CESifo WORKING PAPERS

7310
2018

October 2018

Learning from Failures: Optimal Contract for Experimentation and Production

Fahad Khalil, Jacques Lawarree, Alexander Rodivilov

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7310*
Category 11: Industrial Organisation

Learning from Failures: Optimal Contract for Experimentation and Production

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 agents. Due to the presence of an optimal production decision after experimentation, we find *over*-experimentation to be optimal. The asymmetric information generated during experimentation makes *over*-production optimal. 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.

JEL-Codes: D820, D830, D860

Keywords: information gathering, optimal contracts, strategic experimentation.

Fahad Khalil
Department of Economics
University of Washington
Seattle / WA / USA
khalil@uw.edu

Jacques Lawarree
Department of Economics
University of Washington
Seattle / WA / USA
lawarree@uw.edu

Alexander Rodivilov
School of Business
Stevens Institute of Technology
Hoboken / NJ / USA
arodivil@stevens.edu

August 17, 2018
We are thankful for the helpful comments of Nageeb Ali, Renato Gomes, Marina Halac, Navin Kartik, David Martimort, Dilip Mookherjee, and Larry Samuelson.

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 allows the agent to better determine its profitability but that is also costly and delays production. Therefore, the duration of the experimentation and the optimal scale of the project are interdependent.

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 owner (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 agent become pessimistic about project profitability. We say that the experimentation stage fails if the agent never learns the true cost.

[1] Other applications are the testing of new drugs, the adoption of new technologies or products, the identification of new investment opportunities, the evaluation of the state of the economy, consumer search, etc. See Krähmer and Strausz (2011) and Manso (2011) for other relevant examples.
[2] See, e.g., Bolton and Harris (1999), or Bergemann and Välimäki (2008).
[3] We present our main insights by assuming that the agent's effort and success in experimentation are 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.

In our model, the agent's efficiency is determined by his probability of success in any given period of the experimentation stage when cost is low. Since the agent is privately informed about his efficiency, when experimentation fails, a lying inefficient agent will have a lower expected cost of production compared to the principal. This difference in expected cost implies that the principal (mistakenly believing the agent is efficient) will overcompensate him in the production stage. Therefore, an inefficient agent must be paid a rent to prevent him from overstating his efficiency.

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 is learned during experimentation. Relative to the nascent literature on incentives for experimentation, reviewed below, the novelty of our approach is to study optimal contracts 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. 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.

First, in a model with experimentation and production, we show that *over* experimentation relative to the first-best is an optimal screening strategy for the principal, whereas under experimentation is the standard result in existing models of experimentation.[5] Since increasing the duration of experimentation helps to raise the chance of success, by asking the agent to *over* experiment, the principal makes it less likely for the agent to fail and exploit the asymmetry of information about expected costs. Moreover, we find that the difference in expected costs is non-monotonic in time: we prove that it is increasing for earlier periods but converges to zero if the experimentation stage is sufficiently long. Intuitively, the updated beliefs for each type initially diverge with successive periods without success, but they must

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

[5] 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.

eventually converge. As a result, increasing the duration of experimentation might help reducing the asymmetric information after a series of failed experiments.

Second, we show that experimentation also influences the choice of output in the production stage. We prove that 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 to a standard adverse selection problem.

Interestingly, 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. 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.

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.[6] 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.

[6] 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)). Our analysis suggests that such practices may also have screening properties in situations where innovators have differences in expertise.

We show that the relative likelihood of success for the two types is monotonic, and it determines the timing of rewards. Given risk-neutrality and absence of moral hazard in our base model, this property also implies that each type is paid a reward only once, whereas a more realistic payment structure would involve rewards distributed over multiple periods. This would be true in a model with moral hazard in experimentation, which is beyond the scope of this paper.[7] However, in an extension section, we do introduce ex post moral hazard simply in this model by assuming that success is private. That leads to moral hazard rent in every period in addition to the previously derived asymmetric information rent.[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.

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

To model information gathering, we rely on the growing literature on contracting for experimentation following Bergmann 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

[7] Halac et al. (2016) illustrate the challenges of having both hidden effort and hidden skill in experimentation in a model without production stage.

[8] The monotonic likelihood ratio of success continues to be the key determinant behind the screening properties of the contract. 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.

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

[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 (prior probability) of the project.

Gomes, Gottlieb and Maestri, there is two-dimensional hidden information, where the agent is privately informed about the quality (prior probability) 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 exogenous output and public success. In section 3, we consider extensions and robustness checks. In particular, we allow the principal to choose output optimally and use it as a screening variable, study ex post moral hazard where the agent can hide success, and the case where success is bad news.

2. The Model (Learning good news)

A principal hires an agent to implement a project. 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

[12] They show that, without the moral hazard constraint, the first best can be reached. In our model, we impose a limited liability instead of a moral hazard constraint.
[13] Modeling time as discrete is more convenient to study the optimal timing of payment (section 2.2.3).

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 always optimal to experiment at least once.[14]

In the main part of the paper, information gathering takes the form of looking for good news (see section 3.3 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 3.2, 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 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:

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

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

$$c_t^\theta = \beta_t^\theta \underline{c} + (1 - \beta_t^\theta)\, \overline{c}.$$

Three aspects of learning are worth noting. First, after each period of failure during experimentation, β_t^θ falls, there is more *pessimism* that the true cost is low, and 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 as both $c_t^H$and c_t^L approach $\overline{c}$. An example of how the expected cost c_t^θ converges to $\overline{c}$ for each type is presented in Figure 1 below.

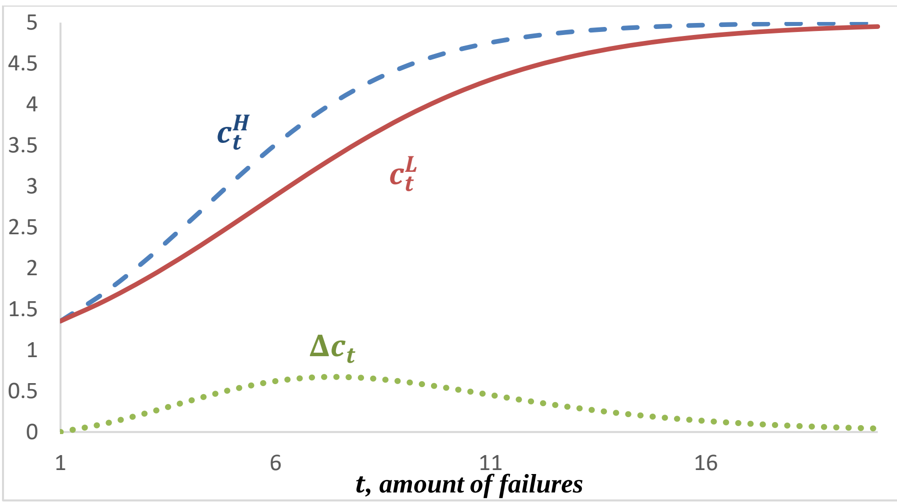

Figure 1. Expected cost with $\lambda^H = 0.35$, $\lambda^L = 0.2$, $\beta_0 = 0.7$, $\underline{c} = 0.5$, $\overline{c} = 5$.

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.[16] Intuitively, each type starts with the same expected cost, which initially diverge as each type of the agent updates differently, but they eventually have to converge to $\overline{c}$.

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

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

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

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] We call q_S the output after success. If experimentation fails, i.e., there is no success during the experimentation stage, production occurs based on the expected cost in period $T + 1$.[19] We call q_F the output after failure. We assume that $q_S > q_F > 0$. Since our main interest is to capture the impact of asymmetric information after failure, it is enough to assume that q_S and q_F are exogenously determined. We relax this assumption in section 3.1, where the output is optimally chosen by the principal given her beliefs.

The contract

Before the experimentation stage takes place, the principal offers the agent a menu of dynamic contracts. Without loss, 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. In terms of notation, in the case of success we include $\underline{c}$ as an argument in the wage and output for each t. In the case of failure, we include the expected cost $c^{\hat{\theta}}_{T^{\hat{\theta}}}$.[20] A contract is defined formally by

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

where $T^{\hat{\theta}}$ is the maximum duration of the experimentation stage for the announced type $\hat{\theta}$, $w_t^{\hat{\theta}}\left(\underline{c}\right)$ is the agent's wage if he observed $c = \underline{c}$ in period $t \leq T^{\hat{\theta}}$ and $w^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}+1} \right)$ is the agent's wage if the agent fails $T^{\hat{\theta}}$ consecutive times.

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

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 $\varpi^{\hat{\theta}}$:

$$U^{\theta}(\varpi^{\hat{\theta}}) = \beta_0 \sum_{t=1}^{T^{\hat{\theta}}} \delta^t \left(1-\lambda^{\theta}\right)^{t-1} \lambda^{\theta} \left(w_t^{\hat{\theta}}(\underline{c}) - \underline{c} q_S\right)$$
$$+\delta^{T^{\hat{\theta}}} \left(1-\beta_0+\beta_0\left(1-\lambda^{\theta}\right)^{T^{\hat{\theta}}}\right)\left(w^{\hat{\theta}}\left(c_{T^{\hat{\theta}}+1}^{\hat{\theta}}\right) - c_{T^{\hat{\theta}}+1}^{\theta} q_F\right).$$

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 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 $\beta_0(1-\lambda^{\theta})^{T^{\hat{\theta}}}$.

