proof of Proposition 1)**

First, we show that if the high type claims to be the low type, the high type is relatively more likely to succeed if the experimentation stage is smaller than a threshold level, $\hat{T}^L$. In terms of notation, we define $f_2(t, T^L) = \frac{P_{T^L}^H}{P_{T^L}^L}(1-\lambda^L)^{t-1}\lambda^L - (1-\lambda^H)^{t-1}\lambda^H$ to track the difference in the likelihood ratios of failure and success for two types.

Lemma 1: There exists a unique $\hat{T}^L > 1$, such that $f_2\left(\hat{T}^L, T^L\right) = 0$, and

$$f_2(t, T^L)\begin{cases} < 0 \text{ for } t < \hat{T}^L \\ > 0 \text{ for } t > \hat{T}^L \end{cases}$$

Proof: Note that $\frac{P_{T^L}^H}{P_{T^L}^L}$ is a ratio of the probability that the high type does not succeed to the probability that the low type does not succeed for T^L periods. At the same time, $\beta_0\left(1-\lambda^\theta\right)^{t-1}\lambda^\theta$ is the probability that the agent of type θ succeeds at period $t \le T^L$ of the experimentation stage and $\frac{\beta_0\left(1-\lambda^H\right)^{t-1}\lambda^H}{\beta_0(1-\lambda^L)^{t-1}\lambda^L} = \frac{\left(1-\lambda^H\right)^{t-1}\lambda^H}{(1-\lambda^L)^{t-1}\lambda^L}$ is a ratio of the probabilities of success at period t by two types. As a result, we can rewrite $f_2(t, T^L) > 0$ as

$$\frac{1-\beta_0+\beta_0\left(1-\lambda^H\right)^{T^L}}{1-\beta_0+\beta_0(1-\lambda^L)^{T^L}} > \frac{\left(1-\lambda^H\right)^{t-1}\lambda^H}{(1-\lambda^L)^{t-1}\lambda^L} \text{ for } 1 \le t \le T^L.$$

We will say that when $f_2(t, T^L) > 0$ (< 0) the high type is relatively more likely to fail (succeed) than the low type during the experimentation stage if he chooses a contract designed for the low type. There exists a unique time period $\hat{T}^L(T^L, \lambda^L, \lambda^H, \beta_0)$ such that $f_2\left(\hat{T}^L, T^L\right) = 0$, where uniqueness follows from $\frac{\left(1-\lambda^H\right)^{t-1}\lambda^H}{(1-\lambda^L)^{t-1}\lambda^L}$ being strictly decreasing in t and $\frac{\lambda^H}{\lambda^L} > 1 > \frac{P_{T^L}^H}{P_{T^L}^L}$.[42] In

[42] To explain, $f_2(t, T^L) = 0$ if and only if $\frac{1-\beta_0+\beta_0\left(1-\lambda^H\right)^{T^L}}{1-\beta_0+\beta_0\left(1-\lambda^L\right)^{T^L}} = \frac{\left(1-\lambda^H\right)^{t-1}\lambda^H}{\left(1-\lambda^L\right)^{t-1}\lambda^L}$. Given that the right hand side of the equation above is strictly decreasing since $\frac{1-\lambda^H}{1-\lambda^L} < 1$ and if evaluated at $t = 1$ is equal to $\frac{\lambda^H}{\lambda^L}$. Since $\frac{1-\beta_0+\beta_0\left(1-\lambda^H\right)^{T^L}}{1-\beta_0+\beta_0\left(1-\lambda^L\right)^{T^L}} < 1$ and $\frac{\lambda^H}{\lambda^L} > 1$ the uniqueness immediately follows. So $\hat{T}^L$ satisfies $\frac{P_{T^L}^H}{P_{T^L}^L} = \frac{\left(1-\lambda^H\right)^{\hat{T}^L-1}\lambda^H}{\left(1-\lambda^L\right)^{\hat{T}^L-1}\lambda^L}$.

addition, for $t < \hat{T}^L$ it follows that $f_2(t, T^L) < 0$ and, as a result, the high type is relatively more likely to succeed than the low type whereas for $t > \hat{T}^L$ the opposite is true. *Q.E.D.*

We will show that the solution to the principal's optimization problem depends on whether the $(IC^{L,H})$ constraint is binding or not; we explore each case separately in what follows.

Case A: The $(\boldsymbol{IC^{L,H}})$ constraint is not binding.

In this case the low type does not receive any rent and it immediately follows that $x^L = 0$ and $y_t^L = 0$ for $1 \le t \le T^L$. Thus, the rent of the high type can be derived from the RHS of $(IC^{H,L})$ as $\delta^{T^L} P_{T^L}^H \Delta c_{T^L+1} q^L(c_{T^L+1}^L)$. Using the binding $(IC^{H,L})$ to replace x^H in the objective function, the principal's optimization problem is to choose $\{y_t^H\}_{t=1}^{T^H}$ to

$$Max\ E_\theta\{\Omega^\theta(\varpi^\theta)\} - \upsilon \delta^{T^L} P_{T^L}^H \Delta c_{T^L+1} q^L(c_{T^L+1}^L) \text{ subject to:}$$

(LLS_t^H) $y_t^H \ge 0$ for $t \le T^H$,

and $(LLF_{T^H}^H)$ $\delta^{T^L} P_{T^L}^H \Delta c_{T^L+1} q^L(c_{T^L+1}^L) - \beta_0 \sum_{t=1}^{T^H} \delta^t (1-\lambda^H)^{t-1} \lambda^H y_t^H \ge 0$.

When the $(IC^{L,H})$ constraint is not binding, the claim below shows that there are no restrictions in choosing $\{y_t^H\}_{t=1}^{T^H}$ except those imposed by the $(IC^{H,L})$ constraint. In other words, the principal can choose any combinations of nonnegative payments to the high type $\left(x^H, \{y_t^H\}_{t=1}^{T^H}\right)$ such that $\beta_0 \sum_{t=1}^{T^H} \delta^t (1-\lambda^H)^{t-1} \lambda^H y_t^H + \delta^{T^H} P_{T^H}^H x^H = \delta^{T^L} P_{T^L}^H \Delta c_{T^L+1} q^L(c_{T^L+1}^L)$.

Labeling by $\{\alpha_t^H\}_{t=1}^{T^H}, \xi^H$ the Lagrange multipliers of the constraints associated with (LLS_t^H) for $t \le T^H$, and $(LLF_{T^H}^H)$ respectively, we have the following claim.

Claim A.1: If $(IC^{L,H})$ is not binding, we have $\xi^H = 0$ and $\alpha_t^H = 0$ for all $t \le T^H$.

Proof: We can rewrite the Kuhn-Tucker conditions as follows:

$\frac{\partial \mathcal{L}}{\partial y_t^H} = \alpha_t^H - \xi^H \beta_0 \delta^t (1-\lambda^H)^{t-1} \lambda^H = 0$ for $1 \le t \le T^H$;

$\frac{\partial \mathcal{L}}{\partial \alpha_t^H} = y_t^H \ge 0$; $\alpha_t^H \ge 0$; $\alpha_t^H y_t^H = 0$ for $1 \le t \le T^L$.

Suppose to the contrary that $\alpha^H > 0$. Then,

$$\delta^{T^L} P_{T^L}^H \Delta c_{T^L+1} q^L(c_{T^L+1}^L) - \beta_0 \sum_{t=1}^{T^H} \delta^t (1-\lambda^H)^{t-1} \lambda^H y_t^H = 0,$$

and there must exist $y_s^H > 0$ for some $1 \le s \le T^H$. Then, we have $\alpha_s^H = 0$, which leads to a contradiction since $\frac{\partial \mathcal{L}}{\partial y_t^H} = 0$ cannot be satisfied unless $\xi^H = 0$.

Suppose to the contrary that $\alpha_s^H > 0$ for some $1 \le s \le T^H$. Then, $\xi^H > 0$, which leads to a contradiction as we have just shown above. *Q.E.D.*

Case B: The ($\boldsymbol{IC^{L,H}}$) constraint is binding.

We will now show that when the $(IC^{L,H})$ becomes binding, there are restrictions on the payment structure to the high type. Denoting by $\psi = P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H}$, we can express x^H and x^L as functions of $\{y^H_t\}_{t=1}^{T^H}$, $\{y^L_t\}_{t=1}^{T^L}$, T^H, T^L, $q^H\left(c^H_{T^H+1}\right)$ and $q^L\left(c^L_{T^L+1}\right)$ from the binding $(IC^{H,L})$ and $(IC^{L,H})$.

$$x^H \delta^{T^H} \psi = \beta_0 \sum_{t=1}^{T^H} \delta^t \left[P^H_{T^L}(1-\lambda^L)^{t-1}\lambda^L - P^L_{T^L}(1-\lambda^H)^{t-1}\lambda^H\right] y^H_t$$

$$+\beta_0 \sum_{t=1}^{T^L} \delta^t \left[P^L_{T^L}(1-\lambda^H)^{t-1}\lambda^H - P^H_{T^L}(1-\lambda^L)^{t-1}\lambda^L\right] y^L_t$$

$$+P^H_{T^L}\left(\delta^{T^L} P^L_{T^L} \Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right) - \delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right)\right), \text{ and}$$

$$x^L \delta^{T^L} \psi = \beta_0 \sum_{t=1}^{T^H} \delta^t \left[P^H_{T^H}(1-\lambda^L)^{t-1}\lambda^L - P^L_{T^H}(1-\lambda^H)^{t-1}\lambda^H\right] y^H_t$$

$$+\beta_0 \sum_{t=1}^{T^L} \delta^t \left[P^L_{T^H}(1-\lambda^H)^{t-1}\lambda^H - P^H_{T^H}(1-\lambda^L)^{t-1}\lambda^L\right] y^L_t$$

$$+P^L_{T^H}\left(\delta^{T^L} P^H_{T^L} \Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right) - \delta^{T^H} P^H_{T^H} \Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right)\right).$$

Outline of the proof: First, we consider the case when $\psi \neq 0$. This is when the likelihood ratio of reaching the last period of the experimentation stage is different for both types i.e., when $\frac{P^H_{T^H}}{P^L_{T^H}} \neq \frac{P^H_{T^L}}{P^L_{T^L}}$ (Case B.1). We showed in Lemma 1 that there exists a time threshold $\hat{T}^L$ such that if type H claims to be type L, he is more likely to fail (resp. succeed) than type L if the experimentation stage is longer (resp. shorter) than $\hat{T}^L$. In Lemma 2 we prove that, if the principal rewards success, it is at most once. In Lemma 3, we establish that the high type is never rewarded for failure. In Lemma 4, we prove that the low type is rewarded for failure if and only if $T^L \leq \hat{T}^L$ and, in Lemma 5, that he is rewarded for the very last success if $T^L > \hat{T}^L$. In Lemma 6, we prove that $\hat{T}^L > T^L (<)$ for high (small) values of γ. Therefore, if the cost of experimentation is large ($\gamma > \gamma^*$), the principal must reward the low type after failure. If the cost of experimentation is small $(\gamma < \gamma^*)$, the principal must reward the low type after late success (last period). We also show that the high type is rewarded only for the very first success.

