Learning from Failures:

Optimal Contracts for Experimentation and Production*

Fahad Khalil, Jacques Lawarree
Alexander Rodivilov**

October 31, 2019

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 his efficiency in 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 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.
** 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 that 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 his efficiency in experimentation. If the agent is not efficient at experimenting, a poor result from his experiments only provides weak evidence of low profitability of the project. However, if the principal is misled into believing that the agent is highly efficient, she becomes more pessimistic than the agent. 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 finding "good news", i.e., the agent finds out that production cost is low.[3] After success, experimentation stops, and production occurs. If experimentation continues without success, the expected cost increases, and both principal and

[1] Other applications are 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] 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 our key insights hold in the case of success being bad news.

agent become more pessimistic 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.[4] 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 nascent 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 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 low. When experimentation fails, an inefficient agent pretending to be efficient will have a lower expected cost of production compared to the principal. Mistakenly believing the agent is efficient, the principal will then overcompensate him in the production stage. Therefore, an inefficient agent must be paid a rent to prevent him from overstating his efficiency. An important element of our setting is that the efficient type may also get a rent. The reason is that the efficiency parameter also enters directly the principal's objective function. As a result, we have what is called a common values problem in contract theory.[5] It is known that in such models both efficient and inefficient types can get rent in equilibrium due to a conflict between the principal's preference and the screening role of contracts. When an efficient agent misreports, he faces a gamble: he can collect the inefficient agent's rent, but he faces a risk of being undercompensated at the production stage if experimentation fails since he is relatively more pessimistic than the

---

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

[5] See, e.g., Laffont and Martimort (2003).

principal. Both efficient and inefficient types can get rent in equilibrium as the principal trades off efficiency in experimentation with rent in the production stage.

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 screening strategy for the principal, whereas under-experimentation is the standard result in existing models of experimentation.[6] 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 costs. The second reason is due to our finding that the difference in expected costs between the principal and the misreporting agent is non-monotonic in time. 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 inefficient agent always gets a rent, we expect, and indeed find, that the output of the efficient agent is distorted downward. This is reminiscent of a standard adverse selection problem.

We find another effect: the output of the inefficient agent is distorted *upward*. This is the case when the efficient 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 efficient type faces a gamble when misreporting his type as inefficient. While he has the chance to collect the rent of the inefficient type, he also faces a cost if experimentation fails. Since he is then relatively more pessimistic than the principal, he will be under-compensated at the production stage relative to the inefficient type. The principal can increase the cost of lying by asking the inefficient type to produce more. A higher output for the inefficient agent makes it costlier for the efficient agent who must produce more output with higher expected costs.

---

[6] 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.[7] Our result suggests that the principal is more likely to tolerate failures in industries where 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.[8] In an extension section, we introduce ex post moral hazard by assuming that success is privately observed by the agent. This leads to moral hazard rent in every period as incentives must be provided to the agent to reveal success. Therefore, in addition to the previously derived asymmetric information rent, the agent receives a moral hazard rent in every period. 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.[9] 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

---

[7] 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 (2017)). 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.

[8] 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.

[9] Early papers are Cremer 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.[10] Recent exceptions that introduce adverse selection are Gomes, Gottlieb and Maestri (2016) and Halac, Kartik and Liu (2016).[11] 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.[12]

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 good-news model under adverse selection with public success. In section 3, we consider extensions and robustness checks. In particular, we study ex post moral hazard where the agent can hide success, and the case where success is bad news.

---

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

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

[12] 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 good 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 $c = \underline{c}$ is denoted by $\beta_0 \in (0,1)$. Before the actual *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.[13] 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.[14]

In the main part of the paper, information gathering takes the form of looking for good news (see section 3.2 for the case of bad news). If the cost is low, the agent learns it with probability λ in each period $t \leq T$. If the agent learns that the cost is low (*good news*) in a period t, we will 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 low. In section 3.1, 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 learn that cost is low 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

[13] 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 2.2.3).

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

beliefs differently during the experimentation stage. The agent's private information about his efficiency λ determines his 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: $\lambda^\theta = Pr(type\ \theta\ learns\ c = \underline{c} | c = \underline{c})$, where $0 < \lambda^L < \lambda^H < 1$.[15] If experimentation fails to reveal low 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 low 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: $c_t^\theta = \beta_t^\theta \underline{c} + (1-\beta_t^\theta)\,\overline{c}$.

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

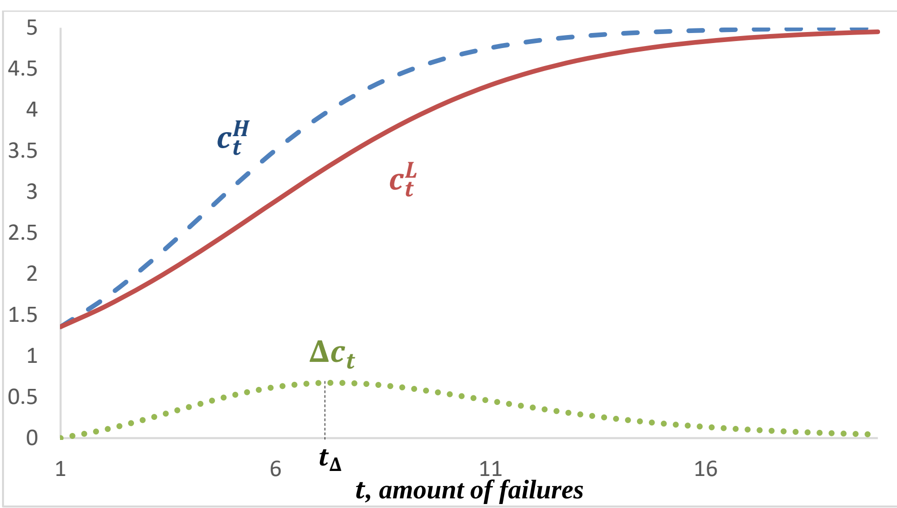

Figure 1. Expected cost with $\lambda^H = 0.35$, $\lambda^L = 0.2$, $\beta_0 = 0.7$, $\underline{c} = 0.5$, $\overline{c} = 5$.

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

Third, we also note the important property that the difference in the expected cost, $\Delta c_t = c_t^H - c_t^L > 0$, is a *non-monotonic* function of time: initially increasing and then decreasing, reaching a maximum at time period t_Δ.[16] Intuitively, each type starts with the same expected cost $\beta_0 \underline{c} + (1-\beta_0)\overline{c}$. The expected costs diverge as each type of agent updates differently, but they eventually have to converge to $\overline{c}$. When λs are close to each other, the Δc_t function is relatively flat. As λ^H becomes larger relative to λ^L, the Δc_t function becomes more skewed to the left, moving t_Δ to the left, which makes the decreasing part of Δc_t relatively larger.

The production stage

After the experimentation stage ends, production takes place. 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.[17] 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 low in a period $t \leq T$, experimentation stops, and production takes place based on $c = \underline{c}$.[18] If experimentation fails, i.e., there is no success during the experimentation stage, production occurs in period $T+1$ based on the expected cost.[19]

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.[20] In terms of notation, in the case of success we

---

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

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

[17] 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.

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

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

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

include $\underline{c}$ as an argument in the wage and output for each t. In the case of failure, we include the expected cost $c^{\hat{\theta}}_{T^{\hat{\theta}}+1}$.[21] A contract is defined formally by

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

where $T^{\hat{\theta}}$ is the maximum duration of the experimentation stage for the announced type $\hat{\theta}$, $w^{\hat{\theta}}_t(\underline{c})$ and $q^{\hat{\theta}}_t(\underline{c})$ 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 $U^{\theta}(\varpi^{\hat{\theta}})$ at time zero from a contract $\varpi^{\hat{\theta}}$:

$$U^{\theta}(\varpi^{\hat{\theta}}) = \beta_0 \sum_{t=1}^{T^{\hat{\theta}}} \delta^t \left(1-\lambda^{\theta}\right)^{t-1} \lambda^{\theta} \left(w^{\hat{\theta}}_t(\underline{c}) - \underline{c} q^{\hat{\theta}}_t(\underline{c})\right)$$

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

We explain the terms in the above expression next. Conditional on the actual cost being low, 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 high ($c = \bar{c}$), which happens with probability $1-\beta_0$, or, if the agent fails $T^{\hat{\theta}}$ times despite $c = \underline{c}$, 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.

[21] 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 $c = \underline{c}$, the experimentation stage stops, and the production stage starts with output and transfers as specified in the contract.

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

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. [22]

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 is contract farming for new crops.[23] In developing countries, large processors (such as exporters, agricultural firms, or supermarket chains) often provide local farmers with new untested, potentially more productive seeds. The local farmers test the seeds in a dynamic process of experimentation and learning, which exhibits features captured by our model. The environment features *i*) common uncertainty at the outset since both parties learn about the quality of new seeds, and *ii*) private information since the local farmers know the relevant properties of their soil and local conditions better. To limit the scope for moral hazard, strictly monitored protocols are part of the contracts which specify actions and

[22] 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.

[23] See Singh (2002).

procedures farmers must follow. Depending on the outcome of learning phase, the scale of production by the farmers is determined.

Our focus is to study the interaction between endogenous asymmetric information due to experimentation and optimal decisions that are made after the experimentation stage. The focus of the existing literature on experimentation has been on providing incentives to experiment, where success is identified as an outcome with a positive payoff. The decision ex post is not explicitly modeled. In contrast, 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 contract farmers testing new seeds before deciding how much to produce, 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.

2.1. 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:

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

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

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 $V'\left(q_{t^\theta}^\theta(\underline{c})\right) = \underline{c}$ for any t^θ. In case the agent fails, the efficient output is based on the current *expected* cost, such that $V'\left(q^\theta\left(c_{T^\theta+1}^\theta\right)\right) = c_{T^\theta+1}^\theta$. Since the expected cost is rising as long as success is not obtained, the termination date T_{FB}^θ is bounded and it is the highest t^θ such that the following condition holds:

$$\delta\beta^{\theta}_{t^\theta}\lambda^\theta\left[V\left(q^{\theta}_{t^\theta}(\underline{c})\right)-\underline{c}q^{\theta}_{t^\theta}(\underline{c})\right]+\delta(1-\beta^{\theta}_{t^\theta}\lambda^\theta)\left[V\left(q^\theta(c^{\theta}_{t^\theta+1})\right)-c^{\theta}_{t^\theta+1}q^\theta(c^{\theta}_{t^\theta+1})\right]$$
$$\geq\gamma+\left[V\left(q^\theta(c^{\theta}_{t^\theta})\right)-c^{\theta}_{t^\theta}q^\theta(c^{\theta}_{t^\theta})\right].$$

Extending the experimentation stage by one additional period costs γ, but an agent of type θ can learn that $c=\underline{c}$ with probability $\beta^{\theta}_{t^\theta}\lambda^\theta$.