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

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

We also assume that the agent must be paid his expected production costs whether experimentation succeeds or fails.[21] Therefore, the individual rationality constraints must be satisfied *ex post* (i.e., after experimentation):

$\left(IRS_t^{\theta}\right)$ $\qquad w_t^{\theta}(\underline{c}) - \underline{c} q_S \geq 0$ for $t \leq T^{\theta}$,

$\left(IRF_{T^{\theta}}^{\theta}\right)$ $\qquad w^{\theta}\left(c_{T^{\theta}+1}^{\theta}\right) - c_{T^{\theta}+1}^{\theta} q_F \geq 0$,

where the S and F are to denote success and failure.

The principal's expected payoff at time zero from a contract ϖ^{θ} offered to the agent of type θ is

$$\pi^{\theta}(\varpi^{\theta}) = \beta_0 \sum_{t=1}^{T^{\theta}} \delta^t \left(1-\lambda^{\theta}\right)^{t-1} \lambda^{\theta} \left(V(q_S) - w_t^{\theta}(\underline{c}) - \Gamma_t\right)$$
$$+\delta^{T^{\theta}} \left(1-\beta_0+\beta_0\left(1-\lambda^{\theta}\right)^{T^{\theta}}\right)\left(V(q_F) - w^{\theta}\left(c_{T^{\theta}+1}^{\theta}\right) - \Gamma_{T^{\theta}}\right),$$

where the cost of experimentation is $\Gamma_t = \frac{\sum_{s=1}^{t} \delta^s \gamma}{\delta^t}$. Thus, the principal's objective function is:

$$\nu \pi^H(\varpi^H) + (1-\nu)\pi^L(\varpi^L).$$

[21] See the recent paper by Krähmer and Strausz (2015) on the importance of ex post participation constraints in a sequential screening model. They provide multiple examples of legal restrictions on penalties on agents prematurely terminating a contract. Our results remain intact as long as there are sufficient restrictions on penalties imposed on the agent. If we assumed unlimited penalties, for example, with only an ex ante participation constraint, we can apply well-known ideas from mechanisms à la Crémer-McLean (1985) that says the principal can still receive the first best profit.

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.

(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 no success is observed during the experimentation stage, the production occurs with output and transfers as specified in the contract.

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 an ex post production stage that is performed by the same agent who experiments. This is common in a wide range of applications such as the regulation of natural monopolies or when a project relies on new, untested technologies.[22]

2.1 The First Best Benchmark

Suppose the agent's type θ is common knowledge *before* the principal offers the contract. The first-best solution is found by maximizing the principal's profit such that the wage to the agent covers the cost in case of success and the expected cost in case of failure.

$$\beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta \left(V(q_S) - w_t^\theta(\underline{c}) - \Gamma_t\right)$$
$$+\delta^{T^\theta}\left(1-\beta_0+\beta_0\left(1-\lambda^\theta\right)^{T^\theta}\right)\left(V(q_F) - w^\theta\left(c^\theta_{T^\theta+1}\right) - \Gamma_{T^\theta}\right)$$

[22] As 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 former 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.

subject to

(IRS_t^θ) $\qquad w_t^\theta(\underline{c}) - \underline{c}q_S \geq 0$ for $t \leq T^\theta$,

$(IRF_{T^\theta}^\theta)$ $\qquad w^\theta(c_{T^\theta+1}^\theta) - c_{T^\theta+1}^\theta q_F \geq 0$.

The individual rationality constraints are binding. If the agent succeeds, the transfers cover the actual cost with no rent given to the agent: $w_t^\theta(\underline{c}) = \underline{c}q_S$ for $t \leq T^\theta$. In case the agent fails, the transfers cover the current *expected* cost and no expected rent is given to the agent: $w^\theta(c_{T^\theta+1}^\theta) = c_{T^\theta+1}^\theta q_F$. 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

$$\delta\beta_{t^\theta}^\theta\lambda^\theta[V(q_S) - \underline{c}q_S] + \delta(1 - \beta_{t^\theta}^\theta\lambda^\theta)[V(q_F) - c_{t^\theta+1}^\theta q_F]$$
$$\geq \gamma + [V(q_F) - c_{t^\theta}^\theta q_F]$$

The intuition is that, by extending the experimentation stage by one additional period, the agent of type θ can learn that $c = \underline{c}$ with probability $\beta_{t^\theta}^\theta\lambda^\theta$.

Note that the first-best termination date of the experimentation stage T_{FB}^θ is a *non-monotonic* function of the agent's type. In the beginning of Appendix A, we formally prove that there exists a unique value of λ^θ called $\hat{\lambda}$, such that:

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

This non-monotonicity is a result of two countervailing forces.[23] 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. 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 $\beta_0(1 - \lambda^\theta)^{t-1}\lambda^\theta$, over time. 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 point. Given these two countervailing forces, the first-best stopping time for the high type agent can be shorter or longer than that of the type L agent depending on the parameters of the problem.[24] Therefore, the first-best stopping time is increasing in the agent's

[23] A similar intuition can be found in Halac et al. (2016) in a model without production.

[24] 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 $T_{FB}^H = 4$, whereas it is optimal to allow the low type agent to experiment for seven periods, $T_{FB}^L = 7$. 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, $T_{FB}^H = 4 > T_{FB}^L = 3$.

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.

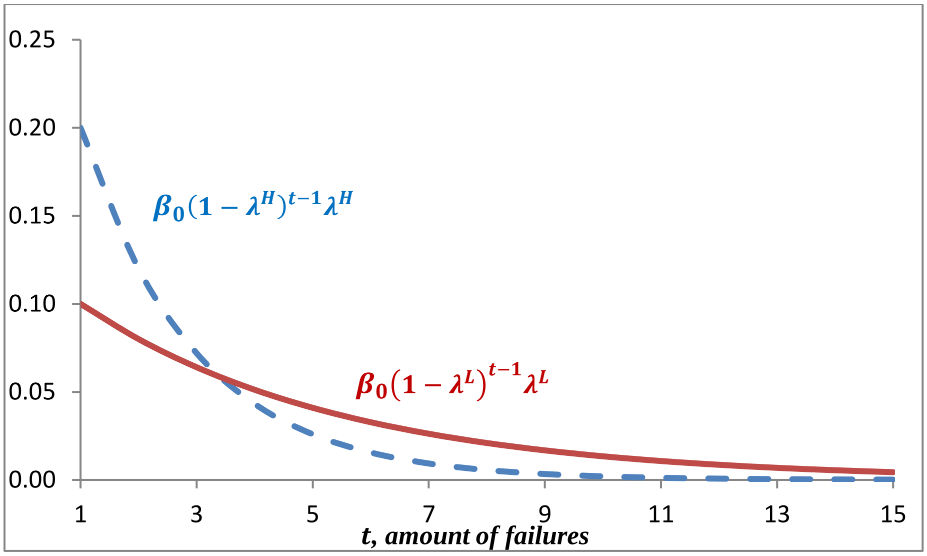

Figure 2. Probability of success with $\lambda^H = 0.4$, $\lambda^L = 0.2$, $\beta_0 = 0.5$.

2.2 Asymmetric information

Assume now that the agent privately knows this type. Recall that all parties have the same expected cost at the outset. Asymmetric information arises in our setting because the two types learn asymmetrically in the experimentation stage, and not because there is any inherent difference in their ability to implement the project. Furthermore, private information can exist only if experimentation fails since the true cost $c = \underline{c}$ is revealed when the agent succeeds.

We now introduce some notation for ex post rent of the agent, which is the rent in the production stage. 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:

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

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

Therefore, the ex post (IR) constraints can be written as:

(IRS_t^θ) $\quad y_t^\theta \geq 0$ for $t \leq T^\theta$,

$(IRF_{T^\theta}^\theta)$ $\quad x^\theta \geq 0$,

where the S and F are to denote success and failure.

To simplify the notation, we denote with P_T^{θ} the probability that an agent of type θ does not succeed 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:

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

Using our notation, the principal maximizes the following objective function subject to (IRS_t^L), $\left(IRF_{T^L}^L\right)$, (IRS_t^H), $\left(IRF_{T^H}^H\right)$, $(IC^{L,H})$, and $(IC^{H,L})$

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

We start by analyzing the two (IC) constraints and first show that the low type always earns a rent.[25] The reason that $(IC^{L,H})$ is binding is that since a high type must be given his expected cost following failure, a low type will have to be given a rent to truthfully report his type as his expected cost is lower. That is, the *RHS* of $(IC^{L,H})$ is strictly positive since $\Delta c_{T^H+1} = c_{T^H+1}^H - c_{T^H+1}^L > 0$. We denote by U^L the rent to the low type by the *LHS* of the $(IC^{L,H})$:

$$U^L \equiv \beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1} \lambda^L y_t^L + \delta^{T^L} P_{T^L}^L x^L > 0.$$

2.2.1. $(IC^{H,L})$ *binding.*

Interestingly, it is also possible that the high type wants to misreport his type such that $(IC^{H,L})$ is binding too. While the low type's benefit from misreporting is positive for sure ($\Delta c_{T^H+1} > 0$), the high type's expected utility from misreporting his type is a *gamble*. There is a positive part since he has a chance to claim the rent U^L of the low type. This part is positively related to Δc_{T^H+1} adjusted by relative probability of collecting the low type's rent. However, there is a negative part as well since he runs the risk of having to produce while being

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

undercompensated since paid as a low type whose expected cost is lower when experimentation fails. This term is positively related to Δc_{T^L+1} adjusted by probability of starting production after failure. This is reflected in $\delta^{T^L} P^H_{T^L} \Delta c_{T^L+1} q_F$ on the RHS of $(IC^{H,L})$. The $(IC^{H,L})$ is binding only when the positive part of the gamble dominates the negative part.[26]

The complexity of the model calls for an illustrative example that demonstrates that $(IC^{H,L})$ might be binding in equilibrium. 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, in the first-best, $T^L = 0$ and $T^H > 0$.[27] 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 ($c^L_{T^H} < c^H_{T^H}$), the low type must be given a rent to induce truth-telling. Consider now the incentives of the high type to claim being low. In this case, production starts immediately without experimentation under identical beliefs about expected cost $\left(\beta_0 \underline{c} + (1-\beta_0)\,\overline{c}\right)$. Therefore, the high type simply collects the rent of the low type without incurring the negative part of the gamble when producing. And, $(IC^{H,L})$ is binding.