Finally, we analyze the case when $\frac{P^H_{T^H}}{P^L_{T^H}} = \frac{P^H_{T^L}}{P^L_{T^L}}$ (Case B.2). In this case, the likelihood ratio of reaching the last period of the experimentation stage is the same for both types and x^H and x^L cannot be used as screening variables. Therefore, the principal must reward both types for success and she chooses $T^L > \hat{T}^L$.

Case B.1: $\psi = P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \neq 0$.

Then x^H and x^L can be expressed as functions of $\{y^H_t\}_{t=1}^{T^H}$, $\{y^L_t\}_{t=1}^{T^L}$, T^H, T^L, $q^H\left(c^H_{T^H+1}\right)$ and $q^L\left(c^L_{T^L+1}\right)$ only from the binding $(IC^{H,L})$ and $(IC^{L,H})$. The principal's optimization problem is to choose $\{y^H_t\}_{t=1}^{T^H}$, T^L, $\{y^L_t\}_{t=1}^{T^L}$ to

$$Max\ E_\theta \left\{ \begin{array}{c} \Omega^\theta\left(\varpi^\theta\right) - \delta^{T^\theta} P^\theta_{T^\theta} x^\theta \left(\{y^H_t\}_{t=1}^{T^H}, \{y^L_t\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c^H_{T^H+1}\right), q^L\left(c^L_{T^L+1}\right)\right) \\ -\beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta y^\theta_t \end{array} \right\} \text{ subject to}$$

$\left(LLS^\theta_t\right)$ $y^\theta_t \geq 0$ for $t \leq T^\theta$,

$\left(LLF_{T^\theta}\right)$ $x^\theta \left(\{y^H_t\}_{t=1}^{T^H}, \{y^L_t\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c^H_{T^H+1}\right), q^L\left(c^L_{T^L+1}\right)\right) \geq 0$ for $\theta = H, L$.

Labeling $\{\alpha^H_t\}_{t=1}^{T^H}$, $\{\alpha^L_t\}_{t=1}^{T^L}$, ξ^H and ξ^L as the Lagrange multipliers of the constraints associated with (LLS^H_t), (LLS^L_t), $\left(LLF^H_{T^H}\right)$ and $\left(LLF^L_{T^L}\right)$ respectively, the Lagrangian is:

$$\mathcal{L} = E_\theta \left\{ \Omega^\theta\left(\varpi^\theta\right) - \beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta y^\theta_t - \right.$$
$$\left. \delta^{T^\theta} P^\theta_{T^\theta} x^\theta \left(\{y^H_t\}_{t=1}^{T^H}, \{y^L_t\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c^H_{T^H+1}\right), q^L\left(c^L_{T^L+1}\right)\right)\right\}$$
$$+ \sum_{t=1}^{T^H} \alpha^H_t y^H_t + \sum_{t=1}^{T^L} \alpha^L_t y^L_t + \xi^H x^H \left(\{y^H_t\}_{t=1}^{T^H}, \{y^L_t\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c^H_{T^H+1}\right), q^L\left(c^L_{T^L+1}\right)\right)$$
$$+ \xi^L x^L \left(\{y^H_t\}_{t=1}^{T^H}, \{y^L_t\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c^H_{T^H+1}\right), q^L\left(c^L_{T^L+1}\right)\right).$$

The Inada conditions give us interior solutions for $q^H_t(\overline{c})$, $q^H\left(c^H_{T^H+1}\right)$, $q^L_t(\overline{c})$ and $q^L\left(c^L_{T^L+1}\right)$. The Kuhn-Tucker conditions with respect to y^H_t and y^L_t are:

$$\frac{\partial \mathcal{L}}{\partial y^H_t} = -\upsilon \left\{ \beta_0 \delta^t (1-\lambda^H)^{t-1} \lambda^H + \delta^{T^H} P^H_{T^H} \frac{\beta_0 \delta^t \left[P^H_{T^L} (1-\lambda^L)^{t-1} \lambda^L - P^L_{T^L} (1-\lambda^H)^{t-1} \lambda^H \right]}{\delta^{T^H} \left(P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \right)} \right\}$$
$$-(1-\upsilon) \delta^{T^L} P^L_{T^L} \frac{\beta_0 \delta^t \left[P^H_{T^H} (1-\lambda^L)^{t-1} \lambda^L - P^L_{T^H} (1-\lambda^H)^{t-1} \lambda^H \right]}{\delta^{T^L} \left(P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \right)} + \alpha^H_t$$
$$+ \xi^H \frac{\beta_0 \delta^t \left[P^H_{T^L} (1-\lambda^L)^{t-1} \lambda^L - P^L_{T^L} (1-\lambda^H)^{t-1} \lambda^H \right]}{\delta^{T^H} \left(P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \right)} + \xi^L \frac{\beta_0 \delta^t \left[P^H_{T^H} (1-\lambda^L)^{t-1} \lambda^L - P^L_{T^H} (1-\lambda^H)^{t-1} \lambda^H \right]}{\delta^{T^L} \left(P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \right)};$$

$$\frac{\partial \mathcal{L}}{\partial y^L_t} = -(1-\upsilon) \left\{ \beta_0 \delta^t (1-\lambda^L)^{t-1} \lambda^L + \delta^{T^L} P^L_{T^L} \frac{\beta_0 \delta^t \left[P^L_{T^H} (1-\lambda^H)^{t-1} \lambda^H - P^H_{T^H} (1-\lambda^L)^{t-1} \lambda^L \right]}{\delta^{T^L} \left(P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \right)} \right\}$$
$$-\upsilon \delta^{T^H} P^H_{T^H} \frac{\beta_0 \delta^t \left[P^L_{T^L} (1-\lambda^H)^{t-1} \lambda^H - P^H_{T^L} (1-\lambda^L)^{t-1} \lambda^L \right]}{\delta^{T^H} \left(P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \right)} + \alpha^L_t$$
$$+ \xi^H \frac{\beta_0 \delta^t \left[P^L_{T^L} (1-\lambda^H)^{t-1} \lambda^H - P^H_{T^L} (1-\lambda^L)^{t-1} \lambda^L \right]}{\delta^{T^H} \left(P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \right)} + \xi^L \frac{\beta_0 \delta^t \left[P^L_{T^H} (1-\lambda^H)^{t-1} \lambda^H - P^H_{T^H} (1-\lambda^L)^{t-1} \lambda^L \right]}{\delta^{T^L} \left(P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} \right)}.$$

We can rewrite the Kuhn-Tucker conditions above as follows:

$$(\boldsymbol{A1}) \quad \frac{\partial \mathcal{L}}{\partial y_t^H} = \frac{\beta_0 \delta^t}{\psi}\left[P_{T^H}^H f_1(t)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L - \frac{\xi^L}{\delta^{T^L}}\right] + \frac{\xi^H}{\delta^{T^H}} P_{T^L}^L f_2(t) + \frac{\alpha_t^H \psi}{\beta_0 \delta^t}\right] = 0,$$

$$(\boldsymbol{A2}) \quad \frac{\partial \mathcal{L}}{\partial y_t^L} = \frac{\beta_0 \delta^t}{\psi}\left[P_{T^L}^L f_2(t)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(t) + \frac{\alpha_t^L \psi}{\beta_0 \delta^t}\right] = 0,$$

where

$$f_1(t, T^H) = \frac{P_{T^H}^L}{P_{T^H}^H}(1-\lambda^H)^{t-1}\lambda^H - (1-\lambda^L)^{t-1}\lambda^L, \text{ and}$$

$$f_2(t, T^L) = \frac{P_{T^L}^H}{P_{T^L}^L}(1-\lambda^L)^{t-1}\lambda^L - (1-\lambda^H)^{t-1}\lambda^H.$$

Next, we show that the principal will reward success in at most one period.

Lemma 2. There exists *at most* one time period $1 \leq j \leq T^L$ such that $y_j^L > 0$ and *at most* one time period $1 \leq s \leq T^H$ such that $y_s^H > 0$.

Proof: Assume to the contrary that there are two distinct periods $1 \leq k, m \leq T^L$ such that $k \neq m$ and $y_k^L, y_m^L > 0$. Then from the Kuhn-Tucker conditions $(A1)$ and $(A2)$ it follows that

$$P_{T^L}^L f_2(k, T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(k, T^H) = 0,$$

and, in addition, $P_{T^L}^L f_2(m, T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(m, T^H) = 0.$

Thus, $\frac{f_2(m,T^L)}{f_1(m,T^H)} = \frac{f_2(k,T^L)}{f_1(k,T^H)}$, which can be rewritten as follows:

$$\left(P_{T^L}^H(1-\lambda^L)^{m-1}\lambda^L - P_{T^L}^L(1-\lambda^H)^{m-1}\lambda^H\right)\left(P_{T^H}^L(1-\lambda^H)^{k-1}\lambda^H - P_{T^H}^H(1-\lambda^L)^{k-1}\lambda^L\right)$$
$$= \left(P_{T^L}^H(1-\lambda^L)^{k-1}\lambda^L - P_{T^L}^L(1-\lambda^H)^{k-1}\lambda^H\right)\left(P_{T^H}^L(1-\lambda^H)^{m-1}\lambda^H - P_{T^H}^H(1-\lambda^L)^{m-1}\lambda^L\right),$$
$$\psi[(1-\lambda^H)^{k-1}(1-\lambda^L)^{m-1} - (1-\lambda^L)^{k-1}(1-\lambda^H)^{m-1}] = 0,$$
$$(1-\lambda^L)^{m-k}(1-\lambda^H)^{k-m} = 1,$$

$\left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{m-k} = 1$, which implies that $m = k$ and we have a contradiction.

Following similar steps, we can show that there exists *at most* one time period $1 \leq s \leq T^H$ such that $y_s^H > 0$. *Q.E.D.*

For later use, we prove the following claim:

Claim B.1.1. $\frac{\xi^L}{\delta^{T^L}} \neq \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$ and $\frac{\xi^H}{\delta^{T^H}} \neq \upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L$.

Proof: By contradiction. Suppose $\frac{\xi^L}{\delta^{T^L}} = \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$. Then combining conditions $(A1)$ and $(A2)$ we have

$$P_{T^L}^L f_2(t, T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(t, T^H)$$
$$= \left(P_{T^L}^H(1-\lambda^L)^{t-1}\lambda^L - P_{T^L}^L(1-\lambda^H)^{t-1}\lambda^H\right)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L\right]$$
$$+\left(P_{T^H}^L(1-\lambda^H)^{t-1}\lambda^H - P_{T^H}^H(1-\lambda^L)^{t-1}\lambda^L\right)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\right]$$
$$= -\psi\left((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H\right),$$

which implies that $-\psi\big((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L+\upsilon(1-\lambda^H)^{t-1}\lambda^H\big)+\frac{\alpha_t^L\psi}{\beta_0\delta^t}=0$ for $1\le t\le T^L$.