Note that the first-best termination date of the experimentation stage T^{θ}_{FB} is a *non-monotonic* function of the agent's type. In Claim 1, Appendix A, we formally prove that there exists a unique value of λ^θ called $\hat{\lambda}$, such that:

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

This non-monotonicity is a result of two countervailing forces.[24] In any given period of the experimentation stage, the high type is more likely to learn $c=\underline{c}$ (conditional on the actual cost being low) since $\lambda^H>\lambda^L$. This suggests that the principal should allow the high type to experiment longer because he is relatively more efficient. However, at the same time, the high type agent becomes relatively more pessimistic with repeated failures. 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 $\left(\frac{\beta_0(1-\lambda^H)^{t-1}\lambda^H}{\beta_0(1-\lambda^L)^{t-1}\lambda^L}\right)$ 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.[25] The first-best termination date is increasing in the agent's type for small values of λ^θ when the first force (relative efficiency) dominates, but becomes decreasing for larger values when the second force (relative pessimism) becomes dominant.

---

[24] A similar intuition can be found in Halac, Kartik and Liu (2016) in a model without production.

[25] For example, if $\lambda^L=0.2$, $\lambda^H=0.4$, $\underline{c}=0.5$, $\overline{c}=20$, $\beta_0=0.5$, $\delta=0.9$, $\gamma=2$, and $V=10\sqrt{q}$, then the first-best termination date for the high type agent is $T^H_{FB}=4$, whereas it is optimal to allow the low type agent to experiment for seven periods, $T^L_{FB}=7$. However, if we now change λ^H to 0.22 and β_0 to 0.4, the low type agent is allowed to experiment less, that is, $T^H_{FB}=4>T^L_{FB}=3$.

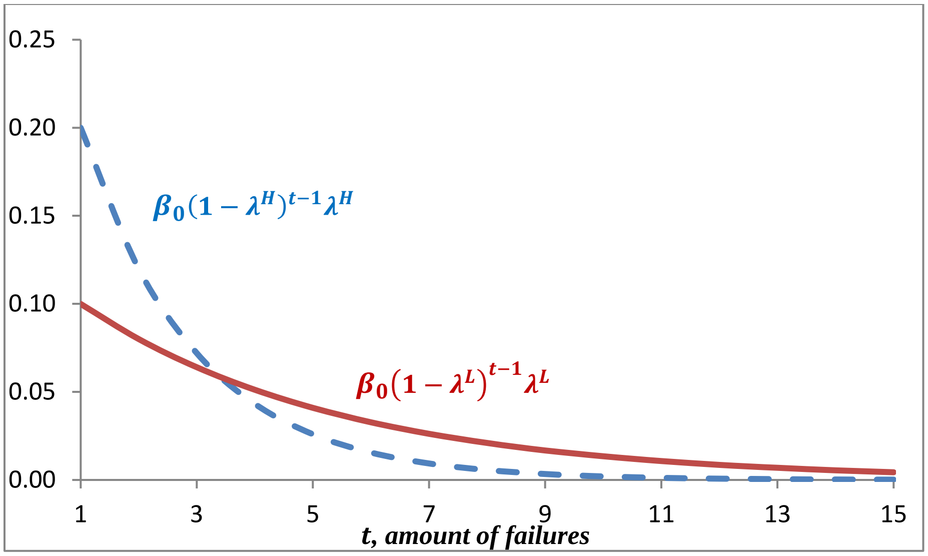

Figure 2. Probability of success with $\lambda^H = 0.4$, $\lambda^L = 0.2$, $\beta_0 = 0.5$.

2.2. Asymmetric information

2.2.1. Benchmark without experimentation

To highlight the implications of experimentation in our model, we now consider a benchmark model without experimentation but with asymmetric beliefs about expected cost in the production stage. We will use this model to illustrate why both types may want to mimic each other because of experimentation. Thus, we assume that a type θ agent's private belief is denoted by β^θ, and we define a high type to be more pessimistic than a low type about the cost being low: $\beta^H < \beta^L$.[26] The expected cost at the production stage is $c^\theta = \beta^\theta \underline{c} + (1-\beta^\theta)\overline{c}$. This implies that $c^H > c^L$, where we denote $\Delta c = c^H - c^L > 0$.

Thus, we have a standard second-best problem where the hidden parameter is the *expected* marginal cost (e.g., Baron and Myerson (1982), Laffont and Tirole (1986)), and the principal can only screen the agents with the output and payments.[27] As is well known, the two incentive constraints can be written in equilibrium as:

[26] This definition may seem counterintuitive, but our goal is pedagogical as we want to analyze a situation similar to when the agent has failed in experimentation and goes to production with private information.

[27] 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 high type's output is distorted downwards and only the low type gets a positive informational rent.

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

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

where the subscript "b" refers to the benchmark without experimentation.

In this model, the high type is not interested in misreporting. When the high type lies, he collects the rent of the low type, $\Delta c\, q(c^H)$, as part of the transfer $w(c^L)$. However, he then must produce $q(c^L)$ while being undercompensated relative to the low type as his true expected cost c^H exceeds that of the low type: $c^H > c^L$. Therefore, the $\left(IC_b^{H,L}\right)$ is never binding. Note that Δc, the low type's cost advantage, is exogenous and is also identical to the high type's cost disadvantage when he has to produce $q(c^L)$.

In contrast, there are two important modifications in our main model where agents have asymmetric efficiency in experimentation. First, the differences in expected cost are non-monotonic in t and vary for each type. They are endogenously determined by the duration of the experimentation stage for each type. Therefore, in the (IC) constraints of the main model, the rents will depend on Δc_{T^H+1} and Δc_{T^L+1} instead of a constant Δc.

Second, because of experimentation, the efficiency parameter (λ^θ) appears directly in the principal's objective function through the probabilities of success/failures. This creates a common values problem. As we know from other contract theory models with common values, the principal's preference for outcome choices can conflict with the monotonicity condition implied by the agent's (IC) constraints.[28] In our model, the principal's preference of termination dates, due the presence of λ^θ in the objective function, creates incentive for the high type to misreport, leading to both (IC) being binding.

2.2.2. Main Model with experimentation:

We now return to the main model, where all parties have the same expected cost at the outset, but asymmetric information arises because the two types learn asymmetrically in the experimentation stage.

[28] See Laffont and Martimort (2003), p. 53.

The optimal contract will have to satisfy the following incentive compatibility constraints for all θ and $\hat{\theta}$:

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

To simply 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:

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

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

We also denote with P_T^{θ} the probability that an agent of type θ 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:

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

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

We also assume that the agent must be paid his expected production costs whether experimentation succeeds or fails.[29] Therefore, we introduce the following limited liability constraints:

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

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

where the S and F denote success and failure.

[29] 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. 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.

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

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

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 $\Gamma_t = \frac{\sum_{s=1}^t \delta^s \gamma}{\delta^t}$.

Both (IC) may be binding

We now focus on why both (IC) constraints can be binding. Consider first why the low-type's $(IC^{L,H})$ constraint is binding. The reason is that a low type has an incentive to claim to be a high type in order to collect the higher transfer given to the high type to cover his higher expected cost following failure.[30] That is, the RHS of $(IC^{L,H})$ is strictly positive since $\Delta c_{T^H+1} = c_{T^H+1}^H - c_{T^H+1}^L > 0$.

Consider now the high-type's $(IC^{H,L})$ constraint. While the low type's benefit from misreporting is positive for sure ($\Delta c_{T^H+1} > 0$), the high type's benefit from misreporting his type is a *gamble*. There is a positive part since he has a chance to claim the rent of the low type. As we just explained, this part is positively related to Δc_{T^H+1} adjusted by the output and the relative probability of collecting the low type's rent. However, there is a negative part as well since the high type who misreports runs the risk of having to produce while being undercompensated. This is because the principal would pay him as a low type whose expected cost is lower when experimentation fails. This term is positively related to Δc_{T^L+1} adjusted by the output and the probability of having to produce after failure. The $(IC^{H,L})$ is binding when the positive part of the gamble dominates the negative part.

[30] We prove this result in a Claim 2 in Appendix A.

The termination dates play a key role in the sign of the gamble since they determine Δc_{T^H+1} and Δc_{T^L+1}. When the duration of the experimentation stage is identical for both types ($T^H = T^L$), we show that the gamble is negative, and the principal pays a rent only to the low type. Intuitively, the magnitudes of Δc_{T^L+1} and Δc_{T^H+1} are the same and, therefore, the cost and benefit of lying have the same magnitude for the high type. See supplementary Appendix E. However, having the same duration for both types might be suboptimal.

Because the efficiency parameter λ^θ enters directly in the principal's objective function, we have a common values problem when choosing the optimal termination dates. As shown in section 2.1, the principal's preference for first-best efficiency can require either $T^H > T^L$ or $T^H < T^L$ depending on the size of λ. When λs are small, first-best efficiency requires that $T^H > T^L$. This choice may conflict with the screening role of termination dates. When the principal chooses $T^H > T^L$, she also makes $\Delta c_{T^H+1} > \Delta c_{T^L+1}$, which implies that the benefit of lying for the high type (positive part of the gamble proportional to Δc_{T^H+1}) may exceed the cost of lying (negative part of the gamble proportional to Δc_{T^L+1}). Thus, choosing $T^H > T^L$ may make the gamble positive. The same would be true for large λs. The principal's preference for first best efficiency requires that $T^H < T^L$, and it conflicts with the screening role of T when Δc_t is decreasing. 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. We provide in Appendix B sufficient conditions for when the ($IC^{H,L}$) constraint will be binding.[31]

We conclude this section with an example with a binding ($IC^{H,L}$) to illustrate the gamble, and show how the principal can affect incentives by altering the termination dates. Consider a case where the two types are significantly different, e.g., λ^L is close to zero and λ^H is close to one so that first-best efficiency requires that $T^L = 0$ and $T^H > 0$. Suppose the low type claims being high. Since his expected cost is lower than the cost of the high type after T^H unsuccessful experiments ($c^L_{T^H+1} < c^H_{T^H+1}$), the low type must be given a rent to induce truth-telling. Consider now the incentive of the high type to claim being low. In this case, production starts immediately without experimentation under identical beliefs about expected cost

[31] These conditions separate the cases for small and large λ to account for the non-monotonicity in Δc_t and the first best termination dates.