In our model, the exact value of the gamble depends on the difference in expected costs and also the relative probabilities of success and failure. These, in turn, are determined by the optimal durations of experimentation stage, T^L and T^H. To see how T^L and T^H affect the value of the gamble, consider again our simple example when the principal asks the low type to (over) experiment (by) one period, $T^L = 1$, and look at the high-type's incentive to misreport again. The high-type now faces a risk. 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 $c^L_2 = \beta^L_2 \underline{c} + (1-\beta^L_2)\,\overline{c} < \overline{c}$. Therefore, by increasing the low-type's duration of experimentation, the principal can use the negative part of the gamble to mitigate the high-type's incentive to lie and, therefore, relax the $(IC^{H,L})$.

[26] Suppose that the principal pays the rent to the low type after an early success. The high type may be interested in claiming to be low type to collect the rent. Indeed, the high type is more likely to succeed early given that the project is low cost. However, misreporting his type is risky for the high type. If he fails to find good news, the principal, believing that he is a low type, will require the agent to produce based on a lower expected cost. Thus, misreporting his type becomes a gamble for the high type: he has a chance to obtain the low-type's rent, but he will be undercompensated relative to the low type in the production stage if he fails during the experimentation stage.

[27] In this example, we emphasize the role of the difference in λs and suppress the impact of the relative probabilities of success and failure.

Another way to illustrate the impact of T^L and T^H on the gamble is to consider a case where the principal must choose an identical length of the experimentation stage for both types ($T^H = T^L = T$).[28] We prove in Proposition 1 below that, in this scenario, the $(IC^{H,L})$ constraint is not binding. Intuitively, since the relevant probabilities P_T^θ and the difference in expected cost Δc_{T+1} are both identical in the positive and negative part of the gamble, they cancel each other. This implies that misreporting his type will be unattractive for the high type.[29]

Proposition 1.

If the duration of experimentation must be chosen identical for both types, $T^H = T^L$, then the high type obtains no rent.

Proof: See Supplementary Appendix B.

Based on Proposition 1, we conclude that it is the principal's choice to have different lengths of the experimentation stage that results in $(IC^{H,L})$ being binding. Since the two types have different efficiencies in experimentation ($\lambda^H > \lambda^L$), the principal optimally chooses different durations of experimentation for each type. This reveals that having both incentive constraints binding might be in the interest of the principal.

In our model, the efficiency in experimentation ($\lambda^H > \lambda^L$) is private information and the principal chooses T^L and T^H to screen the agents. This choice determines equilibrium values of the relative probabilities of success and failure, and the difference in expected costs which determine the gamble. The non-monotonicity of the first-best termination dates (section 2.1) and also non-monotonicity in the difference in expected costs (Figure 1) make it difficult to provide a simple characterization of the optimal durations. This indicates the challenge in deriving a necessary condition for the sign of the gamble.

We provide below sufficient conditions for the $(IC^{H,L})$ constraint to be binding, which are fairly intuitive given the challenges mentioned above. To determine the sufficient conditions, we focus on the adverse selection parameter λ. These conditions say that the constraint is binding as long as the order of termination dates at the optimum remain unaltered from that under the first best. Recalling the definition of $\hat{\lambda}$ from the discussion of first best, for

[28] For example, the FDA requires all the firms to go through the same amount of trials before they are allowed to release new drugs on the market.
[29] We prove formally in Supplementary Appendix B that if T is the same for both types, the gamble is zero.

small values of λ ($\lambda^L < \lambda^H < \hat{\lambda}$) this means $T^L < T^H$, while the opposite is true for high values of λ ($\lambda^H > \lambda^L > \hat{\lambda}$).[30]

Claim. *Sufficient conditions for* $(IC^{H,L})$ *to be binding.*

For any $\lambda^L \in (0,1)$, there exists $0 < \underline{\lambda}^H(\lambda^L) < \overline{\lambda}^H(\lambda^L) < 1$ such that the first best order of termination dates is preserved in equilibrium and $(IC^{H,L})$ binds if either i) $\lambda^H < min\{\underline{\lambda}^H(\lambda^L), \hat{\lambda}\}$ for $\lambda^L < \hat{\lambda}$ or ii) $\lambda^H > \overline{\lambda}^H(\lambda^L) > \lambda^L$ for $\lambda^L \geq \hat{\lambda}$.

Proof: See Appendix A.

The optimal contract is derived formally in Appendix A, and the key results are presented in Propositions 2 and 3. The principal has two tools to screen the agent: the length of the experimentation period and the timing of the payments for each type. We examine each of them first, and later in section 3.2, we let the principal screening by choosing the optimal outputs following both failure and success.

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

While the standard result in the experimentation literature is under-experimentation, we find that *over*-experimentation can also occur when there is a production stage following experimentation. Experimenting longer increases the chance of success, and it can also help reduce information rent. We explain this next.

To give some intuition, consider the case where only $(IC^{L,H})$ binds. The high type gets no rent while rent of the low type is $U^L = \delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q_F$. In this case, there is no benefit from distorting the duration of the experimentation for the low type ($T^L = T^L_{FB}$). However, the principal optimally distorts T^H from its first-best level to mitigate rent of the low type. The reason why the principal may decide to over-experiment is that it might reduce the rent in the production stage, non-existent in standard models of experimentation. First, by extending the experimentation period, the agent is more likely to succeed in experimentation. And, after success, the cost of production is known, and no rent can originate from the production stage. Second, even if experimentation fails, increasing the duration of experimentation can help reduce

[30] As we will see below, when λs are high, the Δc_t function is skewed to the left, and its shape largely determines equilibrium properties as well as our sufficient condition. When λs are small, the Δc_t function is relatively flat, and the relative probabilities of success and failure play a more prominent role.

the asymmetric information and thus the agent's rent in the production stage. This is because the difference in expected cost Δc_t is non-monotonic in t. We show that such over-experimentation is more effective if the agents are sufficiently different in their learning abilities.[31] This does not depend on whether only one or both (IC)s are binding.

When both (IC)s are binding, there is another novel reason for over experimentation. By increasing the duration of experimentation for the low type, T^L, the principal can increase the high type's cost of lying. Recall that the negative part of the high-type's gamble when he lies is represented by $\delta^{T^L} P^H_{T^L} \Delta c_{T^L+1} q_F$ on the RHS of $(IC^{H,L})$. Since Δc_t is non-monotonic, the principal can increase the cost of lying by increasing T^L.

In Proposition 2, we provide sufficient conditions for over-experimentation. In Appendix A, we also give sufficient conditions for under-experimentation to be optimal for the high type. We also provide a numerical example of over-experimentation in Figure 4 below.

Proposition 2. *Sufficient conditions for over-experimentation.*

For any $\lambda^L \in (0,1)$, *there exists* $0 < \underline{\underline{\lambda}}^H(\lambda^L) < \overline{\lambda}^H(\lambda^L) < 1$ *such that the high type over-experiments if* λ^H *is different enough from* λ^L*:*

$$T^H > T^H_{FB} \text{ if } \lambda^H > \overline{\lambda}^H(\lambda^L),$$

and the low type over-experiments if λ^H *is not too different from* λ^L*:*

$$T^L \geq T^L_{FB} \text{ if } \lambda^H < \underline{\underline{\lambda}}^H(\lambda^L).$$

Proof: See Appendix A.

In Figure 3 below, we use an example to illustrate that increasing T^H is more effective when λ^H is higher (relative to λ^L). Start with case when the difference in expected cost is given by the dashed line, and there is over-experimentation ($T^H_{FB} = 10$, while $T^H_{SB} = 11$). By increasing T^H, the principal decreases $P^L_{T^H} \Delta c_{T^H+1}$, which in turn decreases the positive part of the gamble and makes lying less attractive for the high type. This effect is even stronger for a higher λ^H (see plain line), where the difference in expected cost is skewed to the left with a relatively high Δc_{T^H+1} at the first best $T^H_{FB} = 5$. Now, increasing T^H is even more effective since the decrease in $P^L_{T^H} \Delta c_{T^H+1}$ is even sharper.

[31] The opposite is true when the λs are relatively small and close to each other. Then, the principal prefers to under-experiment which reduces the difference in expected cost.