Thus, $\frac{\alpha_t^L}{\beta_0\delta^t}=(1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L+\upsilon(1-\lambda^H)^{t-1}\lambda^H>0$ for $1\le t\le T^L$, which leads to a contradiction since then $x^L=y_t^L=0$ for $1\le t\le T^L$ which implies that the low type does not receive any rent.

Next, assume $\frac{\xi^H}{\delta^{T^H}}=\upsilon P_{T^H}^H+(1-\upsilon)P_{T^H}^L$. Then combining conditions $(A1)$ and $(A2)$ gives

$$\begin{aligned}&P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H+(1-\upsilon)P_{T^L}^L\right]+\frac{\xi^H}{\delta^{T^H}}P_{T^L}^L f_2(t,T^L)\\&=\left(P_{T^H}^L(1-\lambda^H)^{t-1}\lambda^H-P_{T^H}^H(1-\lambda^L)^{t-1}\lambda^L\right)\left[\upsilon P_{T^L}^H+(1-\upsilon)P_{T^L}^L\right]\\&+\left(P_{T^L}^H(1-\lambda^L)^{t-1}\lambda^L-P_{T^L}^L(1-\lambda^H)^{t-1}\lambda^H\right)\left[\upsilon P_{T^H}^H+(1-\upsilon)P_{T^H}^L\right]\\&=-\psi\big((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L+\upsilon(1-\lambda^H)^{t-1}\lambda^H\big),\end{aligned}$$

which implies that $-\psi\big((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L+\upsilon(1-\lambda^H)^{t-1}\lambda^H\big)+\frac{\alpha_t^H\psi}{\beta_0\delta^t}=0$ for $1\le t\le T^H$.

Then $\frac{\alpha_t^H}{\beta_0\delta^t}=(1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L+\upsilon(1-\lambda^H)^{t-1}\lambda^H>0$ for $1\le t\le T^H$, which leads to a contradiction since then $x^H=y_t^H=0$ for $1\le t\le T^H$ (which implies that the high type does not receive any rent and we are back in Case A.) *Q.E.D.*

Now we prove that the high type may be only rewarded for success. Although the proof is long, the result should appear intuitive: Rewarding high type for failure will only exacerbates the problem as the low type is always relatively more optimistic in case he lies, and experimentation fails.

Lemma 3: The high type is not rewarded for failure, i.e., $x^H=0$.

Proof: By contradiction. We consider separately Case (a) $\xi^H=\xi^L=0$, and Case (b) $\xi^H=0$ and $\xi^L>0$.

Case (a): Suppose that $\xi^H=\xi^L=0$, i.e., the $\left(LLF_{T^H}^H\right)$ and $\left(LLF_{T^L}^L\right)$ constraints are not binding.

We can rewrite the Kuhn-Tucker conditions $(A1)$ and $(A2)$ as follows:

$$\frac{\partial\mathcal{L}}{\partial y_t^H}=\frac{\beta_0\delta^t}{\psi}\left[P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H+(1-\upsilon)P_{T^L}^L\right]+\frac{\alpha_t^H\psi}{\beta_0\delta^t}\right]=0 \text{ for } 1\le t\le T^H;$$

$$\frac{\partial\mathcal{L}}{\partial y_t^L}=\frac{\beta_0\delta^t}{\psi}\left[P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H+(1-\upsilon)P_{T^H}^L\right]+\frac{\alpha_t^L\psi}{\beta_0\delta^t}\right]=0 \text{ for } 1\le t\le T^L.$$

Since $f_1(t,T^H)$ is strictly positive for all $t<\hat{T}^H$ from $P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H+(1-\upsilon)P_{T^L}^L\right]=-\frac{\alpha_t^H\psi}{\beta_0\delta^t}$ it must be that $\alpha_t^H>0$ for all $t<\hat{T}^H$ and $\psi<0$. In addition, since $f_2(t,T^L)$ is strictly negative for $t<\hat{T}^L$ from $P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H+(1-\upsilon)P_{T^H}^L\right]=-\frac{\alpha_t^L\psi}{\beta_0\delta^t}$ it must be that that $\alpha_t^L>0$ for $t<\hat{T}^L$ and $\psi>0$, which leads to a contradiction[43].

[43] If there was a solution with $\xi^H=\xi^L=0$ then with necessity it would be possible only if T^H and T^L are such that it holds simultaneously $P_{T^H}^H P_{T^L}^L-P_{T^L}^H P_{T^H}^L>0$ and $P_{T^H}^H P_{T^L}^L-P_{T^L}^H P_{T^H}^L<0$, since the two conditions are mutually exclusive the conclusion immediately follows. Recall that we assumed so far that $\psi\neq 0$; we study $\psi=0$ in details later in Case B.2.

Case (b): Suppose that $\xi^H = 0$ and $\xi^L > 0$, i.e., the $\left(LLF^H_{T^H}\right)$ constraint is not binding but $\left(LLF^L_{T^L}\right)$ is binding.

We can rewrite the Kuhn-Tucker conditions $(A1)$ and $(A2)$ as follows:

$\frac{\partial \mathcal{L}}{\partial y^H_t} = \frac{\beta_0 \delta^t}{\psi}\left[P^H_{T^H} f_1(t, T^H)\left[\upsilon P^H_{T^L} + (1-\upsilon)P^L_{T^L} - \frac{\xi^L}{\delta^{T^L}}\right] + \frac{\alpha^H_t \psi}{\beta_0 \delta^t}\right] = 0$ for $1 \le t \le T^H$;

$\frac{\partial \mathcal{L}}{\partial y^L_t} = \frac{\beta_0 \delta^t}{\psi}\left[P^L_{T^L} f_2(t, T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}\right] + \frac{\xi^L}{\delta^{T^L}} P^H_{T^H} f_1(t, T^H) + \frac{\alpha^L_t \psi}{\beta_0 \delta^t}\right] = 0$ for $1 \le t \le T^L$.

If $\alpha^H_s = 0$ for some $1 \le s \le T^H$ then $P^H_{T^H} f_1(s, T^H)\left[\upsilon P^H_{T^L} + (1-\upsilon)P^L_{T^L} - \frac{\xi^L}{\delta^{T^L}}\right] = 0$,

which implies that $\frac{\xi^L}{\delta^{T^L}} = \upsilon P^H_{T^L} + (1-\upsilon)P^L_{T^L}$[44]. Since we rule out this possibility it immediately follows that all $\alpha^H_t > 0$ for all $1 \le t \le T^H$ which implies that $y^H_t = 0$ for $1 \le t \le T^H$.

Finally, from $P^H_{T^H} f_1(t, T^H)\left[\upsilon P^H_{T^L} + (1-\upsilon)P^L_{T^L} - \frac{\xi^L}{\delta^{T^L}}\right] = -\frac{\alpha^H_t \psi}{\beta_0 \delta^t}$ we conclude that $T^H \le \hat{T}^H$ and there can be one of two sub-cases:[45] (b.1) $\psi > 0$ and $\frac{\xi^L}{\delta^{T^L}} > \upsilon P^H_{T^L} + (1-\upsilon)P^L_{T^L}$, or (b.2) $\psi < 0$ and $\frac{\xi^L}{\delta^{T^L}} < \upsilon P^H_{T^L} + (1-\upsilon)P^L_{T^L}$. We consider each sub-case next.

Case (b.1): $T^H \le \hat{T}^H$, $\psi > 0$, $\frac{\xi^L}{\delta^{T^L}} > \upsilon P^H_{T^L} + (1-\upsilon)P^L_{T^L}$, $\xi^H = 0$, $\alpha^H_t > 0$ for $1 \le t \le T^H$.

We know from Lemma 3 that there exists only one time period $1 \le j \le T^L$ such that $y^L_j > 0$ ($\alpha^L_j = 0$). This implies that

$$P^L_{T^L} f_2(j, T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}\right] + \frac{\xi^L}{\delta^{T^L}} P^H_{T^H} f_1(j, T^H) = 0$$

and $P^L_{T^L} f_2(t, T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}\right] + \frac{\xi^L}{\delta^{T^L}} P^H_{T^H} f_1(t, T^H) = -\frac{\alpha^L_t \psi}{\beta_0 \delta^t} < 0$ for $1 \le t \ne j \le T^L$.

Alternatively, $f_2(t, T^L) < \frac{f_1(t,T^H)}{f_1(j,T^H)} f_2(j, T^L)$ for $1 \le t \ne j \le T^L$.

If $f_1(j, T^H) > 0$ ($j < \hat{T}^H$) then

$$\left(P^H_{T^L}(1-\lambda^L)^{t-1}\lambda^L - P^L_{T^L}(1-\lambda^H)^{t-1}\lambda^H\right)\left(P^L_{T^H}(1-\lambda^H)^{j-1}\lambda^H - P^H_{T^H}(1-\lambda^L)^{j-1}\lambda^L\right)$$
$$< \left(P^H_{T^L}(1-\lambda^L)^{j-1}\lambda^L - P^L_{T^L}(1-\lambda^H)^{j-1}\lambda^H\right)\left(P^L_{T^H}(1-\lambda^H)^{t-1}\lambda^H - P^H_{T^H}(1-\lambda^L)^{t-1}\lambda^L\right).$$
$$\psi\left[(1-\lambda^H)^{t-1}(1-\lambda^L)^{j-1} - (1-\lambda^L)^{t-1}(1-\lambda^H)^{j-1}\right] < 0 \text{ for } 1 \le t \ne j \le T^L.$$

$\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-j}\right] < 0$, which implies that $t > j$ for all $1 \le t \ne j \le T^L$ or, equivalently, $j = 1$.

If $f_1(j, T^H) < 0$ ($j > \hat{T}^H$) then the opposite must be true and $t < j$ for all $1 \le t \ne j \le T^L$ or, equivalently, $j = T^L$.

For $j > \hat{T}^H$ we have $f_1(j, T^H) < 0$ and it follows that $P^L_{T^L} f_2(t, T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}\right] + \frac{\xi^L}{\delta^{T^L}} P^H_{T^H} f_1(t, T^H) < -\psi\left((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H\right) < 0$, which implies that $y^L_j > 0$ is only possible for $j < \hat{T}^H$. Thus, this case is only possible if $j = 1$.

---

[44] If $s = \hat{T}^H$, then both $x^H > 0$ and $y^H_{\hat{T}^H} > 0$ can be optimal.

[45] If $T^H > \hat{T}^H$ then there would be a contradiction since $f_1(t, T^H)$ must be of the same sign for all $t \le T^H$.

Case (b.2): $T^H \le \hat{T}^H$, $\psi < 0$, $\frac{\xi^L}{\delta^{T^L}} < \upsilon P^H_{T^L} + (1-\upsilon)P^L_{T^L}$, $\xi^H = 0$, $\alpha^H_t > 0$ for $1 \le t \le T^H$.

As in the previous case, from Lemma 3 it follows that there exists only one time period $1 \le s \le T^L$ such that $y^L_s > 0$ ($\alpha^L_s = 0$). This implies that $P^L_{T^L} f_2(s, T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}\right] + \frac{\xi^L}{\delta^{T^L}} P^H_{T^H} f_1(s, T^H) = 0$ and $P^L_{T^L} f_2(t, T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}\right] + \frac{\xi^L}{\delta^{T^L}} P^H_{T^H} f_1(t, T^H) = -\frac{\alpha^L_t \psi}{\beta_0 \delta^t} > 0$ for $1 \le t \ne s \le T^L$. Alternatively, $f_2(t, T^L) > \frac{f_1(t,T^H)}{f_1(s,T^H)} f_2(s, T^L)$.