$\left(\beta_0 \underline{c} + (1-\beta_0)\,\overline{c}\right)$. Therefore, the high type simply collects the rent of the low type without incurring the negative part of the gamble when producing. And, $(IC^{H,L})$ would be violated; first-best efficiency is in conflict with incentives.

The principal can affect the value of the gamble by altering the termination dates T^L and T^H. Consider again our simple example and suppose the principal asks the low type to (over) experiment (by) one period, $T^L = 1$. The high type now faces a risk when misreporting. If the project is bad, he will fail with probability $(1-\beta_0)$ and have to produce in period $t = 2$ knowing almost for sure that the cost is $\overline{c}$, while the principal is led to believe that the expected cost is $c_2^L = \beta_2^L \underline{c} + (1-\beta_2^L)\,\overline{c} < \overline{c}$. Therefore, by increasing the low-type's duration of experimentation, the principal can use the negative part of the gamble (under-compensation) to mitigate the high-type's incentive to lie and, therefore, relax the $(IC^{H,L})$. We study the optimal duration of experimentation in section 2.2.4.

2.2.3. 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.

There are two cases to consider. First, when $(IC^{H,L})$ is not binding, $y_t^H = x^H = 0$, the optimal contract is not unique, and the principal can use any combination of y_t^L and x^L to satisfy the binding $(IC^{L,H})$: there is no restriction on when and how the principal pays the rent to the low type as long as $\beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1}\lambda^L y_t^L + \delta^{T^L} P_{T^L}^L x^L = \delta^{T^H} P_{T^H}^L \Delta c_{T^H+1} q^H\left(c_{T^H+1}^H\right)$. Therefore, the principal can reward either early or late success, or even failure.[32] Second, when $(IC^{H,L})$ 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.

[32] See Case A in Appendix A.

Proposition 1. *Optimal timing of payments.*

When only the low type's IC is binding

The high type gets no rent. There is no restriction on when to reward the low 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 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.[33] 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 $y_{T^L}^L \geq 0$.

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 $\left(\frac{P_{T^L}^H}{P_{T^L}^L}\right)$ 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:[34]

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

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

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

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

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.

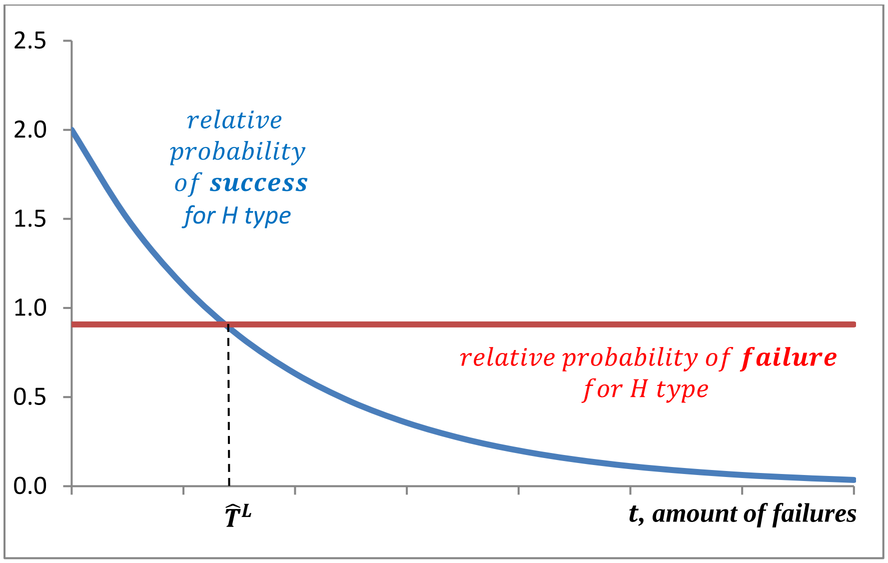

Figure 3. Relative probability of success/failure with $\lambda^H = 0.4$, $\lambda^L = 0.2$, $\beta_0 = 0.5$.

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

Finally, when the high type gets positive rent, we show in Appendix A, that the principal will reward him 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.

2.2.4. 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.[35] With production occurring after failure, the asymmetry of information generated during experimentation leads to information rent. We explain this next in details.

There are two main reasons the principal may ask the agent to over experiment. First, because λ enters directly 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 both increase the cost and reduce the benefit of lying for the high type through Δc_{T^L+1} and 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 in T^H.

Proposition 2: *For any λ^L, there exists $\overline{\lambda}^H(\lambda^L)$ and $\overline{\overline{\lambda}}^H(\lambda^L)$, such that $\lambda^L < \overline{\lambda}^H < \overline{\overline{\lambda}}^H(\lambda^L) \leq 1$, and there is over-experimentation in T^H, i.e., $T^H_{SB} > T^H_{FB}$, when $\overline{\lambda}^H(\lambda^L) < \lambda^H < \overline{\overline{\lambda}}^H(\lambda^L)$.*

Proof: See Appendix A.

To understand the intuition behind these sufficient conditions, it is convenient to focus on the role of Δc_t and separate the cases where the optimal termination date is on the increasing or decreasing part of the Δc_t function. Consider the case where, at the optimum, T^H is in the decreasing part of Δc_t, i.e., when $t_\Delta < T^H$.[36] Increasing T^H decreases Δc_{T^H+1} which is proportional to the benefit of lying for both the low type and the high type (positive part of the

[35] 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, $q^\theta(c^\theta_{T^\theta+1}) \equiv 0$.

[36] This happens when λ^H is large enough relative to λ^L since t_Δ becomes small ($\overline{\lambda}^H(\lambda^L) < \lambda^H$). The reason why λ^H cannot be too high ($\lambda^H < \overline{\overline{\lambda}}^H(\lambda^L)$) is due to the presence of the output in determining whether the rent is increasing or decreasing (see the proof for details).

gamble). Therefore, the principal may benefit from asking the high type to over experiment: $T^H \geq T^H_{FB}$. For the case of over-experimentation in T^L, we need to focus on the increasing part of Δc_t, i.e., when $T^L < t_\Delta$. Increasing T^L increases Δc_{T^L+1} which is proportional to the cost of lying for the high type (negative part of the gamble). Therefore, the principal may benefit from asking the low type to over experiment: $T^L \geq T^L_{FB}$.

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

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

$$V'\left(q_t^{\theta}(\underline{c})\right) = \underline{c} \ \text{ for } t \le T^{\theta}.$$

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

$$q_{SB}^{H}\left(c_{T^H+1}^{H}\right) < q_{FB}^{H}\left(c_{T^H+1}^{H}\right).$$

After failure, the low type overproduces:

$$q_{SB}^{L}\left(c_{T^L+1}^{L}\right) \ge q_{FB}^{L}\left(c_{T^L+1}^{L}\right).$$

Proof: See Appendix A.

3. Extensions

3.1. 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 can remain 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 privately known skill 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. This introduces a moral hazard rent in every period. This moral hazard rent may be so high that both (IC) constraints are slack. However, when adverse selection is a concern, we show that our key insights regarding the screening properties of the optimal contract remain intact. It is still optimal to provide exaggerated rewards for the high type at the beginning and for the low type at the end of experimentation, possibly rewarding failure. Furthermore, the agent's adverse selection rent is still determined by the difference in expected cost, which remains non-monotonic in time. We again find that over-experimentation and over-production can occur.

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 announce or reveal it at any period $t \ge j$.

Thus, we assume that success generates hard information that can be presented to the principal when desired, but it cannot be fabricated.

The (EMH^{θ}) constraint makes it unprofitable for the agent to hide success in the last period. The (EMP_t^{θ}) 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.

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

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

If the agent succeeds but hides it, the principal's expected cost is given by $c_{T^{\theta}+1}^{\theta}$ while the agent knows the true cost is $\underline{c}$ at the production stage. In addition to the existing (LL) and (IC) constraints, the optimal scheme must now satisfy the above ex post moral hazard constraints.

We formally show in the Supplementary Appendix C that both $(IC^{H,L})$ and $(IC^{L,H})$ may be slack, and either or both may be binding.[37] Since the ex post moral hazard constraints imply that both types will receive rent, these rents may be sufficient to satisfy the (IC) constraints.

A key objective for this subsection is to explore the impact of moral hazard on the optimality of delaying rewards or paying after failure. So, we first focus on the timing of payments among the screening instruments. When the principal rewards failure with $x^{\theta} > 0$, the (EMH^{θ}) constraint forces her to also reward success in the last period ($y_{T^{\theta}}^{\theta} > 0$ because of (EMH^{θ})) and in all previous periods ($y_t^{\theta} > 0$ because of (EMP_t^{θ})). An increase of \$1 in x^{θ} causes an increase of \$1 in $y_{T^{\theta}}^{\theta}$, which in turn causes an increase in all the previous y_t^{θ} according to the discount factor.

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 two ex post moral hazard constraints 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.

[37] Unlike the case when success is public, the $(IC^{L,H})$ may not always be binding.

Up to now we have focused on whether the ex post moral hazard constraints affect the timing of payments. Now we consider how those constraints interact with the other screening instruments.

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 $q^{\theta}\left(c^{\theta}_{T^{\theta}+1}\right)$ tightens the moral hazard constraint. A familiar tradeoff emerges: by asking the agent to over produce after failure, the principal mitigates the adverse selection rent but increases the moral hazard rent. In Supplementary Appendix C, we show that over-production remains optimal.

3.2. Learning bad news

In this section, we show that our main results survive if the object of experimentation is to seek bad news, where success in an experiment means discovery of high cost $c = \overline{c}$. For instance, stage 1 of a drug trial looks for bad news by testing the safety of the drug. Following the literature on experimentation we call the event of observing $c = \overline{c}$ by the agent "success" although this is bad news for the principal. If success is not achieved in a particular period, the principal and agent both become more optimistic (instead of pessimistic in a good news model). Also, as time goes by without learning that the cost is high, the expected cost becomes lower. In addition, the difference in the expected cost is now negative, $\Delta c_t = c^H_t - c^L_t < 0$ since the high type is relatively more optimistic after the same amount of failures. 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 good news. The optimization problem mirrors the case for good news and we find results similar to those in Propositions 1, 2, and 3. We present these results formally in Supplementary Appendix D. The parallel between good news and bad news is remarkable but not difficult to explain. In both cases, the agent is looking for news. The types determine how good the agent is at obtaining this news. The contract gives incentives for each type of agent to reveal his type, not the actual news.