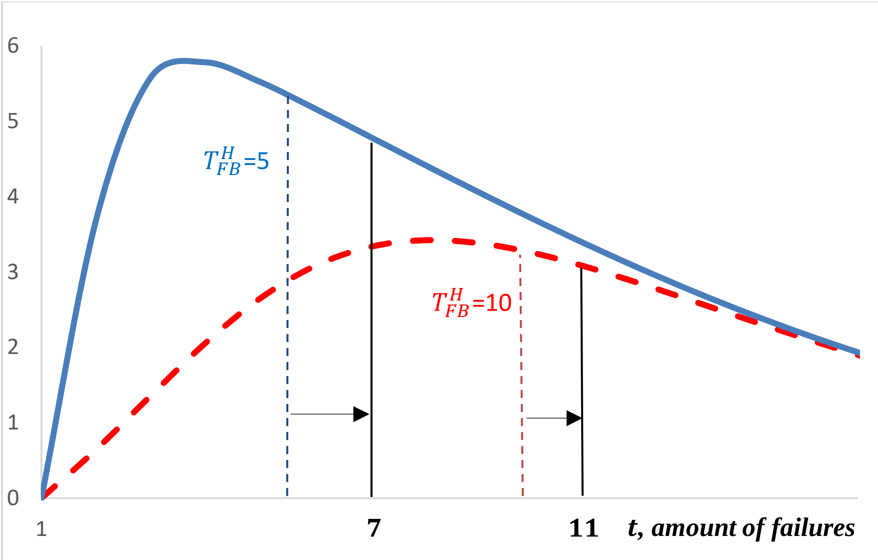

Figure 3. Difference in the expected cost with $\lambda^H = 0.35 (= 0.82)$, $\lambda^L = 0.2$, $\beta_0 = 0.7$, $\underline{c} = 0.1$, $\overline{c} = 10$, $V(q) = 3.5\sqrt{q}$, $\delta = 0.9$, and $\gamma = 1$.

Finally, as in the first best, either type may experiment longer, and T^L can be larger or smaller than T^H.

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

The principal chooses the timing of rewards and the duration of experimentation at the same time as part of the contract, and, in this section, we analyze 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?

Recall that the low type receives a strictly positive rent, $U^L > 0$, and $(IC^{L,H})$ is binding. The principal has to determine when to pay this rent to the low type while taking into account the high type's incentive to misreport. This is achieved by rewarding the low type at events which are relatively more likely for the low type. Since the high type is relatively more likely to succeed early, he is optimally rewarded early, while the reward to the low type is optimally postponed. Furthermore, since the high type is more likely to fail if experimentation lasts long enough, rewarding the low type after late success or failure will depend on the length of the experimentation stage, which is determined by the cost of experimentation (γ).

We will next characterize the optimal timing of payments. There are two cases depending on whether only $(IC^{L,H})$ or both IC constraints are binding.

Proposition 3. *The optimal timing of payments.*

Case A: 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.

Case B: Both types' IC are binding.

The principal must reward the high-type 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.

If $(IC^{H,L})$ is not binding, we show in Case A of Appendix A that 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 $\beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1} \lambda^L y_t^L + \delta^{T^L} P_{T^L}^L x^L = \delta^{T^H} P_{T^H}^L \Delta c_{T^H+1} q_F$. Therefore, the principal can reward either early or late success, or even failure.

If $(IC^{H,L})$ is binding, the high type has an incentive to claim to be a low type and we are in Case B. This is the more interesting case and we focus on its characterization in this section. The timing of rewards to the low type depends on which scheme yields a lower RHS of $(IC^{H,L})$.

We start by analyzing the case where the principal rewards the agent after success. In the optimal timing of payments is determined by the relative likelihood ratio of success in period t,

$$\frac{\beta_0 (1-\lambda^H)^{t-1} \lambda^H}{\beta_0 (1-\lambda^L)^{t-1} \lambda^L},$$

which is strictly decreasing in t. Therefore, if the principal chooses to reward the low type for success, she will optimally postpone this reward till the very last period, T^L, to minimize the high type's incentive to misreport.

To see why the principal may want to reward the low type agent after failure, we need to compare the relative likelihood of ratio 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 $\left(\frac{P_{T^L}^H}{P_{T^L}^L}\right)$. We show in Appendix A that there is a unique period $\hat{T}^L$ such that the two relative probabilities are equal:[32]

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

This critical value $\hat{T}^L$ (depicted in Figure 4 below) determines which type is relatively more likely to succeed or fail during the experimentation stage. In any period $t < \hat{T}^L$, the high type who chooses the contract designed for the low type is relatively more likely to succeed than fail compared to the low type. For $t > \hat{T}^L$, the opposite is true. This feature plays an important role in structuring the optimal contract. The critical value $\hat{T}^L$ determines whether the principal will choose to reward success or failure in the optimal contract.

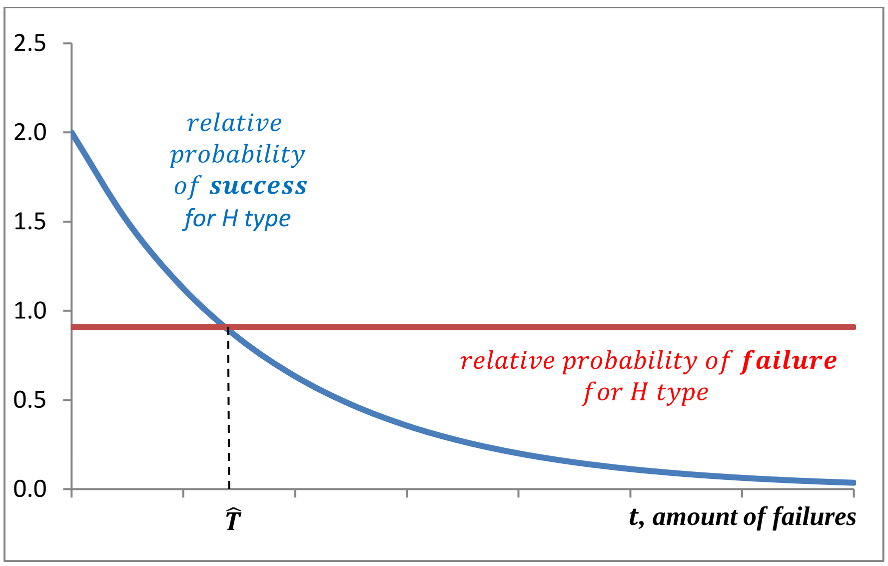

Figure 4. Relative probability of success/failure with $\lambda^H = 0.4$, $\lambda^L = 0.2$, $\beta_0 = 0.5$.

If the principal wants to reward the low type after success, it will only be optimal if the experimentation stage lasts long enough. If $T^L > \hat{T}^L$, the principal can reward success in the last period as the relative probability of success is declining over time and the rent is the smallest in the last period T^L. If the experimentation stage is short, $T^L < \hat{T}^L$, the principal will pay the rent to the low type by rewarding failure since the high type is relatively more likely to succeed during the experimentation stage.

The important question is therefore what the optimal value of T^L is relative to $\hat{T}^L$ and consequently whether the principal should reward success or failure. The optimal value of T^L is

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

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, if the high type also 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.

3. Extensions

3.1. *Over*-production as a screening device

In this section, we allow the principal to choose output optimally after success and after failure, and she can now use output as another screening variable. While our main findings continue to hold, the key new results are that if the experimentation stage fails, the inefficient type is asked to *over*-produce, while the efficient type under-produces. Just like over-experimentation, over-production can be used to increase the cost of lying.

When output is optimally chosen by the principal in the contract, q_S is now replaced by by $q_t^\theta(\underline{c})$ and is determined by $V'\left(q_t^\theta(\underline{c})\right) = \underline{c}$. The main change from the base model is that output after failure which is denoted by $q^\theta(c^\theta_{T^\theta+1})$, can vary continuously depending on the expected cost. We can simply replace q_F by $q^\theta(c^\theta_{T^\theta+1})$ and q_S by $q_t^\theta(\underline{c})$ in the principal's problem.

We derive the formal output scheme in Supplementary Appendix C but present the intuition here. 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, asymmetric information will induce the principal to 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 under-produces 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 under-produces 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. We give below the sufficient conditions such that both incentive constraints are binding when output is variable, and these conditions are similar to the ones identified earlier. When both incentive constraints bind, 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 $q^H\left(c^H_{T^H+1}\right)$. By making this output smaller, the principal can decrease the positive component of the gamble. The negative part is now given by $\delta^{T^L} P^H_{T^L} \Delta c_{T^L+1} q^L\left(c^L_{T^L+1}\right)$, and it comes from the higher expected cost of producing the output required from the low type. By making this output higher, the principal can increase the cost of lying and lower the rent of the high type. We summarize the results in Proposition 4 below.

Proposition 4. *Optimal output.*

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

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

After failure, the high type under-produces relative to the first best output:

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

After failure, the low type over-produces:

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

Proof: See Supplementary Appendix C.

As in our main model, we now derive sufficient conditions for both IC to bind and for over-experimentation to occur and discuss them in turn. We show below that the sufficient conditions for $(IC^{H,L})$ to be binding are stricter than in our main model with an exogenous output. This is not surprising since the principal now has an additional screening instrument to reduce the high type's incentives to misreport. We introduce $\dot{\lambda}^H(\lambda^L) < min\{\underline{\lambda}^H(\lambda^L), \hat{\lambda}\}$ and $\ddot{\lambda}^H(\lambda^L) > \overline{\lambda}^H(\lambda^L)$ to provide sufficient condition for both (IC) to be binding simultaneously.