If $f_1(s, T^H) > 0$ ($s < \hat{T}^H$) then $f_2(t, T^L) f_1(s, T^H) > f_1(t, T^H) f_2(s, T^L)$

$$\left(P^H_{T^L}(1-\lambda^L)^{t-1}\lambda^L - P^L_{T^L}(1-\lambda^H)^{t-1}\lambda^H\right)\left(P^L_{T^H}(1-\lambda^H)^{s-1}\lambda^H - P^H_{T^H}(1-\lambda^L)^{s-1}\lambda^L\right)$$
$$> \left(P^H_{T^L}(1-\lambda^L)^{s-1}\lambda^L - P^L_{T^L}(1-\lambda^H)^{s-1}\lambda^H\right)\left(P^L_{T^H}(1-\lambda^H)^{t-1}\lambda^H - P^H_{T^H}(1-\lambda^L)^{t-1}\lambda^L\right).$$

$\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{s-t}\right] < 0$, which implies that $t > s$ for all$1 \le t \ne s \le T^L$ or, equivalently, $s = 1$.

If $f_1(s, T^H) < 0$ ($s > \hat{T}^H$) then the opposite must be true and $t < s$ for all $1 \le t \ne s \le T^L$ or, equivalently, $s = T^L$.

For $t > \hat{T}^H$ it follows that $P^L_{T^L} f_2(t, T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}\right] + \frac{\xi^L}{\delta^{T^L}} P^H_{T^H} f_1(t, T^H)$ $> -\psi\left((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H\right) > 0$, which implies that $y^L_s > 0$ is only possible for $s < \hat{T}^H$, which is only possible if $s = 1$.

For both cases we just considered, we have

$$x^H = \frac{\beta_0 \delta P^L_{T^L}\left(-f_2(1, T^L)\right) y^L_1}{\delta^{T^H}\psi} + \frac{P^H_{T^L}\left(\delta^{T^L} P^L_{T^L}\Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right) - \delta^{T^H} P^L_{T^H}\Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right)\right)}{\delta^{T^H}\psi} \ge 0;$$

$$x^L = \frac{\beta_0 \delta P^H_{T^H} f_1(1, T^H) y^L_1}{\delta^{T^L}\psi} + \frac{P^L_{T^H}\left(\delta^{T^L} P^H_{T^L}\Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right) - \delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right)\right)}{\delta^{T^L}\psi} = 0.$$

Note that Case B.2 is possible only if $-\delta^{T^H} P^L_{T^H}\Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right) + \delta^{T^L} P^L_{T^L}\Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right) > 0$.[46] This fact together with $x^H \ge 0$ implies that $\psi > 0$. Since $f_1(1, T^H) > 0$, $x^L = 0$ is possible only if $-\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right) + \delta^{T^L} P^H_{T^L}\Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right) < 0$. However, $\delta^{T^H} P^L_{T^H}\Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right) > \delta^{T^L} P^L_{T^L}\Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right)$ implies that $\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right) > \delta^{T^L} \frac{P^H_{T^H}}{P^L_{T^H}} P^L_{T^L}\Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right)$. Note that $P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} > 0$ implies $\frac{P^H_{T^H}}{P^L_{T^H}} P^L_{T^L} > P^H_{T^L}$, and then $\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right) > \delta^{T^L} P^H_{T^L}\Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right)$, which implies $x^L > 0$ and we have a contradiction. Thus, $\xi^H > 0$ and the high type gets rent only after success ($x^H = 0$). *Q.E.D.*

We now prove that the low type is rewarded for failure only if the duration of the experimentation stage for the low type, T^L, is relatively short: $T^L \le \hat{T}^L$.

[46] Otherwise the $(IC^{H,L})$ is not binding.

Lemma 4. $\xi^L = 0 \Rightarrow T^L \leq \hat{T}^L$, $\alpha_t^L > 0$ for $t \leq T^L$ (it is optimal to set $x^L > 0$, $y_t^L = 0$ for $t \leq T^L$) and $\alpha_t^H > 0$ for all $t > 1$ and $\alpha_1^H = 0$ (it is optimal to set $x^H = 0$, $y_t^H = 0$ for all $t > 1$ and $y_1^H > 0$).

Proof: Suppose that $\xi^L = 0$, i.e., the $\left(LLF_{T^L}^L\right)$ constraint is not binding. We can rewrite the Kuhn-Tucker conditions $(A1)$ and $(A2)$ as follows:

$$\frac{\partial\mathcal{L}}{\partial y_t^H} = \frac{\beta_0\delta^t}{\psi}\left[P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\right] + \frac{\xi^H}{\delta^{T^H}}P_{T^L}^L f_2(t,T^L) + \frac{\alpha_t^H\psi}{\beta_0\delta^t}\right] = 0 \text{ for } 1 \leq t \leq T^H;$$

$$\frac{\partial\mathcal{L}}{\partial y_t^L} = \frac{\beta_0\delta^t}{\psi}\left[P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] + \frac{\alpha_t^L\psi}{\beta_0\delta^t}\right] = 0 \text{ for } 1 \leq t \leq T^L.$$

If $\alpha_s^L = 0$ for some $1 \leq s \leq T^L$ then $P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] = 0$, which implies that $\frac{\xi^H}{\delta^{T^H}} = \upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L$[47]. Since we already rule out this possibility it immediately follows that $\alpha_t^L > 0$ for all $1 \leq t \leq T^L$ which implies that $y_t^L = 0$ for $1 \leq t \leq T^L$.

Finally, $P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] = -\frac{\alpha_t^L\psi}{\beta_0\delta^t}$ for $1 \leq t \leq T^L$ and we conclude that $T^L \leq \hat{T}^L$ and there can be one of two sub-cases:[48] (a) $\psi > 0$ and $\frac{\xi^H}{\delta^{T^H}} < \upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L$, or (b) $\psi < 0$ and $\frac{\xi^H}{\delta^{T^H}} > \upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L$. We consider each sub-case next.

Case (a): $T^L \leq \hat{T}^L$, $\psi > 0$, $\frac{\xi^H}{\delta^{T^H}} < \upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L$, $\xi^L = 0$, $\alpha_t^L > 0$ for $1 \leq t \leq T^L$.

From Lemma 2, there exists only one time period $1 \leq j \leq T^H$ such that $y_j^H > 0$ ($\alpha_j^H = 0$). This implies that

$$P_{T^H}^H f_1(j,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\right] + \frac{\xi^H}{\delta^{T^H}}P_{T^L}^L f_2(j,T^L) = 0 \text{ and}$$

$$P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\right] + \frac{\xi^H}{\delta^{T^H}}P_{T^L}^L f_2(t,T^L) = -\frac{\alpha_t^H\psi}{\beta_0\delta^t} < 0 \text{ for } 1 \leq t \neq j \leq T^H.$$

Alternatively, $f_1(t,T^H) < \frac{f_1(j,T^H)}{f_2(j,T^L)}f_2(t,T^L)$ for $1 \leq t \neq j \leq T^H$.

If $f_2(j,T^L) > 0$ $(j > \hat{T}^L)$ then $f_1(t,T^H)f_2(j,T^L) < f_1(j,T^H)f_2(t,T^L)$

$$\left(P_{T^L}^H(1-\lambda^L)^{j-1}\lambda^L - P_{T^L}^L(1-\lambda^H)^{j-1}\lambda^H\right)\left(P_{T^H}^L(1-\lambda^H)^{t-1}\lambda^H - P_{T^H}^H(1-\lambda^L)^{t-1}\lambda^L\right)$$
$$< \left(P_{T^L}^H(1-\lambda^L)^{t-1}\lambda^L - P_{T^L}^L(1-\lambda^H)^{t-1}\lambda^H\right)\left(P_{T^H}^L(1-\lambda^H)^{j-1}\lambda^H - P_{T^H}^H(1-\lambda^L)^{j-1}\lambda^L\right),$$
$$\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t}\right] < 0,$$

which implies that $t < j$ for all $1 \leq t \neq j \leq T^H$ or, equivalently, $j = T^H$.

If $f_2(j,T^L) < 0$ $(j < \hat{T}^L)$ then the opposite must be true and $t > j$ for all $1 \leq t \neq j \leq T^H$ or, equivalently, $j = 1$.

For $t > \hat{T}^L$ it follows that $P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\right] + \frac{\xi^H}{\delta^{T^H}}P_{T^L}^L f_2(t,T^L)$

---

[47] If $t = \hat{T}^L$, then both $x^L > 0$ and $y_{\hat{T}^L}^L > 0$ can be optimal.

[48] If $T^L > \hat{T}^L$, then there would be a contradiction since $f_2(t,T^L)$ must be of the same sign for all $t \leq T^L$.

$< -\psi\big((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H\big) < 0$, which implies that $y_j^H > 0$ is only possible for $j < \hat{T}^L$ and we have $j = 1$.

Case (b): $T^L \leq \hat{T}^L$, $\psi < 0$, $\frac{\xi^H}{\delta^{T^H}} > \upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L$, $\xi^L = 0$, $\alpha_t^L > 0$ for $1 \leq t \leq T^L$.

From Lemma 2, there exists only one time period $1 \leq j \leq T^H$ such that $y_j^H > 0$ ($\alpha_j^H = 0$). This implies that

$$P_{T^H}^H f_1(j,T^H)\big[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\big] + \frac{\xi^H}{\delta^{T^H}} P_{T^L}^L f_2(j,T^L) = 0 \text{ and}$$

$$P_{T^H}^H f_1(t,T^H)\big[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\big] + \frac{\xi^H}{\delta^{T^H}} P_{T^L}^L f_2(t,T^L) = -\frac{\alpha_t^H \psi}{\beta_0 \delta^t} > 0 \text{ for } 1 \leq t \neq j \leq T^H.$$

Alternatively, $f_1(t,T^H) > \frac{f_1(j,T^H)}{f_2(j,T^L)} f_2(t,T^L)$ for $1 \leq t \neq j \leq T^H$.

If $f_2(j,T^L) > 0$ ($j > \hat{T}^L$) then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-j}\right] < 0$, which implies that $t < j$ for all $1 \leq t \neq j \leq T^H$ or, equivalently, $j = T^H$.

If $f_2(j,T^L) < 0$ ($j < \hat{T}^L$) then the opposite must be true and $t > j$ for all $1 \leq t \neq j \leq T^H$ or, equivalently, $j = 1$.

For $t > \hat{T}^L$ ($f_2(t,T^L) > 0$) it follows that

$$P_{T^H}^H f_1(t,T^H)\big[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\big] + \frac{\xi^H}{\delta^{T^H}} P_{T^L}^L f_2(t,T^L)$$
$$> -\psi\big((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H\big) > 0,$$

which implies that $y_j^H > 0$ is only possible for $j < \hat{T}^L$ and we have $j = 1$.