Finally, unlike in the case of good news, if the agent is rewarded for success, he has no incentive to hide success in the last period as he will be under compensated in the production phase.

4. 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 good and bad-news models.

References

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

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

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

Bergemann D. and J. Välimäki (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.

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.

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

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. (1998). 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.

Rodivilov A. (2018). Monitoring Innovation. *Working Paper*.

Rosenberg, D., Salomon, A. and Vieille, N. (2013). On games of strategic experimentation. *Games and Economic Behavior* 82, 31-51.
Singh S. (2002). Contracting Out Solutions: Political Economy of Contract Farming in the Indian Punjab. *World Development* 30(9), 1621-1638.
Sadler E. (2017). Dead Ends. *Working Paper*, April 2017.

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

First best: Characterizing $\hat{\lambda}$.

Claim 1. There exists $\hat{\lambda} \in (0,1)$, such that $\frac{dT_{FB}^{\theta}}{d\lambda^{\theta}} > 0$ for $\lambda^{\theta} < \hat{\lambda}$ and $\frac{dT_{FB}^{\theta}}{d\lambda^{\theta}} \leq 0$ for $\lambda^{\theta} \geq \hat{\lambda}$.

Proof: The first-best termination date t is such that

$$\delta\beta_t^{\theta}\lambda^{\theta}\left[V\left(q_t^{\theta}(\underline{c})\right) - \underline{c}q_{t^{\theta}}^{\theta}(\underline{c})\right] + \delta\left(1 - \beta_t^{\theta}\lambda^{\theta}\right)\left[V\left(q^{\theta}\left(c_{t+1}^{\theta}\right)\right) - c_{t+1}^{\theta}q^{\theta}\left(c_{t+1}^{\theta}\right)\right]$$
$$= \gamma + \left[V\left(q^{\theta}\left(c_t^{\theta}\right)\right) - c_t^{\theta}q^{\theta}\left(c_t^{\theta}\right)\right].$$

Rewriting it next we have

$$\delta\beta_t\lambda\left(\left[V\left(q_t^{\theta}(\underline{c})\right) - \underline{c}q_t^{\theta}(\underline{c})\right] - \left[V\left(q^{\theta}\left(c_{t+1}^{\theta}\right)\right) - c_{t+1}^{\theta}q^{\theta}\left(c_{t+1}^{\theta}\right)\right]\right)$$
$$+\left(\delta\left[V\left(q^{\theta}\left(c_{t+1}^{\theta}\right)\right) - c_{t+1}^{\theta}q^{\theta}\left(c_{t+1}^{\theta}\right)\right] - \left[V\left(q^{\theta}\left(c_t^{\theta}\right)\right) - c_t^{\theta}q^{\theta}\left(c_t^{\theta}\right)\right]\right) = \gamma,$$

which implicitly determines t as a function of λ, $t(\lambda)$. Using the Implicit Function Theorem

$$\frac{dt}{d\lambda} = -\frac{\frac{\partial\Phi(\lambda,t)}{\partial\lambda}}{\frac{\partial\Phi(\lambda,t)}{\partial t}}, \text{ where}$$

$$\Phi(\lambda,t) = \delta\beta_t\lambda\left(\left[V\left(q_t^{\theta}(\underline{c})\right) - \underline{c}q_t^{\theta}(\underline{c})\right] - \left[V\left(q^{\theta}\left(c_{t+1}^{\theta}\right)\right) - c_{t+1}^{\theta}q^{\theta}\left(c_{t+1}^{\theta}\right)\right]\right) +$$
$$+\left(\delta\left[V\left(q^{\theta}\left(c_{t+1}^{\theta}\right)\right) - c_{t+1}^{\theta}q^{\theta}\left(c_{t+1}^{\theta}\right)\right] - \left[V\left(q^{\theta}\left(c_t^{\theta}\right)\right) - c_t^{\theta}q^{\theta}\left(c_t^{\theta}\right)\right]\right) - \gamma.$$

We now determine the sign of both $\frac{\partial\Phi(\lambda,t)}{\partial t}$ and $\frac{\partial\Phi(\lambda,t)}{\partial\lambda}$.

Since $\frac{\partial\beta_t}{\partial t} < 0$, $\frac{\partial\left(\left[V\left(q_t^{\theta}(\underline{c})\right) - \underline{c}q_t^{\theta}(\underline{c})\right] - \left[V\left(q^{\theta}\left(c_{t+1}^{\theta}\right)\right) - c_{t+1}^{\theta}q^{\theta}\left(c_{t+1}^{\theta}\right)\right]\right)}{\partial t} < 0$, and $\frac{\partial\left(\delta\left[V\left(q^{\theta}\left(c_{t+1}^{\theta}\right)\right) - c_{t+1}^{\theta}q^{\theta}\left(c_{t+1}^{\theta}\right)\right] - \left[V\left(q^{\theta}\left(c_t^{\theta}\right)\right) - c_t^{\theta}q^{\theta}\left(c_t^{\theta}\right)\right]\right)}{\partial t} < 0$, we have $\frac{\partial\Phi(\lambda,t)}{\partial t} < 0$. Therefore, the sigh of $\frac{dt}{d\lambda}$ is the same as the sign of $\frac{\partial\Phi(\lambda,t)}{\partial\lambda}$, which we determine next.

Since $\frac{\partial\left(\left[V\left(q_t^\theta(\underline{c})\right)-\underline{c}q_t^\theta(\underline{c})\right]-\left[V\left(q^\theta(c_{t+1}^\theta)\right)-c_{t+1}^\theta q^\theta(c_{t+1}^\theta)\right]\right)}{\partial t} < 0$ and $\frac{\partial\left(\delta\left[V\left(q^\theta(c_{t+1}^\theta)\right)-c_{t+1}^\theta q^\theta(c_{t+1}^\theta)\right]-\left[V\left(q^\theta(c_t^\theta)\right)-c_t^\theta q^\theta(c_t^\theta)\right]\right)}{\partial t} < 0$, for $\frac{dt}{d\lambda}$ to be positive it is necessary that $\frac{\partial(\beta_t\lambda)}{\partial t} > 0$.

Next, $\frac{\partial(\beta_t\lambda)}{\partial t} = \frac{\partial\left(\frac{\lambda\beta_0(1-\lambda)^{t-1}}{\beta_0(1-\lambda)^{t-1}+(1-\beta_0)}\right)}{\partial\lambda} =$

$$\frac{\beta_0\left((1-\lambda)^{t-1}+\lambda(1-\lambda)^{t-2}(t-1)(-1)\right)\left(\beta_0(1-\lambda)^{t-1}+(1-\beta_0)\right)-\beta_0\lambda(1-\lambda)^{t-1}\left(\beta_0(1-\lambda)^{t-1}+(1-\beta_0)\right)}{\left(\beta_0(1-\lambda)^{t-1}+(1-\beta_0)\right)^2} =$$

$$= \frac{\beta_0(1-\lambda)^{t-1}\left[1-\beta_0+\beta_0(1-\lambda)^{t-1}-\frac{(1-\beta_0)\lambda(t-1)}{1-\lambda}\right]}{\left(\beta_0(1-\lambda)^{t-1}+(1-\beta_0)\right)^2}.$$

Therefore, for $\frac{dt}{d\lambda} < 0$ it is necessary that $1-\beta_0+\beta_0(1-\lambda)^{t-1}-\frac{(1-\beta_0)\lambda(t-1)}{1-\lambda} < 0$ or, equivalently, $(1-\beta_0)(1-\lambda t)+\beta_0(1-\lambda)^t < 0$. Since $\frac{d\left[(1-\beta_0)(1-\lambda t)+\beta_0(1-\lambda)^t\right]}{dt} < 0$ for any λ it is sufficient to find $\hat{\lambda}$ such that $(1-\beta_0)(1-2\lambda)+\beta_0(1-\lambda)^2 < 0$ for any $\lambda > \hat{\lambda}$. Since $(1-\beta_0)(1-2\lambda)+\beta_0(1-\lambda)^2 =$
$\beta_0\left(\lambda-\frac{1-\sqrt{1-\beta_0}}{\beta_0}\right)\left(\lambda-\frac{1+\sqrt{1-\beta_0}}{\beta_0}\right)$, we define $\hat{\lambda} = \frac{1-\sqrt{1-\beta_0}}{\beta_0}$.[38] *Q.E.D.*

The Principal's Maximization Problem and Claim 2

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 $\left\{q_t^H(\underline{c})\right\}_{t=1}^{T^H}$, $q^H\left(c_{T^H}^H\right)$, $\left\{q_t^L(\underline{c})\right\}_{t=1}^{T^L}$ and $q^L\left(c_{T^L}^L\right)$ (Proposition 3).

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

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

[38] Note that $\frac{1-\sqrt{1-\beta_0}}{\beta_0}$ is well defined and $0 < \frac{1-\sqrt{1-\beta_0}}{\beta_0} < 1$ for $\beta_0 < 1$.

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:

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

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

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

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

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

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

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

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

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

We begin to solve the problem by first proving the following claim.

Claim 2: The constraint $(IC^{L,H})$ is binding and the low type obtains a strictly positive rent.

Proof: If the $(IC^{L,H})$ constraint was not binding, it would be possible to decrease the payment to the low type until (LLS_t^L) and (LLF_t^L) are binding, but that would violate $(IC^{L,H})$ since $\Delta c_{T^H+1} q^H\left(c_{T^H+1}^H\right) > 0$. *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 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 trace 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,

$\beta_0(1-\lambda^\theta)^{t-1}\lambda^\theta$ is the probability that the agent of type θ succeeds at period $t \le T^L$ of the experimentation stage and $\frac{\beta_0(1-\lambda^H)^{t-1}\lambda^H}{\beta_0(1-\lambda^L)^{t-1}\lambda^L} = \frac{(1-\lambda^H)^{t-1}\lambda^H}{(1-\lambda^L)^{t-1}\lambda^L}$ 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

$$\frac{1-\beta_0+\beta_0(1-\lambda^H)^{T^L}}{1-\beta_0+\beta_0(1-\lambda^L)^{T^L}} > \frac{(1-\lambda^H)^{t-1}\lambda^H}{(1-\lambda^L)^{t-1}\lambda^L} \text{ for } 1 \le t \le T^L \text{ or, equivalently,}$$

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

where $\frac{1-\beta_0+\beta_0(1-\lambda^\theta)^{T^L}}{(1-\lambda^\theta)^{t-1}\lambda^\theta}$ can be interpreted as a likelihood ratio.