Claim. *Sufficient condition for $(IC^{H,L})$ to be binding with endogenous output.*

For any $\lambda^L \in (0,1)$, there exists $0 < \dot{\lambda}^H(\lambda^L) < \ddot{\lambda}^H(\lambda^L) < 1$ such that $(IC^{H,L})$ binds if either i) $\lambda^H < \dot{\lambda}^H(\lambda^L)$ for $\lambda^L < \hat{\lambda}$ or ii) $\lambda^H > \ddot{\lambda}^H(\lambda^L) > \lambda^L$ for $\lambda^L \geq \hat{\lambda}$.

Proof: See Supplementary Appendix C.

The exact distortions in $q^H\left(c^H_{T^H+1}\right)$ and $q^L\left(c^L_{T^L+1}\right)$ are chosen by the principal to mitigate the rent. Since the agent's rent depends on both output and the difference in expected costs after failure, distortions in output are proportional to Δc_{T^H+1} and Δc_{T^L+1}, which are non-monotonic in time. This makes it challenging to derive necessary and sufficient conditions for over/under experimentation when output is chosen optimally. We characterize sufficient conditions for over and under experimentation for the high type below.

Claim. *Sufficient conditions for over-experimentation.*

There exist $0 < \overline{\lambda}^L < \bar{\bar{\lambda}}^L < 1$ and $\bar{\bar{\lambda}}^H > \underline{\lambda}^H(\lambda^L)$ such that

$T^H > T^H_{FB}$ (over experimentation is optimal) if $\bar{\bar{\lambda}}^H > \lambda^H > \overline{\lambda}^H(\lambda^L)$ and $\lambda^L > \overline{\lambda}^L$;

$T^H < T^H_{FB}$ (under experimentation is optimal) if $\lambda^H < \underline{\lambda}^H(\lambda^L)$ and $\lambda^L < \bar{\bar{\lambda}}^L$.

Proof: See Supplementary Appendix C.

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

In the base model, we have suppressed moral hazard to highlight the screening properties of the timing of rewards, which allowed us to isolate the importance of the monotonic likelihood ratio of success in screening the two types. 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 each period, but our key insights regarding the screening properties of the optimal contract remain intact. 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. Furthermore, the agent's rent is still determined by the difference in expected cost, which remains non-monotonic in time. Thus, the reasons for over-experimentation also remain intact.

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 $t \geq j$. Thus, we assume that success generates hard information that can be presented to the principal when desired, but it cannot be fabricated. The agent's decision to reveal success is affected not only by the payment and the output tied to success/failure in the particular period j, but also by the payment and output in all subsequent periods of the experimentation stage.

Note first that if the agent succeeds but hides it, the principal and the agent's beliefs are different at the production stage: the principal's expected cost is given by $c^{\theta}_{T^{\theta}+1}$ while the agent knows the true cost is $\underline{c}$. In addition to the existing (IR) and (IC) constraints, the optimal scheme must now satisfy the following new ex post moral hazard constraints:

(EMH^{θ}) $\quad y^{\theta}_{T^{\theta}} \geq x^{\theta} + (c^{\theta}_{T^{\theta}+1} - \underline{c})q_F$ for $\theta = H, L$, and

(EMP^{θ}_t) $\quad y^{\theta}_t \geq \delta y^{\theta}_{t+1}$ for $t \leq T^{\theta} - 1$.

The (EMH^{θ}) constraint makes it unprofitable for the agent to hide success in the last period. The (EMP^{θ}_t) 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. The principal's problem is exacerbated by having to address the ex post moral hazard constraints in addition to all the constraints presented before. First, as formally shown in the Supplementary Appendix D, both $(IC^{H,L})$ and $(IC^{L,H})$ may be slack, and either or both may be binding.[33] Since the ex post moral hazard constraints imply that both types will receive rent, these rents may be sufficient to satisfy the (IC) constraints. Second, private observation of success increases the cost of paying a reward after failure. When the principal rewards failure with $x^{\theta} > 0$, the (EMH^{θ}) constraint forces her to also reward success in the last period ($y^{\theta}_{T^{\theta}} > 0$ because of (EMH^{θ})) and in all previous periods ($y^{\theta}_t > 0$ because of (EMP^{θ}_t)). However, we show below that it can still be optimal to reward failure.

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

Proposition 5.

When success can be hidden, the principal must reward success in every period for each type. When both the $(IC^{L,H})$ *and* $(IC^{H,L})$ *constraints bind and the optimal* $T^L \leq \hat{T}^L$*, it is optimal to reward failure for the low type.*

Proof: See Supplementary Appendix D.

Details and the formal proof are in the Supplementary Appendix D. Here we provide some intuition why rewarding failure or postponing rewards remain optimal even when the agent privately observes success. We also provide an example below where this occurs in equilibrium. The argument for postponing rewards to the low type to effectively screen the high type applies even when success is privately observed. This is because the relative probability of success between types is not affected by the two ex post moral hazard constraints above. An increase of \$1 in x^θ causes an increase of \$1 in $y^\theta_{T^\theta}$, which in turn causes an increase in all the previous y^θ_t according to the discount factor. Therefore, the increases in $y^\theta_{T^\theta}$ and y^θ_t are not driven by the relative probability of success between types. And, just as in Proposition 3, we again find that it is optimal to postpone the reward for the low type if he experiments for a relatively brief length of time and both $(IC^{H,L})$ and $(IC^{L,H})$ are binding. For example, when $\beta_0 = 0.7$ $\gamma = 2$, $\lambda^L = 0.28$, $\lambda^H = 0.7$ the principal optimally chooses $T^H = 1$, $T^L = 2$ and rewards the low type after failure since $\hat{T}^L = 3$.

While we have focused on how the ex post moral hazard affects the benefit of rewarding failure, those constraints also affect the other optimal variables of the contract. For instance, the constraint (EMP^θ_t) can be relaxed by decreasing either T^θ. So, we expect a shorter experimentation stage and a lower output when success can he hidden.

3.3. 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 $c = \bar{c}$. 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 an event of observing $c = \bar{c}$ by the agent "success" although this is bad news for the principal. If the agent's type were common knowledge, the principal and agent both become more optimistic if success is not achieved in a particular period and relatively more optimistic when the agent is a high type than a low type. Also, as time goes by without learning that the cost is high, the expected cost becomes lower due to Bayesian updating and converges to $\underline{c}$. In addition, the difference in the expected cost is now negative, $\Delta c_t = c_t^H - c_t^L < 0$ since the H 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.

Denoting by β_t^θ the updated belief of agent θ that the cost is actually high, the type θ's expected cost is then $c_t^\theta = \beta_t^\theta \overline{c} + \left(1 - \beta_t^\theta\right) \underline{c}$. An agent of type θ, announcing his type as $\hat{\theta}$, receives expected utility $U^\theta\left(\varpi^{\hat{\theta}}\right)$ at time zero from a contract $\varpi^{\hat{\theta}}$, but now $y_t^{\hat{\theta}} = w_t^{\hat{\theta}}(\overline{c}) - \overline{c} q_t^{\hat{\theta}}(\overline{c})$ is a function of $\overline{c}$.

Under asymmetric information about the agent's type, the intuition behind the key incentive problem is similar to that under learning good news. However, it is now the high type who has an incentive to claim to be a low type. Given the same length of experimentation, following failure, the expected cost is higher for the low type. Thus, a high type now has an incentive to claim to be a low type: since a low type must be given his expected cost following failure, a high type will have to be given a rent to truthfully report his type as his expected cost is lower, that is, $c_{T^L+1}^H < c_{T^L+1}^L$. The details of the optimization problem mirror the case for good news of Propositions 2, 3, and 4 and the results are similar. We present results formally in Proposition 6 in Supplementary Appendix E.

We find similar restrictions when both (IC) constraints bind as in Propositions 2, 3 and 4. The type of news, however, determines length of experimentation decisions. 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.

4. Conclusions

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.

There is also a significant literature on endogenous information gathering in contract theory but typically relying on static models of learning. By modeling experimentation in a dynamic setting, we have endogenized the degree of asymmetry of information in a principal agent model and related it to the length of the learning stage.

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

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 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 (2011), 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.

Sadler E. (2017). Dead Ends. *Working Paper*, April 2017.

Appendix A (Proof of Sufficient conditions for $(IC^{H,L})$ to be binding, Propositions 2 and 3)

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

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

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

$$\beta_t \lambda[V(q_S) - \underline{c} q_S] + (1 - \beta_t \lambda)[V(q_F) - c_{t+1} q_F] = \gamma + [V(q_F) - c_t q_F].$$

Rewriting it next we have

$$\beta_t \lambda[V(q_S) - \underline{c} q_S - V(q_F)] + q_F[c_t - c_{t+1}(1 - \beta_t \lambda)] = \gamma,$$

which given that $c_t - c_{t+1}(1 - \beta_t \lambda) = \beta_t \lambda \underline{c}$, can be rewritten next as

$$\beta_t \lambda = \frac{\gamma}{(V(q_S) - \underline{c} q_S) - (V(q_F) - \underline{c} q_F)},$$

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

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

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

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

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

and $\frac{\partial\left[\frac{\lambda\beta_0(1-\lambda)^{t-1}}{\beta_0(1-\lambda)^{t-1}+(1-\beta_0)}\right]}{\partial t} = \frac{\beta_0(1-\beta_0)\lambda(1-\lambda)^{t-1}\, ln(1-\lambda)}{\left(\beta_0(1-\lambda)^{t-1}+(1-\beta_0)\right)^2}$, we have

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

For $\frac{dt}{d\lambda} < 0$ it is necessary and sufficient that $(1-\beta_0)(1-\lambda t) + \beta_0(1-\lambda)^t < 0$. 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 $\hat{\lambda} = \frac{1-\sqrt{1-\beta_0}}{\beta_0}$.[34] *Q.E.D.*

Sufficient conditions for $(\boldsymbol{IC^{H,L}})$ to be binding.