If $T^L < \hat{T}^L$, from the binding incentive compatibility constraints, we derive the optimal payments:

$$y_1^H = \frac{P_{T^L}^H\Big(\delta^{T^H} P_{T^H}^L \Delta c_{T^H+1} q^H\big(c_{T^H+1}^H\big) - \delta^{T^L} P_{T^L}^L \Delta c_{T^L+1} q^L\big(c_{T^L+1}^L\big)\Big)}{\beta_0 \delta P_{T^L}^L f_2(1,T^L)} \geq 0;$$

$$x^L = \frac{\delta^{T^H} \lambda^H P_{T^H}^L \Delta c_{T^H+1} q^H\big(c_{T^H+1}^H\big) - \delta^{T^L} \lambda^L P_{T^L}^H \Delta c_{T^L+1} q^L\big(c_{T^L+1}^L\big)}{\delta^{T^L} P_{T^L}^L f_2(1,T^L)} > 0. \qquad \textit{Q.E.D.}$$

We now prove that the low type is rewarded for success only if the duration of the experimentation stage for the low type, T^L, is relatively long: $T^L > \hat{T}^L$.

Lemma 5: $\xi^L > 0 \Rightarrow T^L > \hat{T}^L$, $\alpha_t^L > 0$ for $t < T^L$, $\alpha_{T^L}^L = 0$ and $\alpha_t^H > 0$ for $t > 1$, $\alpha_1^H = 0$ (it is optimal to set $x^L = 0$, $y_t^L = 0$ for $t < T^L$, $y_{T^L}^L > 0$ and $x^H = 0$, $y_t^H = 0$ for $t > 1$, $y_1^H > 0$)

Proof: Suppose that $\xi^L > 0$, i.e., the $(LLF_{T^L}^L)$ constraint is binding. We can rewrite the Kuhn-Tucker conditions $(A1)$ and $(A2)$ as follows:

$$\left[P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L - \frac{\xi^L}{\delta^{T^L}}\right] + \frac{\xi^H}{\delta^{T^H}} P_{T^L}^L f_2(t,T^L) + \frac{\alpha_t^H \psi}{\beta_0 \delta^t}\right] = 0 \text{ for } 1 \leq t \leq T^H;$$

$$\left[P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(t,T^H) + \frac{\alpha_t^L \psi}{\beta_0 \delta^t}\right] = 0 \text{ for } 1 \leq t \leq T^L.$$

As part of the proof of Lemma 5, we prove the following claim:

Claim B.1.2: If both types are rewarded for success, it must be at *extreme* time periods, i.e. only at *the last* or *the first* period of the experimentation stage.

Proof: Since (See Lemma 2) there exists only one time period $1 \le j \le T^L$ such that $y_j^L > 0$ ($\alpha_j^L = 0$) it follows that

$$P_{T^L}^L f_2(j,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(j,T^H) = 0 \text{ and}$$

$$P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L - \frac{\xi^H}{\delta^{T^H}}\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(t,T^H) = -\frac{\alpha_t^L \psi}{\beta_0 \delta^t} \text{ for } 1 \le t \ne j \le T^L.$$

Alternatively, $\frac{\xi^L}{\delta^{T^L}}\left[f_1(t,T^H) - \frac{f_2(t,T^L) f_1(j,T^H)}{f_2(j,T^L)}\right] = -\frac{\alpha_t^L \psi}{\beta_0 \delta^t P_{T^H}^H}$ for $1 \le t \ne j \le T^L$.

Suppose $\psi > 0$. Then $f_1(t,T^H) - \frac{f_2(t,T^L) f_1(j,T^H)}{f_2(j,T^L)} < 0$ for $1 \le t \ne j \le T^L$.

If $f_2(j,T^L) > 0$ ($j > \hat{T}^L$) then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t}\right] < 0$ which implies $1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t} < 0$ or, equivalently, $j > t$ for $1 \le t \ne j \le T^L$ which implies that $j = T^L > \hat{T}^L$.

If $f_2(j,T^L) < 0$ ($j < \hat{T}^L$) then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t}\right] > 0$ which implies $1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t} > 0$ or, equivalently, $j < t$ for $1 \le t \ne j \le T^L$ which implies that $j = 1$.

Suppose $\psi < 0$. Then $f_1(t,T^H) - \frac{f_2(t,T^L) f_1(j,T^H)}{f_2(j,T^L)} > 0$ for $1 \le t \ne j \le T^L$.

If $f_2(j,T^L) > 0$ ($j > \hat{T}^L$) then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t}\right] > 0$ which implies $1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t} < 0$ or, equivalently, $j > t$ for $1 \le t \ne j \le T^L$ which implies that $j = T^L > \hat{T}^L$.

If $f_2(j,T^L) < 0$ ($j < \hat{T}^L$) then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t}\right] < 0$ which implies $1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{j-t} > 0$ or, equivalently, $j < t$ for $1 \le t \ne j \le T^L$ which implies that $j = 1$.

Since (from Lemma 2) there exists only one time period $1 \le s \le T^H$ such that $y_s^H > 0$ ($\alpha_s^H = 0$) it follows that

$$P_{T^H}^H f_1(s,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L - \frac{\xi^L}{\delta^{T^L}}\right] + \frac{\xi^H}{\delta^{T^H}} P_{T^L}^L f_2(s,T^L) = 0,$$

$$P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L - \frac{\xi^L}{\delta^{T^L}}\right] + \frac{\xi^H}{\delta^{T^H}} P_{T^L}^L f_2(t,T^L) = -\frac{\alpha_t^H \psi}{\beta_0 \delta^t} < 0 \text{ for } 1 \le t \ne s \le T^H.$$

Alternatively, $\frac{\xi^H}{\delta^{T^H}}\left[f_2(t,T^L) - \frac{f_2(s,T^L) f_1(t,T^H)}{f_1(s,T^H)}\right] = -\frac{\alpha_t^H \psi}{\beta_0 \delta^t P_{T^L}^L}$ for $1 \le t \ne s \le T^H$.

Suppose $\psi > 0$. Then $f_2(t,T^L) - \frac{f_2(s,T^L) f_1(t,T^H)}{f_1(s,T^H)} < 0$ for $1 \le t \ne s \le T^H$.

If $f_1(s,T^H) > 0$ ($s < \hat{T}^H$) then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-s}\right] < 0$ which implies $1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-s} < 0$ or, equivalently, $t > s$ for $1 \le t \ne s \le T^H$ which implies that $s = 1$.

If $f_1(s,T^H) < 0$ ($s > \hat{T}^H$) then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-s}\right] > 0$ which implies $1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-s} > 0$ or, equivalently, $t < s$ for $1 \le t \ne s \le T^H$ which implies that $s = T^H > \hat{T}^H$.

Suppose $\psi < 0$. Then $f_2(t, T^L) - \frac{f_2(s,T^L)f_1(t,T^H)}{f_1(s,T^H)} > 0$ for $1 \leq t \neq s \leq T^H$.

If $f_1(s, T^H) > 0$ $(s < \hat{T}^H)$ then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-s}\right] > 0$ which implies $1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-s} < 0$ or, equivalently, $t > s$ for $1 \leq t \neq s \leq T^H$ which implies that $s = 1$.

If $f_1(s, T^H) < 0$ $(s > \hat{T}^H)$ then $\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-s}\right] < 0$ which implies $1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{t-s} > 0$ or, equivalently, $t < s$ for $1 \leq t \neq s \leq T^H$ which implies that $s = T^H > \hat{T}^H$.

This completes the proof of Claim B.1.2. *Q.E.D.*

The Lagrange multipliers are uniquely determined from $(A1)$ and $(A2)$ as follows:

$$\frac{\xi^L}{\delta^{T^L}} = \frac{\psi\left[\upsilon(1-\lambda^H)^{s-1}\lambda^H + (1-\upsilon)(1-\lambda^L)^{s-1}\lambda^L\right]f_2(j,T^L)}{P^H_{T^H}[f_1(j,T^H)f_2(s,T^L) - f_1(s,T^H)f_2(j,T^L)]} > 0,$$

$$\frac{\xi^H}{\delta^{T^H}} = \frac{\psi\left[\upsilon(1-\lambda^H)^{j-1}\lambda^H + (1-\upsilon)(1-\lambda^L)^{j-1}\lambda^L\right]f_1(s,T^H)}{P^L_{T^L}[f_1(j,T^H)f_2(s,T^L) - f_1(s,T^H)f_2(j,T^L)]} > 0,$$

which also implies that $f_2(j, T^L)$ and $f_1(s, T^H)$ must be of the same sign.

Assume $s = T^H > \hat{T}^H$. Then $f_1(s, T^H) < 0$ and the optimal contract involves

$$x^H = \frac{\beta_0\delta^{T^H}P^L_{T^L}f_2(T^H,T^L)y^H_{T^H} - \beta_0\delta P^L_{T^L}f_2(1,T^L)y^L_1}{\delta^{T^H}\psi} + \frac{P^H_{T^L}\left(\delta^{T^L}P^L_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right) - \delta^{T^H}P^L_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right)\right)}{\delta^{T^H}\psi} = 0;$$

$$x^L = \frac{\beta_0P^H_{T^H}\delta f_1(1,T^H)y^L_1 - \beta_0\delta^{T^H}P^H_{T^H}f_1(T^H,T^H)y^H_{T^H}}{\delta^{T^L}\psi} + \frac{P^L_{T^H}\left(\delta^{T^L}P^H_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right) - \delta^{T^H}P^H_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right)\right)}{\delta^{T^L}\psi} = 0.$$

Since Case B.2 is possible only if $\delta^{T^H}P^L_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right) - \delta^{T^L}P^L_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right) > 0$[49], we have a contradiction since $-f_2(1, T^L) > 0$ and $f_2(T^H, T^L) > 0$ imply that $x^H > 0$. As a result, $s = 1$. Since $f_2(j, T^L)$ and $f_1(s, T^H)$ must be of the same sign we have $j = T^L > \hat{T}^L$.

If $T^L > \hat{T}^L$, from the binding incentive compatibility constraints, we derive the optimal payments:

$$y^H_1 = \frac{\delta^{T^L}P^H_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right)(1-\lambda^L)^{T^L-1}\lambda^L - \delta^{T^H}P^L_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right)(1-\lambda^H)^{T^L-1}\lambda^H}{\beta_0\delta\lambda^H\lambda^L\left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)} \geq 0;$$

$$y^L_{T^L} = \frac{\left(\delta^{T^L}\lambda^L P^H_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right) - \delta^{T^H}\lambda^H P^L_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right)\right)}{\beta_0\delta^{T^L}\lambda^L\lambda^H\left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)} > 0.$$

This completes the proof of Lemma 5. *Q.E.D.*

We now prove that $\hat{T}^L > T^L(<)$ for high (small) values of γ.

[49] Otherwise the $(IC^{H,L})$ is not binding.

Lemma 6. There exists a unique value of γ^* such that $\hat{T}^L > T^L$ (<) for any $\gamma > \gamma^*$ (<).