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 $f_2(\hat{T}^L, T^L) = 0$ defined as

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

where uniqueness follows from $\frac{(1-\lambda^H)^{t-1}\lambda^H}{(1-\lambda^L)^{t-1}\lambda^L}$ being strictly decreasing in t and $\frac{\lambda^H}{\lambda^L} > 1 > \frac{P^H_{T^L}}{P^L_{T^L}}$.[39] In 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 $(IC^{H,L})$ constraint is binding or not; we explore each case separately in what follows.

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

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

[39] To explain, $f_2(t, T^L) = 0$ if and only if $\frac{1-\beta_0+\beta_0(1-\lambda^H)^{T^L}}{1-\beta_0+\beta_0(1-\lambda^L)^{T^L}} = \frac{(1-\lambda^H)^{t-1}\lambda^H}{(1-\lambda^L)^{t-1}\lambda^L}$. 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 $\frac{1-\beta_0+\beta_0(1-\lambda^H)^{T^L}}{1-\beta_0+\beta_0(1-\lambda^L)^{T^L}} < 1$ and $\frac{\lambda^H}{\lambda^L} > 1$ the uniqueness immediately follows. So $\hat{T}^L$ satisfies $\frac{P^H_{T^L}}{P^L_{T^L}} = \frac{(1-\lambda^H)^{\hat{T}^L-1}\lambda^H}{(1-\lambda^L)^{\hat{T}^L-1}\lambda^L}$.

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

(LLS^L_t) $y^L_t \geq 0$ for $t \leq T^L$,

and (LLF_{T^L}) $\delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q^H(c^H_{T^H+1}) - \beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1}\lambda^L y^L_t \geq 0$.

When the $(IC^{H,L})$ constraint is not binding, the claim below shows that there are no restrictions in choosing $\{y^L_t\}_{t=1}^{T^L}$ except those imposed by the $(IC^{L,H})$ constraint. In other words, the principal can choose any combinations of nonnegative payments to the low type $\left(x_L, \{y^L_t\}_{t=1}^{T^L}\right)$ such that $\beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1}\lambda^L y^L_t + \delta^{T^L} P^L_{T^L} x^L = \delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q^H(c^H_{T^H+1})$. Labeling by $\{\alpha^L_t\}_{t=1}^{T^L}, \alpha^L$ the Lagrange multipliers of the constraints associated with (LLS^L_t) for $t \leq T^L$, and (LLF_{T^L}) respectively, we have the following claim.

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

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

$\frac{\partial \mathcal{L}}{\partial y^L_t} = \alpha^L_t - \alpha^L \beta_0 \delta^t (1-\lambda^L)^{t-1}\lambda^L = 0$ for $1 \leq t \leq T^L$;

$\frac{\partial \mathcal{L}}{\partial \alpha^L_t} = y^L_t \geq 0$; $\alpha^L_t \geq 0$; $\alpha^L_t y^L_t = 0$ for $1 \leq t \leq T^L$.

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

$$\delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q^H(c^H_{T^H+1}) - \beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1}\lambda^L y^L_t = 0,$$

and there must exist $y^L_s > 0$ for some $1 \leq s \leq T^L$. Then, we have $\alpha^L_s = 0$, which leads to a contradiction since $\frac{\partial \mathcal{L}}{\partial y^L_t} = 0$ cannot be satisfied unless $\alpha^L = 0$.

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

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

We will now show that when the $(IC^{H,L})$ becomes binding, there are restrictions on the payment structure to the low type. Denoting by $\psi = P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H}$, we can re-write the incentive compatibility constraints as:

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

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

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

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

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

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 $\frac{P_{T^H}^H}{P_{T^H}^L} \neq \frac{P_{T^L}^H}{P_{T^L}^L}$ (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 may be rewarded only for the very first success.

Finally, we analyze the case when $\frac{P_{T^H}^H}{P_{T^H}^L} = \frac{P_{T^L}^H}{P_{T^L}^L}$ (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: $\psi = P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L \neq 0.$

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

$$Max\ E_\theta \left\{ \begin{array}{c} \Omega^\theta(\varpi^\theta) - \delta^{T^\theta} P_{T^\theta}^\theta x^\theta \left(\{y_t^H\}_{t=1}^{T^H}, \{y_t^L\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c_{T^H+1}^H\right), q^L\left(c_{T^L+1}^L\right)\right) \\ -\beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta y_t^\theta \end{array} \right\} \text{ subject to}$$

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

$\left(LLF_{T^\theta}\right)$ $x^\theta \left(\{y_t^H\}_{t=1}^{T^H}, \{y_t^L\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c_{T^H+1}^H\right), q^L\left(c_{T^L+1}^L\right)\right) \geq 0$ for $\theta = H, L$.

Labeling $\{\alpha_t^H\}_{t=1}^{T^H}$, $\{\alpha_t^L\}_{t=1}^{T^L}$, ξ^H and ξ^L as the Lagrange multipliers of the constraints associated with (LLS_t^H), (LLS_t^L), (LLF_{T^H}) and (LLF_{T^L}) respectively, the Lagrangian is:

$$\mathcal{L} = E_\theta \left\{ \Omega^\theta(\varpi^\theta) - \beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta y_t^\theta - \right.$$
$$\left. \delta^{T^\theta} P_{T^\theta}^\theta x^\theta \left(\{y_t^H\}_{t=1}^{T^H}, \{y_t^L\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c_{T^H+1}^H\right), q^L\left(c_{T^L+1}^L\right) \right) \right\}$$
$$+ \sum_{t=1}^{T^H} \alpha_t^H y_t^H + \sum_{t=1}^{T^L} \alpha_t^L y_t^L + \xi^H x^H \left(\{y_t^H\}_{t=1}^{T^H}, \{y_t^L\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c_{T^H+1}^H\right), q^L\left(c_{T^L+1}^L\right) \right)$$
$$+ \xi^L x^L \left(\{y_t^H\}_{t=1}^{T^H}, \{y_t^L\}_{t=1}^{T^L}, T^H, T^L, q^H\left(c_{T^H+1}^H\right), q^L\left(c_{T^L+1}^L\right) \right).$$

The Inada conditions give us interior solutions for $q_t^H(\underline{c})$, $q^H\left(c_{T^H+1}^H\right)$, $q_t^L(\underline{c})$ and $q^L\left(c_{T^L+1}^L\right)$. We also assumed that $T^L > 0$ and $T^H > 0$. The Kuhn-Tucker conditions with respect to y_t^H and y_t^L are:

$$\frac{\partial \mathcal{L}}{\partial y_t^H} = -\upsilon \left\{ \beta_0 \delta^t (1-\lambda^H)^{t-1} \lambda^H + \delta^{T^H} P_{T^H}^H \frac{\beta_0 \delta^t \left[P_{T^L}^H (1-\lambda^L)^{t-1} \lambda^L - P_{T^L}^L (1-\lambda^H)^{t-1} \lambda^H \right]}{\delta^{T^H}\left(P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L\right)} \right\}$$
$$-(1-\upsilon)\delta^{T^L} P_{T^L}^L \frac{\beta_0 \delta^t \left[P_{T^H}^H (1-\lambda^L)^{t-1} \lambda^L - P_{T^H}^L (1-\lambda^H)^{t-1} \lambda^H \right]}{\delta^{T^L}\left(P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L\right)} + \alpha_t^H$$
$$+\xi^H \frac{\beta_0 \delta^t \left[P_{T^L}^H (1-\lambda^L)^{t-1} \lambda^L - P_{T^L}^L (1-\lambda^H)^{t-1} \lambda^H \right]}{\delta^{T^H}\left(P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L\right)} + \xi^L \frac{\beta_0 \delta^t \left[P_{T^H}^H (1-\lambda^L)^{t-1} \lambda^L - P_{T^H}^L (1-\lambda^H)^{t-1} \lambda^H \right]}{\delta^{T^L}\left(P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L\right)};$$

$$\frac{\partial \mathcal{L}}{\partial y_t^L} = -(1-\upsilon) \left\{ \beta_0 \delta^t (1-\lambda^L)^{t-1} \lambda^L + \delta^{T^L} P_{T^L}^L \frac{\beta_0 \delta^t \left[P_{T^H}^L (1-\lambda^H)^{t-1} \lambda^H - P_{T^H}^H (1-\lambda^L)^{t-1} \lambda^L \right]}{\delta^{T^L}\left(P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L\right)} \right\}$$
$$-\upsilon \delta^{T^H} P_{T^H}^H \frac{\beta_0 \delta^t \left[P_{T^L}^L (1-\lambda^H)^{t-1} \lambda^H - P_{T^L}^H (1-\lambda^L)^{t-1} \lambda^L \right]}{\delta^{T^H}\left(P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L\right)} + \alpha_t^L$$
$$+\xi^H \frac{\beta_0 \delta^t \left[P_{T^L}^L (1-\lambda^H)^{t-1} \lambda^H - P_{T^L}^H (1-\lambda^L)^{t-1} \lambda^L \right]}{\delta^{T^H}\left(P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L\right)} + \xi^L \frac{\beta_0 \delta^t \left[P_{T^H}^L (1-\lambda^H)^{t-1} \lambda^H - P_{T^H}^H (1-\lambda^L)^{t-1} \lambda^L \right]}{\delta^{T^L}\left(P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L\right)}.$$

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

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

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

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 $1 \le j \le T^L$ such that $y_j^L > 0$ and *at most* one time period $1 \le s \le T^H$ such that $y_s^H > 0$.

Proof: Assume to the contrary that there are two distinct periods $1 \le k, m \le T^L$ such that $k \ne m$ 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:

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

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

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

For later use, we prove the following claim:

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

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 $-\psi\left((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H\right) + \frac{\alpha_t^L \psi}{\beta_0 \delta^t} = 0$ 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

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

which implies that $-\psi\left((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L+\upsilon(1-\lambda^H)^{t-1}\lambda^H\right)+\frac{\alpha_t^H\psi}{\beta_0\delta^t}=0$ 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 $(LLF_{T^H}^H)$ and $(LLF_{T^L}^L)$ 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 $\alpha_t^L>0$ for $t<\hat{T}^L$ and $\psi>0$, which leads to a contradiction[40].