Claim. For any $\lambda^L \in (0,1)$, there exists $0 < \underline{\lambda}^H(\lambda^L) < \overline{\lambda}^H(\lambda^L) < 1$ such that the first best order of termination dates is preserved in equilibrium and $(IC^{H,L})$ binds if either i) $\lambda^H < min\{\underline{\lambda}^H(\lambda^L), \hat{\lambda}\}$ for $\lambda^L < \hat{\lambda}$ or ii) $\lambda^H > \overline{\lambda}^H(\lambda^L) > \lambda^L$ for $\lambda^L \geq \hat{\lambda}$.

Proof: We will prove later in this Appendix A (see Optimal payment structure) that when the $(IC^{H,L})$ binds, the principal will pay U^L by only rewarding failure ($x^L > 0$), or only rewarding success in the last period T^L ($y^L_{T^L} > 0$). In step 1 below, we characterize a function $\zeta(t)$ that determines the sign of the high-type's gamble, i.e., if $(IC^{H,L})$ is binding, regardless of whether the agent is optimally paid after failure or success. In step 2, we characterize values of λ^L and λ^H such that $\zeta(t)$ is monotonic. These two steps together imply that the gamble is positive under this set of λ^L and λ^H if the order of optimal termination dates (T^L and T^H) is the same as in the first best. In step 3, we prove that under this set of λ^L and λ^H it is indeed optimal for the principal to make the $(IC^{H,L})$ binding in equilibrium. Therefore, the sufficient conditions characterize values of λ^L and λ^H such that the principal finds it optimal to preserve the first best order of termination dates in equilibrium.

We begin by deriving the gamble if the agent is optimally rewarded after failure and after success in the last period T^L. If the principal rewards the low type after failure, the high type's expected utility from misreporting (i.e., the RHS of the $(IC^{H,L})$ constraint) is:

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

[34] 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 $\beta_0 < 1$.

If the principal rewards the low type only after success in period T^L, the high type's expected utility from misreporting (i.e., the RHS of the $(IC^{H,L})$ constraint) is:

$$\frac{\beta_0(1-\lambda^H)^{T^L-1}\lambda^H}{\beta_0(1-\lambda^L)^{T^L-1}\lambda^L}\delta^{T^H}P^L_{T^H}\Delta c_{T^H+1}q_F - \delta^{T^L}P^H_{T^L}\Delta c_{T^L}q_F =$$

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

Step 1. The gamble for the high type is positive (regardless of the payment scheme) if and only if $\zeta(T^H) > \zeta(T^L)$, where $\zeta(t) \equiv \delta^t P^L_t(\beta^L_{t+1} - \beta^H_{t+1})$.

a) Consider the case of reward after failure first, $q_F P^H_{T^L}\left(\delta^{T^H}\frac{P^L_{T^H}}{P^L_{T^L}}\Delta c_{T^H+1} - \delta^{T^L}\Delta c_{T^L+1}\right)$. Given that $\Delta c_t = (\overline{c} - \underline{c})(\beta^L_t - \beta^H_t)$, the gamble is positive if and only if

$$\delta^{T^H}\frac{P^L_{T^H}}{P^L_{T^L}}\Delta c_{T^H+1} - \delta^{T^L}\Delta c_{T^L+1} > 0,$$

$$\delta^{T^H}P^L_{T^H}\left(\beta^L_{T^H+1} - \beta^H_{T^H+1}\right) > \delta^{T^L}P^L_{T^L}\left(\beta^L_{T^L+1} - \beta^H_{T^L+1}\right),$$

which can be re-written as $\zeta(T^H) > \zeta(T^L)$.

b) Consider now the case of reward after success, $q_F\left(\frac{(1-\lambda^H)^{T^L-1}\lambda^H}{(1-\lambda^L)^{T^L-1}\lambda^L}\delta^{T^H}P^L_{T^H}\Delta c_{T^H+1} - \delta^{T^L}P^H_{T^L}\Delta c_{T^L+1}\right)$. Similarly, the gamble is positive if and only if

$$\frac{(1-\lambda^H)^{T^L-1}\lambda^H}{(1-\lambda^L)^{T^L-1}\lambda^L}\delta^{T^H}P^L_{T^H}\Delta c_{T^H+1} - \delta^{T^L}P^H_{T^L}\Delta c_{T^L+1} > 0,$$

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

We will prove later in this Appendix A (see Optimal payment structure) that when gamble for the high type is positive, the low type is rewarded for success if and only if $\frac{(1-\lambda^H)^{T^L-1}\lambda^H}{(1-\lambda^L)^{T^L-1}\lambda^L} > \frac{P^H_{T^L}}{P^L_{T^L}}$ or, alternatively, $\frac{(1-\lambda^H)^{T^L-1}\lambda^H P^L_{T^L}}{(1-\lambda^L)^{T^L-1}\lambda^L P^H_{T^L}} > 1$. Therefore, a sufficient condition for the gamble to be positive when the low type is rewarded for success is $1 > \frac{\delta^{T^L}P^L_{T^L}\left(\beta^L_{T^L+1}-\beta^H_{T^L+1}\right)}{\delta^{T^H}P^L_{T^H}\left(\beta^L_{T^H+1}-\beta^H_{T^H+1}\right)}$. Given that $\frac{\delta^{T^L}P^L_{T^L}\left(\beta^L_{T^L+1}-\beta^H_{T^L+1}\right)}{\delta^{T^H}P^L_{T^H}\left(\beta^L_{T^H+1}-\beta^H_{T^H+1}\right)} = \frac{\zeta(T^L)}{\zeta(T^H)}$, the condition above may be rewritten as $1 > \frac{\zeta(T^L)}{\zeta(T^H)}$.

Therefore, we have proved that a sufficient condition for the gamble to be positive (regardless of the payment scheme) is $\zeta(T^H) > \zeta(T^L)$.

We now explore the properties of $\zeta(t)$ function.

Step 2. Next, we prove that for any $\lambda^L \in (0,1)$, there exist $\overline{\lambda}^H(\lambda^L)$ and $\underline{\lambda}^H(\lambda^L)$, $1 > \overline{\lambda}^H > \underline{\lambda}^H > 0$ such that $\frac{d\zeta(t)}{dt} < 0$ if $\lambda^H > \overline{\lambda}^H$ and $\frac{d\zeta(t)}{dt} > 0$ if $\lambda^H < \underline{\lambda}^H$.

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

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

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

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

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

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

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

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 first that $\phi(\lambda^H) < 0$ for any t if λ^H is sufficiently high.

Since both $(1-\lambda^L)^t P_t^H$ and $(1-\lambda^H)^t(P_t^L ln(1-\lambda^H) + P_t^H ln\delta)$ are increasing in t, we have

$$\phi(\lambda^H) < (1-\beta_0)(1-\lambda^L)^{\overline{T}}\, ln[\delta(1-\lambda^L)] - (1-\lambda^H)(P_1^L ln(1-\lambda^H) + P_1^H ln\delta),$$

where $\overline{T} = \max\{T^H, T^L\}$.

Next, $(1-\lambda^H)(P_1^L ln(1-\lambda^H) + P_1^H ln\delta) > (1-\lambda^H)(1-\beta_0\lambda^L)\, ln[\delta(1-\lambda^H)]$ and $\frac{ln[\delta(1-\lambda^H)]}{ln[\delta(1-\lambda^L)]} > 1$. Therefore, $\frac{(1-\beta_0)(1-\lambda^L)^{\overline{T}}}{(1-\lambda^H)(1-\beta_0\lambda^L)} > 1 \Rightarrow \phi(\lambda^H) < 0$.

Rearranging the above, we have $\phi(\lambda^H) < 0$ for any t if $\lambda^H > 1 - \frac{(1-\beta_0)(1-\lambda^L)^{\overline{T}}}{(1-\beta_0\lambda^L)}$.

Denote

$$\overline{\lambda}^H \equiv 1 - \frac{(1-\beta_0)(1-\lambda^L)^{\overline{T}}}{(1-\beta_0\lambda^L)}.$$

Therefore, for any t, $\phi(\lambda^H) < 0$ if $\lambda^H > \overline{\lambda}^H$ and, consequently, $\frac{d\zeta(t)}{dt} < 0$ for any t if $\lambda^H > \overline{\lambda}^H$.

We next prove $\phi(\lambda^H) > 0$ for any t if λ^H is sufficiently low.