Proof: We formally defined $\hat{T}^L$ as: $\frac{(1-\lambda^H)^{\hat{T}^L-1}\lambda^H}{(1-\lambda^L)^{\hat{T}^L-1}\lambda^L} \equiv \frac{P^H_{T^L}}{P^L_{T^L}}$, for any T^L. This explicitly determines $\hat{T}^L$ as a function of T^L: $\hat{T}^L(T^L) = 1 + \log_{\left(\frac{1-\lambda^H}{1-\lambda^L}\right)} \frac{P^H_{T^L}}{P^L_{T^L}} \frac{\lambda^H}{\lambda^L}$.

We will prove next that there exist a unique value of $\ddot{T}^L > 0$ such that $\hat{T}^L > T^L$ (<) for any $T^L < \ddot{T}^L$ (>). With that aim, we define the function $f(T^L) \equiv \hat{T}^L(T^L) - T^L = 1 + \log_{\left(\frac{1-\lambda^H}{1-\lambda^L}\right)} \frac{P^H_{T^L}}{P^L_{T^L}} \frac{\lambda^H}{\lambda^L} - T^L$

$$= 1 + \log_{\left(\frac{1-\lambda^H}{1-\lambda^L}\right)} \frac{\lambda^H}{\lambda^L} + \log_{\left(\frac{1-\lambda^H}{1-\lambda^L}\right)} \frac{P^H_{T^L}}{P^L_{T^L}} - T^L.$$

$$\text{Then } \frac{df}{dT^L} = \frac{\left(\beta_0(1-\lambda^H)^{T^L}\ln(1-\lambda^H)\right)P^L_{T^L} - P^H_{T^L}\left(\beta_0(1-\lambda^L)^{T^L}\ln(1-\lambda^L)\right)}{\frac{P^H_{T^L}}{P^L_{T^L}}\ln\left(\frac{1-\lambda^H}{1-\lambda^L}\right)\left(P^L_{T^L}\right)^2} - 1$$

$$= \frac{\left(\beta_0(1-\lambda^H)^{T^L}\ln(1-\lambda^H)\right)P^L_{T^L} - P^H_{T^L}\left(\beta_0(1-\lambda^L)^{T^L}\ln(1-\lambda^L)\right)}{P^L_{T^L}P^H_{T^L}\ln\left(\frac{1-\lambda^H}{1-\lambda^L}\right)} - 1$$

$$= \frac{P^L_{T^L}\ln(1-\lambda^H)\left(\beta_0(1-\lambda^H)^{T^L} - P^H_{T^L}\right) + P^H_{T^L}\ln(1-\lambda^L)\left(P^L_{T^L} - \beta_0(1-\lambda^L)^{T^L}\right)}{P^L_{T^L}P^H_{T^L}\ln\left(\frac{1-\lambda^H}{1-\lambda^L}\right)} = \frac{(1-\beta_0)\left(P^H_{T^L}\ln(1-\lambda^L) - P^L_{T^L}\ln(1-\lambda^H)\right)}{P^L_{T^L}P^H_{T^L}\ln\left(\frac{1-\lambda^H}{1-\lambda^L}\right)}.$$

Since $P^H_{T^L} < P^L_{T^L}$ and $|\ln(1-\lambda^H)| > |\ln(1-\lambda^L)|$, $P^H_{T^L}\ln(1-\lambda^L) - P^L_{T^L}\ln(1-\lambda^H) > 0$ and, as a result, $\frac{df}{dT^L} < 0$. Since $f(0) > 0$ there is only one point where $f(\ddot{T}^L) = 0$. Thus, there exist a unique value of $\ddot{T}^L$ such that $\hat{T}^L > T^L$ (<) for any $T^L < \ddot{T}^L$ (>). Furthermore, $\ddot{T}^L$ >0. Finally, since the optimal T^L is strictly decreasing in γ, and $f(\cdot)$ is independent of γ, it follows that there exists a unique value of γ^* such that $\hat{T}^L > T^L$ (<) for any $\gamma > \gamma^*$ (<). *Q.E.D.*

Finally, we consider the case when the likelihood ratio of reaching the last period of the experimentation stage is the same for both types, $\frac{P^H_{T^H}}{P^L_{T^H}} = \frac{P^H_{T^L}}{P^L_{T^L}}$.

Case B.2: *knife-edge case when* $\psi = P^H_{T^H}P^L_{T^L} - P^H_{T^L}P^L_{T^H} = 0$.

Define a $\hat{T}^H$ similarly to $\hat{T}^L$, as done in Lemma 1, by $\frac{P^L_{T^H}}{P^H_{T^H}} = \frac{(1-\lambda^H)^{\hat{T}^H-1}\lambda^H}{(1-\lambda^L)^{\hat{T}^H-1}\lambda^L}$.

Claim B.2.1. $P^H_{T^H}P^L_{T^L} - P^H_{T^L}P^L_{T^H} = 0 \Leftrightarrow \hat{T}^H = \hat{T}^L$ for any T^H, T^L.

Proof: Recall that $\hat{T}^L$ was determined by $\frac{P^H_{T^L}}{P^L_{T^L}} = \frac{(1-\lambda^L)^{\hat{T}^L-1}\lambda^L}{(1-\lambda^H)^{\hat{T}^L-1}\lambda^H}$. Next, $P^H_{T^H}P^L_{T^L} - P^H_{T^L}P^L_{T^H} = 0 \Leftrightarrow \frac{P^L_{T^H}}{P^H_{T^H}} = \frac{P^L_{T^L}}{P^H_{T^L}}$, which immediately implies that

$$P^H_{T^H}P^L_{T^L} - P^H_{T^L}P^L_{T^H} = 0 \Leftrightarrow \frac{(1-\lambda^H)^{\hat{T}^H-1}\lambda^H}{(1-\lambda^L)^{\hat{T}^H-1}\lambda^L} = \frac{(1-\lambda^H)^{\hat{T}^L-1}\lambda^H}{(1-\lambda^L)^{\hat{T}^L-1}\lambda^L};$$

$$\left(\frac{1-\lambda^H}{1-\lambda^L}\right)^{\hat{T}^H-\hat{T}^L} = 1 \text{ or, equivalently } \hat{T}^H = \hat{T}^L. \qquad \textit{Q.E.D.}$$

We prove now that the principal will choose T^L and T^H optimally such that $\psi = 0$ only if $T^L > \hat{T}^L$.

Lemma B.2.1. $P^H_{T^H}P^L_{T^L} - P^H_{T^L}P^L_{T^H} = 0 \Rightarrow T^L > \hat{T}^L$, $\xi^H > 0$, $\xi^L > 0$, $\alpha^H_t > 0$ for $t > 1$ and $\alpha^L_t > 0$ for $t < T^L$ (it is optimal to set $x^L = x^H = 0$, $y^H_t = 0$ for $t > 1$ and $y^L_t = 0$ for $t < T^L$).

Proof: Labeling $\{\alpha^H_t\}_{t=1}^{T^H}$, $\{\alpha^L_t\}_{t=1}^{T^L}$, α^H, α^L, ξ^H and ξ^L as the Lagrange multipliers of the constraints associated with (LLS^H_t), (LLS^L_t), $(IC^{H,L})$, $(IC^{L,H})$, $\left(LLF^H_{T^H}\right)$ and $\left(LLF^L_{T^L}\right)$ respectively, we can rewrite the Kuhn-Tucker conditions as follows:

$\frac{\partial \mathcal{L}}{\partial x^H} = -\upsilon\delta^{T^H}P^H_{T^H} + \xi^H = 0$, which implies that $\xi^H > 0$ and, as a result, $x^H = 0$;

$\frac{\partial \mathcal{L}}{\partial x^L} = -(1-\upsilon)\delta^{T^L}P^L_{T^L} + \xi^L = 0$, which implies that $\xi^L > 0$ and, as a result, $x^L = 0$;

$\frac{\partial \mathcal{L}}{\partial y^H_t} = -\upsilon(1-\lambda^H)^{t-1}\lambda^H + \alpha^H P^L_{T^L}f_2(t,T^L) - \alpha^L P^H_{T^H}f_1(t,T^H) + \frac{\alpha^H_t}{\delta^t\beta_0} = 0$ for $1 \le t \le T^H$;

$\frac{\partial \mathcal{L}}{\partial y^L_t} = -(1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L - \alpha^H P^L_{T^L}f_2(t,T^L) + \alpha^L P^H_{T^H}f_1(t,T^H) + \frac{\alpha^L_t}{\delta^t\beta_0} = 0$ for $1 \le t \le T^L$.

Q.E.D.

Similar results to those from Lemma 2 hold in this case as well.

Lemma B.2.2. There exists *at most* one time period $1 \le j \le T^L$ such that $y^L_j > 0$ and *at most* one time period $1 \le s \le T^H$ such that $y^H_s > 0$.

Proof: Assume to the contrary that there are two distinct periods $1 \le k, m \le T^H$ such that $k \ne m$ and $y^H_k, y^H_m > 0$. Then from the Kuhn-Tucker conditions it follows that

$$-\upsilon(1-\lambda^H)^{k-1}\lambda^H + \alpha^H P^L_{T^L}f_2(k,T^L) - \alpha^L P^H_{T^H}f_1(k,T^H) = 0,$$

and, in addition, $-\upsilon(1-\lambda^H)^{m-1}\lambda^H + \alpha^H P^L_{T^L}f_2(m,T^L) - \alpha^L P^H_{T^H}f_1(m,T^H) = 0$.

Combining the two equations together, $\alpha^L P^H_{T^H}\left(f_1(k,T^H)f_2(m,T^L) - f_1(m,T^H)f_2(k,T^L)\right)$ $+\upsilon\lambda^H\left((1-\lambda^H)^{k-1}f_2(m,T^L) - (1-\lambda^H)^{m-1}f_2(k,T^L)\right) = 0$, which can be rewritten as follows[50]:

[50] It is straightforward that $f_1(k,T^H)f_2(m,T^L) - f_1(m,T^H)f_2(k,T^L)$ $= \psi\frac{\lambda^H\lambda^L}{P^H_{T^H}P^L_{T^L}}[(1-\lambda^H)^{m-1}(1-\lambda^L)^{k-1} - (1-\lambda^L)^{m-1}(1-\lambda^H)^{k-1}]$.

$$\frac{P_{T^L}^H}{P_{T^L}^L}\lambda^L((1-\lambda^H)^{k-1}(1-\lambda^L)^{m-1}-(1-\lambda^H)^{m-1}(1-\lambda^L)^{k-1})=0,$$

$\left(\frac{1-\lambda^H}{1-\lambda^L}\right)^{m-k}=1$, which implies that $m=k$ and we have a contradiction.

In the same way, there exists *at most* one time period $1\le j\le T^L$ such that $y_j^L>0$. *Q.E.D.*

Lemma B.2.3: Both types may be rewarded for success only at *extreme* time periods, i.e. only at *the last* or *the first* period of the experimentation stage.