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

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

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

If $\alpha_s^H=0$ for some $1\le s\le T^H$ then $P_{T^H}^H f_1(s,T^H)\left[\upsilon P_{T^L}^H+(1-\upsilon)P_{T^L}^L-\frac{\xi^L}{\delta^{T^L}}\right]=0$, which implies that $\frac{\xi^L}{\delta^{T^L}}=\upsilon P_{T^L}^H+(1-\upsilon)P_{T^L}^L$[41]. Since we rule out this possibility it immediately follows that all $\alpha_t^H>0$ for all $1\le t\le T^H$ which implies that $y_t^H=0$ for $1\le t\le T^H$.

[40] 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 $\psi\ne0$; we study $\psi=0$ in details later in Case B.2.

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

Finally, from $P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L - \frac{\xi^L}{\delta^{T^L}}\right] = -\frac{\alpha_t^H \psi}{\beta_0 \delta^t}$ we conclude that $T^H \leq \hat{T}^H$ and there can be one of two sub-cases:[42] (b.1) $\psi > 0$ and $\frac{\xi^L}{\delta^{T^L}} > \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$, or (b.2) $\psi < 0$ and $\frac{\xi^L}{\delta^{T^L}} < \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$. We consider each sub-case next.

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

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

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

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

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

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

$$\left(P_{T^L}^H(1-\lambda^L)^{t-1}\lambda^L - P_{T^L}^L(1-\lambda^H)^{t-1}\lambda^H\right)\left(P_{T^H}^L(1-\lambda^H)^{j-1}\lambda^H - P_{T^H}^H(1-\lambda^L)^{j-1}\lambda^L\right)$$
$$< \left(P_{T^L}^H(1-\lambda^L)^{j-1}\lambda^L - P_{T^L}^L(1-\lambda^H)^{j-1}\lambda^H\right)\left(P_{T^H}^L(1-\lambda^H)^{t-1}\lambda^H - P_{T^H}^H(1-\lambda^L)^{t-1}\lambda^L\right).$$
$$\psi\left[(1-\lambda^H)^{t-1}(1-\lambda^L)^{j-1} - (1-\lambda^L)^{t-1}(1-\lambda^H)^{j-1}\right] < 0 \text{ for } 1 \leq t \neq j \leq T^L.$$

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

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

Case (b.2): $T^H \leq \hat{T}^H$, $\psi < 0$, $\frac{\xi^L}{\delta^{T^L}} < \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$, $\xi^H = 0$, $\alpha_t^H > 0$ for $1 \leq t \leq T^H$.

As in the previous case, from Lemma 3 it follows that there exists only one time period $1 \leq s \leq T^L$ such that $y_s^L > 0$ ($\alpha_s^L = 0$). This implies that $P_{T^L}^L f_2(s,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(s,T^H) = 0$ and $P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(t,T^H) = -\frac{\alpha_t^L \psi}{\beta_0 \delta^t} > 0$ for $1 \leq t \neq s \leq T^L$. 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 $f_2(t,T^L) f_1(s,T^H) > f_1(t,T^H) f_2(s,T^L)$

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

$\psi\left[1 - \left(\frac{1-\lambda^L}{1-\lambda^H}\right)^{s-t}\right] < 0$, which implies that $t > s$ for all1 $\leq t \neq s \leq T^L$ or, equivalently, $s = 1$.

[42] 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 $t \leq T^H$.

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

For $t > \hat{T}^H$ it follows that $P^L_{T^L} f_2(t, T^L)[\upsilon P^H_{T^H} + (1-\upsilon) P^L_{T^H}] + \frac{\xi^L}{\delta^{T^L}} P^H_{T^H} f_1(t, T^H)$ $> -\psi((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H) > 0$, 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

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

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

Note that Case B.2 is possible only if $\delta^{T^H} P^L_{T^H}\Delta c_{T^H+1} q^H(c^H_{T^H+1}) - \delta^{T^L} P^L_{T^L}\Delta c_{T^L+1} q^L(c^L_{T^L+1}) > 0$[43]. This fact together with $x^H \geq 0$ implies that $\psi > 0$. Since $f_1(1, T^H) > 0$, $x^L = 0$ is possible only if $\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q^H(c^H_{T^H+1}) - \delta^{T^L} P^H_{T^L}\Delta c_{T^L+1} q^L(c^L_{T^L+1}) < 0$. However, $\delta^{T^H} P^L_{T^H}\Delta c_{T^H+1} q^H(c^H_{T^H+1}) > \delta^{T^L} P^L_{T^L}\Delta c_{T^L+1} q^L(c^L_{T^L+1})$ implies that $\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q^H(c^H_{T^H+1}) > \delta^{T^L} \frac{P^H_{T^H}}{P^L_{T^H}} P^L_{T^L}\Delta c_{T^L+1} q^L(c^L_{T^L+1})$. 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 $\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q^H(c^H_{T^H+1}) > \delta^{T^L} P^H_{T^L}\Delta c_{T^L+1} q^L(c^L_{T^L+1})$, 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: $T^L \leq \hat{T}^L$.

Lemma 4. $\xi^L = 0 \Rightarrow T^L \leq \hat{T}^L$, $\alpha^L_t > 0$ for $t \leq T^L$ (it is optimal to set $x^L > 0$, $y^L_t = 0$ for $t \leq T^L$) and $\alpha^H_t > 0$ for all $t > 1$ and $\alpha^H_1 = 0$ (it is optimal to set $x^H = 0$, $y^H_t = 0$ for all $t > 1$ and $y^H_1 > 0$).

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

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

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

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

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 $\frac{\xi^H}{\delta^{T^H}} = \upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L$[44]. 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:[45] (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 $P_{T^H}^H f_1(t, T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\right] + \frac{\xi^H}{\delta^{T^H}} P_{T^L}^L f_2(t, T^L)$ $< -\psi\left((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H\right) < 0$, which implies that $y_j^H > 0$ 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

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

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

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

[45] 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 $t \leq T^L$.

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

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 $1 \le t \ne j \le T^H$ 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 $1 \le t \ne j \le T^H$ or, equivalently, $j = 1$.

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

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

which implies that $y^H_j > 0$ 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:

$$y^H_1 = \frac{P^H_{T^L}\left(\delta^{T^L} P^L_{T^L} \Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right) - \delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right)\right)}{\beta_0 \delta P^L_{T^L} f_2(1,T^L)} \ge 0;$$

$$x^L = \frac{\delta^{T^L}\lambda^L P^H_{T^L} \Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right) - \delta^{T^H}\lambda^H P^L_{T^H} \Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right)}{\delta^{T^L} P^L_{T^L} f_2(1,T^L)} > 0.$$

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: $\xi^L > 0 \Rightarrow T^L > \hat{T}^L$, $\alpha^L_t > 0$ for $t < T^L$, $\alpha^L_{T^L} = 0$ and $\alpha^H_t > 0$ for $t > 1$, $\alpha^H_1 = 0$ (it is optimal to set $x^L = 0$, $y^L_t = 0$ for $t < T^L$, $y^L_{T^L} > 0$ and $x^H = 0$, $y^H_t = 0$ for $t > 1$, $y^H_1 > 0$)

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

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

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

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 $y^L_j > 0$ ($\alpha^L_j = 0$) it follows that

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

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

Alternatively, $\frac{\xi^L}{\delta^{T^L}}\left[f_1(t,T^H) - \frac{f_2(t,T^L) f_1(j,T^H)}{f_2(j,T^L)}\right] = -\frac{\alpha^L_t \psi}{\beta_0 \delta^t P^H_{T^H}}$ 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

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

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

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

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 $s=T^H>\hat{T}^H$. *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

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

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

Since Case B.2 is possible only if $\delta^{T^H}P^L_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right) - \delta^{T^L}P^L_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right) > 0$[46], 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:

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

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

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

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

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

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

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

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

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

Since $P^H_{TL} < P^L_{TL}$ and $|\ln(1-\lambda^H)| > |\ln(1-\lambda^L)|$, $P^H_{TL}\ln(1-\lambda^L) - P^L_{TL}\ln(1-\lambda^H) > 0$ 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, $\frac{P^H_{TH}}{P^L_{TH}} = \frac{P^H_{TL}}{P^L_{TL}}$.

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

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

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

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

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

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

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. $P^H_{TH}P^L_{TL} - P^H_{TL}P^L_{TH} = 0 \Rightarrow T^L > \hat{T}^L$, $\xi^H > 0$, $\xi^L > 0$, $\alpha^H_t > 0$ 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 $\{\alpha_t^H\}_{t=1}^{T^H}$, $\{\alpha_t^L\}_{t=1}^{T^L}$, α^H, α^L, ξ^H and ξ^L as the Lagrange multipliers of the constraints associated with (LLS_t^H), (LLS_t^L), $(IC^{H,L})$, $(IC^{L,H})$, (LLF_{T^H}) and (LLF_{T^L}) respectively, we can rewrite the Kuhn-Tucker conditions as follows:

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

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

$\frac{\partial \mathcal{L}}{\partial y_t^H} = -\upsilon(1-\lambda^H)^{t-1}\lambda^H + \alpha^H P_{T^L}^L f_2(t,T^L) - \alpha^L P_{T^H}^H f_1(t,T^H) + \frac{\alpha_t^H}{\delta^t \beta_0} = 0$ for $1 \le t \le T^H$;

$\frac{\partial \mathcal{L}}{\partial y_t^L} = -(1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L - \alpha^H P_{T^L}^L f_2(t,T^L) + \alpha^L P_{T^H}^H f_1(t,T^H) + \frac{\alpha_t^L}{\delta^t \beta_0} = 0$ for $1 \le t \le T^L$.

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 $y_j^L > 0$ and *at most* one time period $1 \le s \le T^H$ such that $y_s^H > 0$.