Similarly, because both $(1-\lambda^L)^t P_t^H$ and $(1-\lambda^H)^t (P_t^L ln(1-\lambda^H) + P_t^H ln\delta)$ are increasing in t,

$$\phi(\lambda^H) > P_1^H (1-\lambda^L)\, ln[\delta(1-\lambda^L)] - (1-\lambda^H)^{\overline{T}} \left(P_{\overline{T}}^L ln(1-\lambda^H) + P_{\overline{T}}^H ln\delta\right).$$

Next, $(1-\lambda^H)^{\overline{T}} P_{\overline{T}}^H\, ln[\delta(1-\lambda^H)] > (1-\lambda^H)^{\overline{T}} \left(P_{\overline{T}}^L ln(1-\lambda^H) + P_{\overline{T}}^H ln\delta\right)$ and $\frac{ln[\delta(1-\lambda^H)]}{ln[\delta(1-\lambda^L)]} > 1$.

Therefore, $(1-\beta_0\lambda^H)(1-\lambda^L) < (1-\lambda^H)^{\overline{T}}\left(1-\beta_0+\beta_0(1-\lambda^H)^{\overline{T}}\right) \Rightarrow \phi(\lambda^H) > 0$.

Rearranging the above, we have $\phi(\lambda^H) > 0$ for any t if

$$\frac{(1-\beta_0\lambda^H)}{(1-\lambda^H)^{\overline{T}}}(1-\lambda^L) < \left(1-\beta_0+\beta_0(1-\lambda^H)^{\overline{T}}\right).$$

The left-hand side of the above inequality, $\frac{(1-\beta_0\lambda^H)}{(1-\lambda^H)^{\overline{T}}}(1-\lambda^L)$, is increasing in λ^H:

$$\frac{d\left[\frac{(1-\beta_0\lambda^H)}{(1-\lambda^H)^{\overline{T}}}\right]}{d\lambda^H} = \frac{-\beta_0(1-\lambda^H)+\overline{T}(1-\beta_0\lambda^H)}{(1-\lambda^H)^{\overline{T}+1}} > 0.$$

The right-hand side of the above inequality, $1-\beta_0+\beta_0(1-\lambda^H)^{\overline{T}}$, is decreasing in λ^H:

$$\frac{d\left[1-\beta_0+\beta_0(1-\lambda^H)^{\overline{T}}\right]}{d\lambda^H} = -\beta_0\overline{T}(1-\lambda^H)^{\overline{T}-1} < 0.$$

Therefore, there exists a unique value of $\lambda^H = \underline{\lambda}^H$, such that if $\lambda^H < \underline{\lambda}^H$: $\frac{(1-\beta_0\lambda^H)}{(1-\lambda^H)^{\overline{T}}}(1-\lambda^L) < \left(1-\beta_0+\beta_0(1-\lambda^H)^{\overline{T}}\right)$ and $\frac{(1-\beta_0\lambda^H)}{(1-\lambda^H)^{\overline{T}}}(1-\lambda^L) > \left(1-\beta_0+\beta_0(1-\lambda^H)^{\overline{T}}\right)$ if $\lambda^H > \underline{\lambda}^H$.

Denote $\underline{\lambda}^H$ as the unique value of λ^H that makes the LHS and RHS equal

$$\underline{\lambda}^H: \frac{(1-\beta_0\underline{\lambda}^H)}{(1-\underline{\lambda}^H)^{\overline{T}}}(1-\lambda^L) = \left(1-\beta_0+\beta_0\left(1-\underline{\lambda}^H\right)^{\overline{T}}\right).$$

Therefore, for any t, $\phi(\lambda^H) > 0$ if $\lambda^H < \underline{\lambda}^H$ and, consequently, $\frac{d\zeta(t)}{dt} > 0$ for any t if $\lambda^H < \underline{\lambda}^H$.

Step 3. We now prove by contradiction that for $\lambda^H < min\{\underline{\lambda}^H(\lambda^L), \hat{\lambda}\}$ and $\hat{\lambda} < \lambda^L < \overline{\lambda}^H < \lambda^H$, it is optimal to have $(IC^{H,L})$ binding, i.e., pay rent to both types rather than the low type only.

We consider $\hat{\lambda} < \lambda^L < \overline{\lambda}^H < \lambda^H$ and argue later that case $0 < \lambda^L < \lambda^H < \min\{\underline{\lambda}^H, \hat{\lambda}\}$ can be proven by repeating the same procedure.

Step 3a. Assume $(IC^{H,L})$ is not binding and it is optimal to pay rent only to the low type. Denote optimal duration of experimentation stage for both types T_{SB}^L and T_{SB}^H, respectively. We prove later in this Appendix A (see Optimal length of experimentation, Case A) that in this case $T_{FB}^H < T_{SB}^H$ and $T_{FB}^L = T_{SB}^L$.

Since gamble for the high type is positive if $\zeta(T^H) > \zeta(T^L)$ and $\frac{d\zeta(t)}{dt} < 0$ for $\lambda^H > \overline{\lambda}^H$, it must be that $T_{SB}^H > T_{FB}^L$ (otherwise the gamble for the high type would be positive and both types would collect rent). Denote the expected surplus net of costs for $\theta = H, L$ by $\Omega^\theta(\varpi^\theta) = \beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta [V(q_S) - \underline{c} q_S - \Gamma_t] + \delta^{T^\theta} P_{T^\theta}^\theta [V(q_F) - c_{T^\theta+1}^\theta q_F - \Gamma_{T^\theta}]$. Since the principal optimally distorts T^H, it must be that

$$\nu\Omega^H(T_{SB}^H) + (1-\nu)\Omega^L(T_{FB}^L) - (1-\nu)U^L(T_{SB}^H) >$$
$$\nu\Omega^H(T_{FB}^H) + (1-\nu)\Omega^L(T_{FB}^L) - \nu U^H(T_{FB}^H, T_{FB}^L) - (1-\nu)U^L(T_{FB}^H, T_{FB}^L),$$

where $U^H(T_{FB}^H, T_{FB}^L)$ and $U^L(T_{FB}^H, T_{FB}^L)$ appear on the RHS because the $T_{FB}^H < T_{FB}^L$ and, given and $\frac{d\zeta(t)}{dt} < 0$, the gamble is positive (so the principal has to pay rent to both types).

Given that $T_{FB}^L = T_{SB}^L$, the condition above simplifies to

$$\nu U^H(T_{FB}^H, T_{FB}^L) + (1-\nu)U^L(T_{FB}^H, T_{FB}^L) - (1-\nu)U^L(T_{SB}^H) > \nu[\Omega^H(T_{FB}^H) - \Omega^H(T_{SB}^H)].$$

Step 3b. We now show that our assumption (that $(IC^{H,L})$ is not binding and it is optimal to pay rent only to the low type) leads to a contradiction by proving that the principal can be better off by paying rent to both types instead.

Consider another contract with $T^L = T_{FB}^L$ and $T^H = T_{FB}^L - \varepsilon$, where agent's rewards are chosen optimally given these Ts.[35] With $T^L = T_{FB}^L$ and $T^H = T_{FB}^L - \varepsilon$, the principal's expected profit becomes

$$\nu\Omega^H(T_{FB}^L - \varepsilon) + (1-\nu)\Omega^L(T_{FB}^L) - \nu U^H(T_{FB}^L - \varepsilon, T_{FB}^L) - (1-\nu)U^L(T_{FB}^L - \varepsilon, T_{FB}^L).$$

Therefore, the principal is better off with the newly introduced contract ($T^L = T_{FB}^L$ and $T^H = T_{FB}^L - \varepsilon$) than with (T_{FB}^L, T_{SB}^H) if

$$\nu\Omega^H(T_{FB}^L - \varepsilon) + (1-\nu)\Omega^L(T_{FB}^L) - \nu U^H(T_{FB}^L - \varepsilon, T_{FB}^L) - (1-\nu)U^L(T_{FB}^L - \varepsilon, T_{FB}^L) >$$

[35] Since we consider λ^H sufficiently higher than λ^L, we can always choose $T_{FB}^H < T_{FB}^L + 1$ for $\hat{\lambda} < \lambda^L$ and, there always exists ε such that $T_{FB}^L - \varepsilon > T_{FB}^H$.

$$\nu\Omega^H(T_{SB}^H) + (1-\nu)\Omega^L(T_{FB}^L) - (1-\nu)U^L(T_{SB}^H) \text{ or, equivalently}$$

$$\nu[\Omega^H(T_{FB}^L - \varepsilon) - \Omega^H(T_{SB}^H)] >$$

$$\nu U^H(T_{FB}^L - \varepsilon, T_{FB}^L) + (1-\nu)U^L(T_{FB}^L - \varepsilon, T_{FB}^L) - (1-\nu)U^L(T_{SB}^H).$$

Since $\nu U^H(T_{FB}^H, T_{FB}^L) + (1-\nu)U^L(T_{FB}^H, T_{FB}^L) - (1-\nu)U^L(T_{SB}^H) > \nu[\Omega^H(T_{FB}^H) - \Omega^H(T_{SB}^H)]$ by assumption (see step 3a) and $\nu[\Omega^H(T_{FB}^H) - \Omega^H(T_{SB}^H)] > \nu[\Omega^H(T_{FB}^L - \varepsilon) - \Omega^H(T_{SB}^H)]$ because $T_{SB}^H - T_{FB}^H > T_{SB}^H - (T_{FB}^L - \varepsilon)$, the principal is better off with the newly introduced contract ($T^L = T_{FB}^L$ and $T^H = T_{FB}^L - \varepsilon$) than with (T_{FB}^L, T_{SB}^H) if

$$\nu U^H(T_{FB}^H, T_{FB}^L) + (1-\nu)U^L(T_{FB}^H, T_{FB}^L) - (1-\nu)U^L(T_{SB}^H) >$$

$$\nu U^H(T_{FB}^L - \varepsilon, T_{FB}^L) + (1-\nu)U^L(T_{FB}^L - \varepsilon, T_{FB}^L) - (1-\nu)U^L(T_{SB}^H).$$

The above inequality holds for any $\varepsilon < T_{FB}^L - T_{FB}^H$, since we prove later in this Appendix A (see *Sufficient conditions for over/under experimentation*) that for $\overline{\lambda}^H < \lambda^H$ both $\frac{\partial U^H(T^H,T^L)}{\partial T^H} < 0$ and $\frac{\partial U^L(T^H,T^L)}{\partial T^H} < 0$. Therefore, the principal is better off by paying rent to both types rather than to the low type only.