Proof: Since (See Lemma B.2.2) there exists only one time period $1\le s\le T^H$ such that $y_s^H>0$ ($\alpha_s^H=0$) it follows that $-\upsilon(1-\lambda^H)^{s-1}\lambda^H+\alpha^H P_{T^L}^L f_2(s,T^L)-\alpha^L P_{T^H}^H f_1(s,T^H)=0$ and $-\upsilon(1-\lambda^H)^{t-1}\lambda^H+\alpha^H P_{T^L}^L f_2(t,T^L)-\alpha^L P_{T^H}^H f_1(t,T^H)=-\frac{\alpha_t^H}{\delta^t\beta_0}$ for $1\le t\ne s\le T^H$.

Combining the equations together, $\alpha^L P_{T^H}^H\big(f_1(s,T^H)f_2(t,T^L)-f_1(t,T^H)f_2(s,T^L)\big)$ $+\upsilon\lambda^H\big((1-\lambda^H)^{s-1}f_2(t,T^L)-(1-\lambda^H)^{t-1}f_2(s,T^L)\big)=-\frac{\alpha_t^H}{\delta^t\beta_0}f_2(s,T^L)$, which can be rewritten as follows:

$$\frac{P_{T^L}^H(1-\lambda^H)^{t-1}(1-\lambda^L)^{t-1}}{P_{T^L}^L}((1-\lambda^H)^{s-t}-(1-\lambda^L)^{s-t})=-\frac{\alpha_t^H}{\delta^t\beta_0}f_2(s,T^L)\text{ for }1\le t\ne s\le T^H.$$

If $f_2(s,T^L)>0$ ($s>\hat{T}^H$) then $(1-\lambda^H)^{s-t}-(1-\lambda^L)^{s-t}<0$, which implies that $t<s$ for $1\le t\ne s\le T^H$ and it must be that $s=T^H>\hat{T}^H$. If $f_2(s,T^L)<0$ ($s<\hat{T}^H$) then $(1-\lambda^H)^{s-t}-(1-\lambda^L)^{s-t}<0$, which implies that $t>s$ for $1\le t\ne s\le T^H$ and it must be that $s=1$. In a similar way, for $1\le j\le T^L$ such that $y_j^L>0$ it must be that either $j=1$ or $j=T^L>\hat{T}^L$.

Finally, from $\frac{\partial\mathcal{L}}{\partial y_1^H}=-\upsilon\lambda^H+\alpha^H P_{T^L}^L f_2(1,T^L)-\alpha^L P_{T^H}^H f_1(1,T^H)=0$ when $y_1^H>0$ and $\frac{\partial\mathcal{L}}{\partial y_1^L}=-(1-\upsilon)\lambda^L-\alpha^H P_{T^L}^L f_2(1,T^L)+\alpha^L P_{T^H}^H f_1(1,T^H)=0$ when $y_1^L>0$ we have a contradiction. As a result, $y_1^H>0$ implies $y_{T^L}^L>0$ with $T^L>\hat{T}^L$. *Q.E.D.*

II. Optimal length of experimentation **(Proof of Proposition 2)**

Using the binding (IC) constraints and the results in Proposition 1, we can now derive the expected utility or rent for each type. In Case A in the proof of Proposition 1, only $(IC^{H,L})$ is binding, and the rents to the high and low types are

$$U_A^H=\delta^{T^L}P_{T^L}^H\Delta c_{T^L+1}q^L(c_{T^L+1}^L),$$
$$U_A^L=0,$$

where the subscript A refers to Case A. In Case B in the proof of Proposition 1, both $(IC^{H,L})$ and $(IC^{L,H})$ are binding, and the rents to the low and high types are

$$U_B^L = \frac{(1-\lambda^L)^{T^L-1}\left(\delta^{T^L}\lambda^L P_{T^L}^H \Delta c_{T^L+1} q^L\left(c_{T^L+1}^L\right) - \delta^{T^H}\lambda^H P_{T^H}^L \Delta c_{T^H+1} q^H\left(c_{T^H+1}^H\right)\right)}{\lambda^H\left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)},$$

$$U_B^H = \frac{\delta^{T^L}\Delta c_{T^L+1}(1-\lambda^L)^{T^L-1}\lambda^L q^L\left(c_{T^L+1}^L\right) - \delta^{T^H} P_{T^H}^L \Delta c_{T^H+1}(1-\lambda^H)^{T^L-1}\lambda^H q^H\left(c_{T^H+1}^H\right)}{\lambda^L\left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)},$$

where the subscript B refers to Case B.

Since T^L and T^H affect the information rents, there will be a distortion in the duration of the experimentation stage for both types depending on whether we are in Case A ($(IC^{L,H})$ is slack) or Case B (both $(IC^{H,L})$ and $(IC^{L,H})$ are binding.)

In Case A, the high type's rent U_A^H is not affected by T^H. Therefore, the F.O.C. with respect to T^H is identical to that under first best since $\Omega^L(\varpi^L)$ does not depend on T^H:

$$\frac{\partial\, \Omega^H(\varpi^H)}{\partial\, T^H} = 0,$$

or, equivalently, $T_{SB}^H = T_{FB}^H$ when $(IC^{L,H})$ is not binding.

Similarly, since $\Omega^H(\varpi^H)$ does not depend on T^L but U_A^H does depend on T^L, there will be a distortion in the duration of the experimentation stage for the low type:

$$\frac{\partial\, (\Omega^L(\varpi^L) - \upsilon U_A^H)}{\partial\, T^L} = 0.$$

Since U_A^H is non-monotonic in T^L, it is possible, in general, to have $T_{SB}^L > T_{FB}^L$ or $T_{SB}^L < T_{FB}^L$.

In Case B, the exact values of the rent depend on both T^L and T^H, and $U_B^H > 0$ and $U_B^L \geq 0$. The *F.O.C.* is given by

$$\frac{\partial\, (\nu\Omega^H(\varpi^H) + (1-\nu)\Omega^L(\varpi^L) - \upsilon U_B^H - (1-\upsilon)U_B^L)}{\partial\, T^\theta} = 0.$$

It is possible, in general, to have $T_{SB}^H > T_{FB}^H$ or $T_{SB}^H < T_{FB}^H$ and $T_{SB}^L > T_{FB}^L$ or $T_{SB}^L < T_{FB}^L$.

We next provide sufficient conditions for over-experimentation in T^H.

Define a function $\zeta(t) \equiv P_t^L(\beta_{t+1}^L - \beta_{t+1}^H)$. Note that this function $\zeta(t)$ is directly related to the difference in expected costs as $(\overline{c} - \underline{c})\zeta(t) \equiv P_t^L \Delta c_{t+1}$. In step 1, we characterize values of λs and β_0 such that $\zeta(t)$ is monotonically increasing in t. In step 2, we characterize the set of λs and β_0 such that both rents are decreasing in T^H, which implies over-experimentation in T^H.

Step 1. **We show that $\frac{d\zeta(t)}{dt} > 0$ if λ^L and λ^H are close to each other and β_0 is high enough ($\lambda^L < \lambda^H < \dot{\lambda}^H(\lambda^L)$ and $\beta_0 > \underline{\underline{\beta}}_0$).**

Proof of step 1:

Recalling that $P_T^\theta = 1 - \beta_0 + \beta_0(1-\lambda^\theta)^T$, and $\beta_t^\theta = \frac{\beta_0(1-\lambda^\theta)^{t-1}}{\beta_0(1-\lambda^\theta)^{t-1}+(1-\beta_0)}$, we can rewrite $\zeta(t)$:

$$\zeta(t) = (1-\beta_0+\beta_0(1-\lambda^L)^t)\left(\frac{\beta_0(1-\lambda^L)^t}{\beta_0(1-\lambda^L)^t+(1-\beta_0)} - \frac{\beta_0(1-\lambda^H)^t}{\beta_0(1-\lambda^H)^t+(1-\beta_0)}\right)$$

$$= \frac{\beta_0\left((1-\lambda^L)^t\left(\beta_0(1-\lambda^H)^t+(1-\beta_0)\right)-(1-\lambda^H)^t\left(1-\beta_0+\beta_0(1-\lambda^L)^t\right)\right)}{\beta_0(1-\lambda^H)^t+(1-\beta_0)}$$

$$= \frac{\beta_0(1-\beta_0)\left((1-\lambda^L)^t-(1-\lambda^H)^t\right)}{\left(\beta_0(1-\lambda^H)^t+(1-\beta_0)\right)} = \frac{\beta_0(1-\beta_0)\left((1-\lambda^L)^t-(1-\lambda^H)^t\right)}{P_t^H}.$$

Therefore, $\frac{d\zeta(t)}{dt} = \frac{\left((1-\lambda^L)^t ln(1-\lambda^L)-(1-\lambda^H)^t ln(1-\lambda^H)\right)P_t^H-\beta_0(1-\lambda^H)^t ln(1-\lambda^H)\left((1-\lambda^L)^t-(1-\lambda^H)^t\right)}{(P_t^H)^2\frac{1}{\beta_0(1-\beta_0)}} =$

$$\frac{(1-\lambda^L)^t P_t^H ln(1-\lambda^L)-(1-\lambda^H)^t P_t^L ln(1-\lambda^H)}{(P_t^H)^2\frac{1}{\beta_0(1-\beta_0)}}.$$

The function $\zeta(t)$ increases with t if and only if $\phi(\lambda^H) > 0$, where

$$\phi(\lambda^H) = (1-\lambda^L)^t P_t^H\, ln(1-\lambda^L) - (1-\lambda^H)^t P_t^L ln(1-\lambda^H).$$

We prove next that $\phi(\lambda^H) > 0$ if λ^H is not too far apart from λ^L and β_0 is high enough: there exists $\dot{\lambda}^H(\lambda^L) > \lambda^L$ and $\underline{\underline{\beta}}_0 < 1$ such that $\phi(\lambda^H) > 0$ if $\lambda^L < \lambda^H < \dot{\lambda}^H(\lambda^L)$ and $\beta_0 > \underline{\underline{\beta}}_0$.

We first prove that $\frac{d\,\phi(\lambda^H)}{d\,\lambda^H} > 0$ if λ^H is sufficiently close to λ^L and β_0 is high enough.

Consider the derivative of $\phi(\lambda^H)$ with respect to λ^H:

$$\frac{d\,\phi(\lambda^H)}{d\,\lambda^H} =$$

$$-\beta_0 t(1-\lambda^H)^{t-1}(1-\lambda^L)^t\, ln(1-\lambda^L) - P_t^L\left(-t(1-\lambda^H)^{t-1} ln(1-\lambda^H) + (1-\lambda^H)^t \frac{(-1)}{1-\lambda^H}\right) =$$

$$\left(-\beta_0 t(1-\lambda^L)^t\, ln(1-\lambda^L) + P_t^L(tln(1-\lambda^H)+1)\right)(1-\lambda^H)^{t-1}.$$

Consider next $\frac{d\,\phi(\lambda^H)}{d\,\lambda^H}$ and take the limit for $\lambda^H \to \lambda^L$

$$\lim_{\lambda^H\to\lambda^L}\frac{d\,\phi(\lambda^H)}{d\,\lambda^H} = \left(-\beta_0 t(1-\lambda^L)^t\, ln(1-\lambda^L) + P_t^L(tln(1-\lambda^L)+1)\right)(1-\lambda^L)^{t-1} =$$

$$(t(1-\beta_0)\, ln(1-\lambda^L) + 1 - \beta_0 + \beta_0(1-\lambda^L)^t)\;(1-\lambda^L)^{t-1},$$

which is positive if $t(1-\beta_0)\,ln(1-\lambda^L)+1-\beta_0+\beta_0(1-\lambda^L)^t>0$.