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

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

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

Combining the two equations together, $\alpha^L P_{T^H}^H\big(f_1(k,T^H)f_2(m,T^L) - f_1(m,T^H)f_2(k,T^L)\big)$ $+\upsilon\lambda^H\Big((1-\lambda^H)^{k-1}f_2(m,T^L) - (1-\lambda^H)^{m-1}f_2(k,T^L)\Big) = 0$, which can be rewritten as follows[47]:

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

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

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:

---

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

$$\frac{P^H_{T^L}(1-\lambda^H)^{t-1}(1-\lambda^L)^{t-1}}{P^L_{T^L}}\left((1-\lambda^H)^{s-t}-(1-\lambda^L)^{s-t}\right)=-\frac{\alpha^H_t}{\delta^t\beta_0}f_2(s,T^L) \text{ for } 1\le t\ne s\le T^H.$$

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 $y^L_j>0$ it must be that either $j=1$ or $j=T^L>\hat{T}^L$. *Q.E.D.*

Finally, from $\frac{\partial\mathcal{L}}{\partial y^H_1}=-\upsilon\lambda^H+\alpha^H P^L_{T^L}f_2(1,T^L)-\alpha^L P^H_{T^H}f_1(1,T^H)=0$ when $y^H_1>0$ and $\frac{\partial\mathcal{L}}{\partial y^L_1}=-(1-\upsilon)\lambda^L-\alpha^H P^L_{T^L}f_2(1,T^L)+\alpha^L P^H_{T^H}f_1(1,T^H)=0$ when $y^L_1>0$ we have a contradiction. As a result, $y^H_1>0$ implies $y^L_{T^L}>0$ with $T^L>\hat{T}^L$. *Q.E.D.*

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

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

$$U^L_A=\delta^{T^H}P^L_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right),$$
$$U^H_A=0,$$

where the subscript A refers to case A. In case B, both $(IC^{H,L})$ and $(IC^{L,H})$ are binding, and the rents to the low and high types are

$$U^L_B=\frac{(1-\lambda^L)^{T^L-1}\left(\delta^{T^H}\lambda^H P^L_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right)-\delta^{T^L}\lambda^L P^H_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right)\right)}{\lambda^H\left((1-\lambda^L)^{T^L-1}-(1-\lambda^H)^{T^L-1}\right)}$$

$$U^H_B=\frac{\delta^{T^H}P^L_{T^H}\Delta c_{T^H+1}(1-\lambda^H)^{T^L-1}\lambda^H q^H\left(c^H_{T^H+1}\right)-\delta^{T^L}\Delta c_{T^L+1}(1-\lambda^L)^{T^L-1}\lambda^L q^L\left(c^L_{T^L+1}\right)}{\lambda^L\left((1-\lambda^L)^{T^L-1}-(1-\lambda^H)^{T^L-1}\right)},$$

where the subscript B refers to case B in the proof of proposition 1.

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 ($(IC^{H,L})$ is slack) or Case B (both $(IC^{L,H})$ and $(IC^{H,L})$ are binding.)

In Case A, the low type's rent U^L_A is not affected by T^L. Therefore, the F.O.C. with respect to T^L is identical to that under first best: $\frac{\partial E_\theta\,\Omega^\theta(\varpi^\theta)}{\partial T^L}=0$, or, equivalently, $T^L_{SB}=T^L_{FB}$ when $(IC^{H,L})$ is not binding. However, since U^L_A depends on T^H, there will be a distortion in the duration of the experimentation stage for the high type:

$$\frac{\partial\left(E_\theta\,\Omega^\theta(\varpi^\theta)-(1-\upsilon)U_A^L\right)}{\partial T^H}=0.$$

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

In Case B, the exact values of the rent to each type U_B^H and U_B^L depend on whether $T^L < \hat{T}^L$ (Lemma 4) or $T^L > \hat{T}^L$ (Lemma 5), but in each case $U_B^L > 0$ and $U_B^H \geq 0$. The FOC is given by

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

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. We can use similar steps to provide sufficient conditions for over-experimentation in T^L.

Proposition 2: *For any λ^L, there exists $\overline{\lambda}^H(\lambda^L)$ and $\overline{\overline{\lambda}}^H(\lambda^L)$, such that $\lambda^L < \overline{\lambda}^H < \overline{\overline{\lambda}}^H(\lambda^L) \leq 1$, and there is over-experimentation in T^H, i.e., $T_{SB}^H > T_{FB}^H$, when $\overline{\lambda}^H(\lambda^L) < \lambda^H < \overline{\overline{\lambda}}^H(\lambda^L)$.*

Proof of proposition 2:

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

Step 1. We show that $\boldsymbol{\frac{d\zeta(t)}{dt} < 0}$ if $\boldsymbol{\lambda^H}$ is high enough $\left(\boldsymbol{\overline{\lambda}^H(\lambda^L) < \lambda^H}\right)$.

Proof of step 1: Recalling that $P_T^\theta = 1-\beta_0+\beta_0\left(1-\lambda^\theta\right)^T$, and $\beta_t^\theta = \frac{\beta_0\left(1-\lambda^\theta\right)^{t-1}}{\beta_0\left(1-\lambda^\theta\right)^{t-1}+(1-\beta_0)}$, we can rewrite $\zeta(t)$: $\zeta(t) = \delta^t(1-\beta_0+\beta_0(1-\lambda^L)^t)\left(\frac{\beta_0\left(1-\lambda^L\right)^t}{\beta_0(1-\lambda^L)^t+(1-\beta_0)} - \frac{\beta_0\left(1-\lambda^H\right)^t}{\beta_0(1-\lambda^H)^t+(1-\beta_0)}\right)$

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

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

$$\frac{d\zeta(t)}{dt} =$$

$$= \delta^t \frac{\left((1-\lambda^L)^t ln(1-\lambda^L) - (1-\lambda^H)^t ln(1-\lambda^H)\right)P_t^H - \beta_0(1-\lambda^H)^t ln(1-\lambda^H)((1-\lambda^L)^t - (1-\lambda^H)^t)}{(P_t^H)^2 \frac{1}{\beta_0\left(1-\beta_0\right)}}$$

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

$$= \delta^t \frac{[ln(1-\lambda^L) + ln\,\delta](1-\lambda^L)^t P_t^H - (1-\lambda^H)^t (P_t^L ln(1-\lambda^H) + P_t^H ln\delta)}{(P_t^H)^2 \frac{1}{\beta_0(1-\beta_0)}}.$$

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

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

We prove next that $\phi(\lambda^H) < 0$ if λ^H is sufficiently large, i.e., there exists $\overline{\lambda}^H$ such that $\phi(\lambda^H) < 0$ if $\lambda^H > \overline{\lambda}^H$. Consider the derivative of $\phi(\lambda^H)$ with respect to λ^H:

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

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

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

$$(1-\lambda^H)^{t-1}(-\beta_0 t(1-\lambda^L)^t \ln[\delta(1-\lambda^L)] + P_t^L(t ln(1-\lambda^H) + 1) - ln\,\delta\, t(\beta_0(1-\lambda^H)^t + P_t^H))$$

Since $-\beta_0 t(1-\lambda^L)^t \ln[\delta(1-\lambda^L)] > 0$ and $-ln\,\delta\, t(\beta_0(1-\lambda^H)^t + P_t^H) > 0$, there exists a value of $\dot{\lambda}^H(\lambda^L)$ such that if $\lambda^H > (<)\dot{\lambda}^H(\lambda^L)$ then

$$-\beta_0 t(1-\lambda^L)^t \ln[\delta(1-\lambda^L)] + P_t^L(t ln(1-\lambda^H) + 1) - ln\,\delta\, t(\beta_0(1-\lambda^H)^t + P_t^H) > (<)0.$$

Therefore, the function $\phi(\lambda^H)$ is increasing in λ^H if $\lambda^H < \dot{\lambda}^H$ and decreasing in λ^H if $\lambda^H > \dot{\lambda}^H$. In addition, $\phi = 0$ if $\lambda^H = \lambda^L$ and $\lim_{\lambda^H \to 1} \phi(\lambda^H) < 0$. Since the function $\phi(\lambda^H)$ is continuous in λ^H, there exists $\overline{\lambda}^H < 1$, such that $\phi(\lambda^H) < 0$ if $\overline{\lambda}^H(\lambda^L) < \lambda^H$. We define a value $\overline{\lambda}^H$ such that the function is equal to zero:

$$\phi\left(\overline{\lambda}^H\right) \equiv 0.$$

As a result, the function $\zeta(t)$ is a decreasing function of t if $\overline{\lambda}^H(\lambda^L) < \lambda^H$.

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

Proof of Step 2:

If $(IC^{H,L})$ is not binding, the rent to the low type is $U^L = \delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right) =$. If both $(IC^{H,L})$ and $(IC^{L,H})$ are binding, using the function $\zeta(t) \equiv \delta^t P^L_t(\beta^L_{t+1} - \beta^H_{t+1})$, we can rewrite U^L and U^H as:

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

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

Note that $q^H\left(c^H_{T^H+1}\right)$ decreases proportionately to $\frac{P^L_{T^H}}{P^H_{T^H}} \Delta c_{T^H+1}$.[48] Therefore, if $\frac{d\zeta(t)}{dt} < 0$ and $\frac{P^L_{T^H}}{P^H_{T^H}} \Delta c_{T^H+1}$ is increasing in T^H, then both rents U^H and U^L are decreasing in T^H, and over-experimentation in T^H is optimal.

Note that:

$$\frac{P^L_{T^H}}{P^H_{T^H}} \Delta c_{T^H+1} = \beta_0(1-\beta_0)(\overline{c} - \underline{c}) \frac{\left((1-\lambda^L)^{T^H} - (1-\lambda^H)^{T^H}\right)}{\left[P^H_{T^H}\right]^2}.$$

We next prove that for any λ^L there exists $\overline{\overline{\lambda}}^H(\lambda^L)$ such that for $\lambda^H < \overline{\overline{\lambda}}^H(\lambda^L)$, $\frac{\left((1-\lambda^L)^t - (1-\lambda^H)^t\right)}{\left[P^H_t\right]^2}$ is an increasing function of t.