We define a value $\underline{\underline{\beta}}_0$ by choosing $t=1$ or $\overline{T}\equiv\max\{T^H,T^L\}$ such that the smallest value of $\lim_{\lambda^H\to\lambda^L}\frac{d\,\phi(\lambda^H)}{d\,\lambda^H}$ is equal to zero:

$$\underline{\underline{\beta}}_0:\quad \overline{T}(1-\beta_0)\,ln(1-\lambda^L)+1-\beta_0+\beta_0(1-\lambda^L)=0 \text{ or, equivalently}$$

$$\underline{\underline{\beta}}_0\equiv\frac{\overline{T}\,ln(1-\lambda^L)+1}{\overline{T}\,ln(1-\lambda^L)+\lambda^L}.$$ [51]

Therefore, $\lim_{\lambda^H\to\lambda^L}\frac{d\,\phi(\lambda^H)}{d\,\lambda^H}>0$ for any λ^L if $\beta_0>\underline{\underline{\beta}}_0$.

Since the function $\phi(\lambda^H)$ is continuous in λ^H and becomes negative as λ^H approaches one ($\lim_{\lambda^H\to 1}\phi(\lambda^H)<0$), there exists $\dot{\lambda}^H(\lambda^L)>\lambda^L$, such that $\phi(\lambda^H)>0$ if $\lambda^L<\lambda^H<\dot{\lambda}^H(\lambda^L)$.[52] We define a value $\dot{\lambda}^H$ such that the function $\phi(\lambda^H)$ is equal to zero:

$$\phi(\dot{\lambda}^H)\equiv 0.$$

As a result, the function $\zeta(t)$ is an increasing function of t if $\lambda^L<\lambda^H<\dot{\lambda}^H(\lambda^L)$ and $\beta_0>\underline{\underline{\beta}}_0$.

Step 2. Both rents U^H and U^L are decreasing in T^H, and there is over-experimentation in T^H.

Proof of Step 2: over experimentation in T^H.

If $(IC^{L,H})$ is not binding, the rent to the high type is $U^H=\delta^{T^L}P^H_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right)$, which does not depend on T^H. Therefore, there is no distortion in T^H, $T^H_{FB}=T^H_{SB}$.

If both $(IC^{H,L})$ and $(IC^{L,H})$ are binding, using the function $\zeta(t)\equiv P^L_t(\beta^L_{t+1}-\beta^H_{t+1})$, we can rewrite U^L and U^H as:

$$U^L=\frac{\left(\frac{\lambda^L\delta^{T^L}\zeta(T^L)P^H_{T^L}}{P^L_{T^L}}q^L\left(c^L_{T^L+1}\right)-\lambda^H\delta^{T^H}\zeta(T^H)q^H\left(c^H_{T^H+1}\right)\right)}{\lambda^H\left(1-\left(\frac{1-\lambda^H}{1-\lambda^L}\right)^{T^L-1}\right)}(\overline{c}-\underline{c}),\text{ and}$$

[51] Note that $\frac{\overline{T}\,ln(1-\lambda^L)+1}{\overline{T}\,ln(1-\lambda^L)+\lambda^L}<1$ for any $\lambda^L<1$.

[52] $\lim_{\lambda^H\to 1}\phi(\lambda^H)=\lim_{\lambda^H\to 1}[(1-\lambda^L)^t\,P^H_t ln(1-\lambda^L)-(1-\lambda^H)^tP^L_t ln(1-\lambda^H)]=(1-\lambda^L)^t(1-\beta_0)ln(1-\lambda^L)<0$ (using l'Hôpital's rule).

$$U^H = \frac{\frac{\lambda^L \delta^{T^L} \zeta(T^L) P^H_{T^L}}{P^L_{T^L}} \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{T^L-1} q^L(c^L_{T^L+1}) - \lambda^H \delta^{T^H} \zeta(T^H) q^H(c^H_{T^H+1})}{\lambda^L \left(\left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{T^L-1} - 1\right)} (\overline{c} - \underline{c}), \text{ respectively.}$$

Note that then term δ^{T^H} cancels out from the *F.O.C.* with respect to T^H is given by

$$\frac{\partial\, (\nu\Omega^H(\varpi^H) + (1-\nu)\Omega^L(\varpi^L) - \upsilon U^H_B - (1-\upsilon)U^L_B)}{\partial\, T^\theta} = 0.$$

Therefore, if $\zeta(T^H) q^H(c^H_{T^H+1})$ is increasing in T^H, then the principal will mitigate both rents U^H and U^L by increasing T^H, and over-experimentation in T^H is optimal. First note that $q^H(c^H_{T^H+1})$ increases proportionately to $\frac{P^L_{T^H}}{P^H_{T^H}} \Delta c_{T^H+1}$.[53] Note also that $\frac{P^L_{T^H}}{P^H_{T^H}} \Delta c_{T^H+1} = \frac{\zeta(T^H)}{P^H_{T^H}} \Delta c$. Since $P^H_{T^H}$ is decreasing in T^H, $\frac{d\zeta(t)}{dt} > 0$ implies that $\frac{P^L_{T^H}}{P^H_{T^H}} \Delta c_{T^H+1}$ is increasing in T^H. Therefore, if $\frac{d\zeta(t)}{dt} > 0$, then over-experimentation in T^H is optimal. Since we proved (1) in Step 1 that $\frac{d\zeta(t)}{dt} > 0$ if $\lambda^L < \lambda^H < \dot{\lambda}^H$ and $\beta_0 > \underline{\underline{\beta}}_0$, and (2) in Supplementary Appendix B that both $(IC^{H,L})$ and $(IC^{L,H})$ are binding if $\hat{\lambda} < \lambda^L < \lambda^H < \overline{\overline{\lambda}}^H(\lambda^L)$ and $\beta_0 > \hat{\beta}_0$. Therefore we can define the sufficient values of β_0 and λ^H, as $\underline{\beta}_0 \equiv \max\left\{\underline{\underline{\beta}}_0, \hat{\beta}_0\right\}$ and $\overline{\lambda}^H(\lambda^L) \equiv \min\left\{\dot{\lambda}^H, \overline{\overline{\lambda}}^H(\lambda^L)\right\}$. As a result, we have $T^H_{SB} \geq T^H_{FB}$ if $\hat{\lambda} < \lambda^L < \lambda^H < \overline{\lambda}^H(\lambda^L)$ and $\beta_0 > \underline{\beta}_0$.

This completes the proof of Proposition 2. *Q.E.D.*

III. Optimal outputs (Proof of Proposition 3)

After success, the optimal $q^\theta_t(\overline{c})$ is efficient as it chosen to maximize $E_\theta\, \Omega^\theta(\varpi^\theta)$. After failure, we have to consider whether we are in case A or B.

Case A [when $(IC^{L,H})$ is not binding]

The following two FOCs imply that there is no distortion after failure by the high type but there will be underproduction by the low type after failure, that is, $q^L_{SB}(c^L_{T^L+1}) < q^L_{FB}(c^L_{T^L+1})$:

$$V'\left(q^L_{SB}(c^L_{T^L+1})\right) - c^L_{T^L+1} = \frac{\nu}{(1-\upsilon)} \frac{P^H_{T^L}}{P^L_{T^L}} \Delta c_{T^L+1} > 0,$$

$$V'\left(q^H_{SB}(c^H_{T^H+1})\right) - c^H_{T^H+1} = 0.$$

[53] In Proposition 3, we formally prove that $\left(V'\left(q^H(c^H_{T^H+1})\right) - c^H_{T^H+1}\right) = -\frac{P^L_{T^H} E_\theta\{(1-\lambda^\theta)^{T^L-1}\lambda^\theta\}}{\upsilon P^H_{T^H} \lambda^H \left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)} \Delta c_{T^H+1}$ if $(IC^{L,H})$ is binding.

Case B. [when $(IC^{L,H})$ is binding]
The following two FOCs imply that there will be overproduction for the high type ($q_{SB}^H(c_{T^H+1}^H) > q_{FB}^H(c_{T^H+1}^H)$) and underproduction for the low type ($q_{SB}^L(c_{T^L+1}^L) < q_{FB}^L(c_{T^L+1}^L)$) after failure. We start with the main case B.1, when $\psi \neq 0$, and consider cases when $T^L \leq \hat{T}^L$ and $T^L > \hat{T}^L$ separately.

When $T^L \leq \hat{T}^L$, we have:

$$V'\left(q^L(c_{T^L+1}^L)\right) - c_{T^L+1}^L = \frac{E_\theta\{\lambda^\theta\}P_{T^L}^H}{(1-\upsilon)\left(P_{T^L}^L\lambda^H - P_{T^L}^H\lambda^L\right)}\Delta c_{T^L+1} > 0,$$

$$V'\left(q^H(c_{T^H+1}^H)\right) - c_{T^H+1}^H = -\frac{P_{T^H}^L E_\theta\left\{P_{T^L}^\theta\right\}}{\upsilon P_{T^H}^H\left(P_{T^L}^L\lambda^H - P_{T^L}^H\lambda^L\right)}\Delta c_{T^H+1} < 0.$$

When $T^L > \hat{T}^L$, we have:

$$V'\left(q^L(c_{T^L+1}^L)\right) - c_{T^L+1}^L = \frac{P_{T^L}^H(1-\lambda^L)^{T^L-1}E_\theta\{\lambda^\theta\}}{(1-\upsilon)P_{T^L}^L\lambda^L\left((1-\lambda^L)^{T^L-1}-(1-\lambda^H)^{T^L-1}\right)}\Delta c_{T^L+1} > 0,$$

$$V'\left(q^H(c_{T^H+1}^H)\right) - c_{T^H+1}^H = -\frac{P_{T^H}^L E_\theta\left\{(1-\lambda^\theta)^{T^L-1}\lambda^\theta\right\}}{\upsilon P_{T^H}^H\lambda^H\left((1-\lambda^L)^{T^L-1}-(1-\lambda^H)^{T^L-1}\right)}\Delta c_{T^H+1} < 0,$$

In the knife-edge case B.2, when $\psi = 0$, the relevant FOCs are:

$$V'\left(q^L(c_{T^L+1}^L)\right) - c_{T^L+1}^L = \frac{\delta^{T^L}\nu\lambda^H P_{T^L}^H P_{T^H}^L}{f_1(1,T^H)P_{T^H}^H}\Delta c_{T^L+1} > 0,$$

$$V'\left(q^H(c_{T^H+1}^H)\right) - c_{T^H+1}^H = -\frac{\delta^{T^H}\nu\lambda^H P_{T^H}^L}{f_1(1,T^H)}\Delta c_{T^H+1} < 0.$$

Q.E.D.