[48] In Proposition 2, we formally prove that $V'\left(q^H\left(c^H_{T^H+1}\right)\right) - c^H_{T^H+1} = \frac{(1-\nu) P^L_{T^H}}{\nu P^H_{T^H}} \Delta c_{T^H+1}$ if $(IC^{H,L})$ is not binding, and $V'\left(q^H\left(c^H_{T^H+1}\right)\right) - c^H_{T^H+1} = \frac{P^L_{T^H} E_\theta\left\{(1-\lambda^\theta)^{T^L-1}\lambda^\theta\right\}}{\upsilon P^H_{T^H} \lambda^L\left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)} \Delta c_{T^H+1}$ if $(IC^{H,L})$ is binding.

$$\frac{d\left[\frac{\left((1-\lambda^L)^t-(1-\lambda^H)^t\right)}{[P_t^H]^2}\right]}{dt}=$$

$$\frac{\left((1-\lambda^L)^t\ln(1-\lambda^L)-(1-\lambda^H)^t\ln(1-\lambda^H)\right)[P_t^H]^2}{[P_t^H]^4}-\frac{\left((1-\lambda^L)^t-(1-\lambda^H)^t\right)2P_t^H\beta_0(1-\lambda^H)^t\ln(1-\lambda^H)}{[P_t^H]^4}=$$

$$\frac{\left((1-\lambda^L)^t\ln(1-\lambda^L)-(1-\lambda^H)^t\ln(1-\lambda^H)\right)P_t^H-\left((1-\lambda^L)^t-(1-\lambda^H)^t\right)2\beta_0(1-\lambda^H)^t\ln(1-\lambda^H)}{[P_t^H]^3}.$$

Therefore, $\frac{\left((1-\lambda^L)^t-(1-\lambda^H)^t\right)}{[P_t^H]^2}$ is an increasing function of t if

$$((1-\lambda^L)^t\ln(1-\lambda^L)-(1-\lambda^H)^t\ln(1-\lambda^H))P_t^H>$$

$$2\beta_0(1-\lambda^H)^t((1-\lambda^L)^t-(1-\lambda^H)^t)\ln(1-\lambda^H),$$

$$P_t^H(1-\lambda^L)^t\ln(1-\lambda^L)-(1-\lambda^H)^t(P_t^H+2\beta_0(1-\lambda^L)^t-2\beta_0(1-\lambda^H)^t)\ln(1-\lambda^H)>0.$$

Given that $P_t^H+2\beta_0(1-\lambda^L)^t-2\beta_0(1-\lambda^H)^t=1-\beta_0-\beta_0(1-\lambda^H)^t+2\beta_0(1-\lambda^L)^t$, the inequality above can be rewritten as

$$P_t^H(1-\lambda^L)^t\ln(1-\lambda^L)-(1-\lambda^H)^t(1-\beta_0-\beta_0(1-\lambda^H)^t+2\beta_0(1-\lambda^L)^t)\ln(1-\lambda^H)>0.$$

We prove next that the function £ defined as

$$£=P_t^H(1-\lambda^L)^t\ln(1-\lambda^L)-(1-\lambda^H)^t(1-\beta_0-\beta_0(1-\lambda^H)^t+2\beta_0(1-\lambda^L)^t)\ln(1-\lambda^H),$$

is increasing for small values of λ^H, when $\lambda^H<\ddot{\lambda}^H(\lambda^L)$, and decreasing for high values of λ^H, when $\lambda^H>\ddot{\lambda}^H(\lambda^L)$.

$$\frac{d\,£}{d\,\lambda^H}=-\beta_0t(1-\lambda^H)^{t-1}(1-\lambda^L)^t\ln(1-\lambda^L)-$$

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

$$-(1-\lambda^H)^t\left[\frac{\left(1-\beta_0-\beta_0(1-\lambda^H)^t+2\beta_0(1-\lambda^L)^t\right)(-1)}{1-\lambda^H}+\ln(1-\lambda^H)\,\beta_0t(1-\lambda^H)^{t-1}\right]$$

$$=(1-\lambda^H)^{t-1}\begin{bmatrix}-\beta_0t(1-\lambda^L)^t\ln(1-\lambda^L)+1-\beta_0-\beta_0(1-\lambda^H)^t+2\beta_0(1-\lambda^L)^t+\\+\ln(1-\lambda^H)\,t(1-\beta_0-2\beta_0(1-\lambda^H)^t+2\beta_0(1-\lambda^L)^t)\end{bmatrix}.$$

Since $-\beta_0t(1-\lambda^L)^t\ln(1-\lambda^L)+1-\beta_0-\beta_0(1-\lambda^H)^t+2\beta_0(1-\lambda^L)^t>0$ and

$1-\beta_0-2\beta_0(1-\lambda^H)^t+2\beta_0(1-\lambda^L)^t>0$, there exists a value of $\ddot{\lambda}^H(\lambda^L)>\lambda^L$ such that

$$-\beta_0 t(1-\lambda^L)^t \ln(1-\lambda^L) + 1 - \beta_0 - \beta_0(1-\lambda^H)^t + 2\beta_0(1-\lambda^L)^t + \\ +\ln(1-\lambda^H)\, t(1-\beta_0 - 2\beta_0(1-\lambda^H)^t + 2\beta_0(1-\lambda^L)^t) < 0 \text{ if } \lambda^H > \ddot{\lambda}^H(\lambda^L) \text{ and}$$

$$-\beta_0 t(1-\lambda^L)^t \ln(1-\lambda^L) + 1 - \beta_0 - \beta_0(1-\lambda^H)^t + 2\beta_0(1-\lambda^L)^t + \\ +\ln(1-\lambda^H)\, t(1-\beta_0 - 2\beta_0(1-\lambda^H)^t + 2\beta_0(1-\lambda^L)^t) > 0 \text{ if } \lambda^H < \ddot{\lambda}^H(\lambda^L).$$

We define a value $\ddot{\lambda}^H(\lambda^L)$ such that $\frac{d\,£}{d\,\lambda^H}$ is equal to zero:

$$\frac{d\,£}{d\,\lambda^H}(\ddot{\lambda}^H) \equiv 0.$$

Therefore, the function £ is increasing in λ^H if $\lambda^H < \ddot{\lambda}^H(\lambda^L)$ and decreasing in λ^H if $\lambda^H > \ddot{\lambda}^H(\lambda^L)$. In addition, if $\lambda^H = \lambda^L$, then $£ = 0$. Since the function £ is continuous in λ^H, there exists $\bar{\bar{\lambda}}^H(\lambda^L) \leq 1$, such that $£ > 0$ for $\lambda^L < \lambda^H < \bar{\bar{\lambda}}^H(\lambda^L)$. We define a value $\bar{\bar{\lambda}}^H(\lambda^L)$ such that the function £ is equal to zero:

$$£\left(\bar{\bar{\lambda}}^H\right) \equiv 0.$$ [49]

As a result, $\frac{\left((1-\lambda^L)^t-(1-\lambda^H)^t\right)}{[P_t^H]^2}$ is an increasing function of t if $\lambda^L < \lambda^H < \bar{\bar{\lambda}}^H(\lambda^L)$.

We established that if

$$[ln(1-\lambda^L) + ln\,\delta](1-\lambda^L)^t P_t^H - (1-\lambda^H)^t(P_t^L ln(1-\lambda^H) + P_t^H ln\delta) < 0 \text{ and}$$

$$P_t^H(1-\lambda^L)^t \ln(1-\lambda^L) >$$

$$(1-\lambda^H)^t(P_t^H + 2\beta_0(1-\lambda^L)^t - 2\beta_0(1-\lambda^H)^t)\ln(1-\lambda^H),$$

then $\frac{d\zeta(t)}{dt} < 0$ and $\frac{P_{T^H}^L}{P_{T^H}^H}\Delta c_{T^H+1}$ is increasing in T^H. Therefore, if both inequalities are satisfied simultaneously, then both rents U^H and U^L are decreasing in T^H, and over-experimentation in T^H is optimal. We next prove that the two inequalities are satisfied simultaneously for a non-empty set of parameters, i.e., $\bar{\lambda}^H(\lambda^L) < \bar{\bar{\lambda}}^H(\lambda^L)$.

The first inequality can be rewritten as

$$(1-\lambda^L)^t P_t^H ln(1-\lambda^L) < (1-\lambda^H)^t(P_t^L ln(1-\lambda^H) + P_t^H ln\delta) - (1-\lambda^L)^t P_t^H\, ln\,\delta.$$

Then the two inequalities are satisfied for a non-empty set of parameters if

$$(1-\lambda^H)^t(P_t^H + 2\beta_0(1-\lambda^L)^t - 2\beta_0(1-\lambda^H)^t)\ln(1-\lambda^H)$$

[49] If $£ > 0$ for all $\lambda^H > \lambda^L$, we then define $\bar{\bar{\lambda}}^H(\lambda^L) \equiv 1$.

$$< (1-\lambda^H)^t (P_t^L ln(1-\lambda^H) + P_t^H ln\delta) - (1-\lambda^L)^t P_t^H \, ln\,\delta$$

$$(1-\lambda^H)^t (P_t^H + 2\beta_0(1-\lambda^L)^t - 2\beta_0(1-\lambda^H)^t) \ln(1-\lambda^H)$$

$$< (1-\lambda^H)^t P_t^L ln(1-\lambda^H) + ((1-\lambda^H)^t - (1-\lambda^L)^t) P_t^H \, ln\,\delta.$$

Since $((1-\lambda^H)^t - (1-\lambda^L)^t) P_t^H \, ln\,\delta > 0$, the inequality above follows from

$$(P_t^H + 2\beta_0(1-\lambda^L)^t - 2\beta_0(1-\lambda^H)^t) > P_t^L,$$

$$1-\beta_0+\beta_0(1-\lambda^H)^t + 2\beta_0(1-\lambda^L)^t - 2\beta_0(1-\lambda^H)^t > 1-\beta_0+\beta_0(1-\lambda^L)^t,$$

$(1-\lambda^L)^t > (1-\lambda^H)^t$, which holds for any t.

Therefore, if $\overline{\lambda}^H(\lambda^L) < \lambda^H < \overline{\overline{\lambda}}^H(\lambda^L)$, then both $U^H(T^H, T^L)$ and $U^L(T^H, T^L)$ are decreasing in T^H, and there is over-experimentation in T^H. *Q.E.D.*

III. Optimal outputs (Proof of Proposition 3)

After success, the optimal $q_t^\theta(\underline{c})$ 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 $(IC^{H,L})$ is not binding]

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

$$V'\left(q_{SB}^H(c_{T^H+1}^H)\right) - c_{T^H+1}^H = \frac{(1-\nu)P_{T^H}^L}{\nu P_{T^H}^H} \Delta c_{T^H+1},$$

$$V'\left(q^L(c_{T^L+1}^L)\right) - c_{T^L+1}^L = 0.$$

Case B. [when $(IC^{H,L})$ is binding]

The following two FOCs imply that there will be overproduction for the low type ($q_{SB}^L(c_{T^L+1}^L) > q_{FB}^L(c_{T^L+1}^L)$) and underproduction for the high type ($q_{SB}^H(c_{T^H+1}^H) < q_{FB}^H(c_{T^H+1}^H)$) 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:

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

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

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

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

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

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

$$V'\left(q^H\left(c^H_{T^H+1}\right)\right) - c^H_{T^H+1} = \frac{\nu\lambda^H P^L_{T^H} \Delta c_{T^H+1}}{f_1(1,T^H)},$$

$$V'\left(q^L\left(c^L_{T^L+1}\right)\right) - c^L_{T^L+1} = -\frac{\nu\lambda^H P^H_{T^L} P^L_{T^H} \Delta c_{T^L+1}}{f_1(1,T^H) P^H_{T^H}}.$$

Q.E.D.