Consider now $0 < \lambda^L < \lambda^H < \min\{\underline{\lambda}^H, \hat{\lambda}\}$. Repeating the same procedure with $T^L = T_{FB}^H - \varepsilon$ and $T^H = T_{FB}^H$ and given that for $\lambda^L < \lambda^H < \min\{\underline{\lambda}^H, \hat{\lambda}\}$ both $\frac{d\zeta(t)}{dt} < 0$ and $T_{FB}^H > T_{FB}^L$ a similar conclusion follows.

Q.E.D.

Proof of Propositions 2 and 3

We first characterize the optimal payment structure given T^L and T^H, x_L, $\{y_t^L\}_{t=1}^{T^L}$, x_H and $\{y_t^H\}_{t=1}^{T^H}$ (Proposition 3), then the optimal length of experimentation, T^L and T^H (Proposition 2).

Denote the expected surplus net of costs for $\theta = H, L$ by $\Omega^\theta(\varpi^\theta) = \beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1 - \lambda^\theta\right)^{t-1} \lambda^\theta [V(q_S) - \underline{c} q_S - \Gamma_t] + \delta^{T^\theta} P_{T^\theta}^\theta [V(q_F) - c_{T^\theta+1}^\theta q_F - \Gamma_{T^\theta}]$. 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 IR constraints given below:

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

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

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

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

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

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

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

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

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

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

Claim: 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 (IRS_t^L) and (IRF_t^L) are binding, but that would violate $(IC^{L,H})$ since $\Delta c_{T^H+1} q_F > 0$. *Q.E.D.*

I. Optimal payment structure, x_L, $\{y_t^L\}_{t=1}^{T^L}$, x_H and $\{y_t^H\}_{t=1}^{T^H}$ (Proof of Proposition 3)

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

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

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

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

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

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 $\frac{\lambda^H}{\lambda^L} > 1 > \frac{P^H_{T^L}}{P^L_{T^L}}$.[36] 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 $(IC^{H,L})$ 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 $1 \le t \le T^H$, which implies that the rent of the low type in this case becomes $\delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q_F$. Replacing x_L in the objective function, the principal's optimization problem is to choose T^H, T^L, $\{y^L_t\}_{t=1}^{T^L}$ to

$$Max\ E_\theta\{\pi^\theta_{FB}(\varpi^\theta) - (1-\upsilon)\delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q_F\} \text{ subject to:}$$

(IRS^L_t) $y^L_t \ge 0$ for $t \le T^L$,

and (IRF_{T^L}) $\delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q_F - \beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1} \lambda^L y^L_t \ge 0$.

---

[36] 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 $\frac{P^H_{T^L}}{P^L_{T^L}} = \frac{(1-\lambda^H)^{\hat{T}^L-1}\lambda^H}{(1-\lambda^L)^{\hat{T}^L-1}\lambda^L}$.

When the $(IC^{H,L})$ constraint is not binding, the claim below shows that there are no restrictions in choosing $\{y_t^L\}_{t=1}^{T^L}$ 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 $\left(x_L, \{y_t^L\}_{t=1}^{T^L}\right)$ such that $\beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1}\lambda^L y_t^L + \delta^{T^L} P_{T^L}^L x^L = \delta^{T^H} P_{T^H}^L \Delta c_{T^H+1} q_F$. Labeling by $\{\alpha_t^L\}_{t=1}^{T^L}, \alpha^L$ the Lagrange multipliers of the constraints associated with (IRS_t^L) for $t \le T^L$, and (IRF_{T^L}) respectively, we have the following claim.

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

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

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

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

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

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

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

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

Case B: The $(IC^{H,L})$ 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 $\psi = P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L$, we can re-write the incentive compatibility constraints as:

$$x^H \delta^{T^H} \psi = \beta_0 \sum_{t=1}^{T^H} \delta^t \left[P_{T^L}^H (1-\lambda^L)^{t-1}\lambda^L - P_{T^L}^L (1-\lambda^H)^{t-1}\lambda^H\right] y_t^H$$

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

$$+P_{T^L}^H q_F \left(\delta^{T^H} P_{T^H}^L \Delta c_{T^H+1} - \delta^{T^L} P_{T^L}^L \Delta c_{T^L+1}\right), \text{ and}$$

$$x^L \delta^{T^L} \psi = \beta_0 \sum_{t=1}^{T^H} \delta^t \left[P_{T^H}^H (1-\lambda^L)^{t-1}\lambda^L - P_{T^H}^L (1-\lambda^H)^{t-1}\lambda^H\right] y_t^H$$

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

$$+P_{T^H}^L q_F \left(\delta^{T^H} P_{T^H}^H \Delta c_{T^H+1} - \delta^{T^L} P_{T^L}^H \Delta c_{T^L+1}\right).$$

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

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

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

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

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

(IRS^θ_t) $y^\theta_t \geq 0$ for $t \leq T^\theta$,

(IRF_{T^θ}) $x^\theta\left(\{y^H_t\}^{T^H}_{t=1}, \{y^L_t\}^{T^L}_{t=1}, T^H, T^L\right) \geq 0$ for $\theta = H, L$.

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

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

$$+\sum_{t=1}^{T^H} \alpha_t^H y_t^H + \sum_{t=1}^{T^L} \alpha_t^L y_t^L + \xi^H x^H\left(\{y_t^H\}_{t=1}^{T^H}, \{y_t^L\}_{t=1}^{T^L}, T^H, T^L\right) + \xi^L x^L\left(\{y_t^H\}_{t=1}^{T^H}, \{y_t^L\}_{t=1}^{T^L}, T^H, T^L\right).$$

We 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:

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

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

where

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

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

Next, we show that the principal will not commit to reward success in two different periods for either type (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

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

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

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

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

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

Following similar steps, 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. $\frac{\xi^L}{\delta^{T^L}} \ne \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$ 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 $\left(IRF_{T^H}^H\right)$ and $\left(IRF_{T^L}^L\right)$ constraints are not binding.

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

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

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

Since $f_1(t,T^H)$ is strictly positive for all $t < \hat{T}^H$ from $P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\right] = -\frac{\alpha_t^H\psi}{\beta_0\delta^t}$ it must be that $\alpha_t^H > 0$ for all $t < \hat{T}^H$ and $\psi < 0$. In addition, since $f_2(t,T^L)$ is strictly negative for $t < \hat{T}^L$ from $P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L\right] = -\frac{\alpha_t^L\psi}{\beta_0\delta^t}$ it must be that that $\alpha_t^L > 0$ for $t < \hat{T}^L$ and $\psi > 0$, which leads to a contradiction[37].

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

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

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

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

---

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

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 $\frac{\xi^L}{\delta^{T^L}} = \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$[38]. 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 $1 \le t \le T^H$.

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 $T^H \le \hat{T}^H$ and there can be one of two sub-cases:[39] (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): $T^H \le \hat{T}^H$, $\psi > 0$, $\frac{\xi^L}{\delta^{T^L}} > \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$, $\xi^H = 0$, $\alpha_t^H > 0$ for $1 \le t \le T^H$.

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

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

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

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

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

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

For $j > \hat{T}^H$ we have $f_1(j, T^H) < 0$ and it follows that $P_{T^L}^L f_2(t, T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L\right] + \frac{\xi^L}{\delta^{T^L}} P_{T^H}^H f_1(t, T^H) < -\psi\left((1-\upsilon)(1-\lambda^L)^{t-1}\lambda^L + \upsilon(1-\lambda^H)^{t-1}\lambda^H\right) < 0$, 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): $T^H \le \hat{T}^H$, $\psi < 0$, $\frac{\xi^L}{\delta^{T^L}} < \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$, $\xi^H = 0$, $\alpha_t^H > 0$ for $1 \le t \le T^H$.

As in the previous case, from Lemma 3 it follows that there exists only one time period $1 \le s \le T^L$ such that $y_s^L > 0$ ($\alpha_s^L = 0$). This implies that $P_{T^L}^L f_2(s, T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L\right] +$

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

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

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

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

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

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

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

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

For both cases we just considered, we have

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

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

Note that Case B.2 is possible only if $\delta^{T^H} P^L_{T^H}\Delta c_{T^H+1} q_F - \delta^{T^L} P^L_{T^L}\Delta c_{T^L+1} q_F > 0$[40]. This fact together with $x^H \ge 0$ implies that $\psi > 0$. Since $f_1(1,T^H) > 0$, $x^L = 0$ is possible only if $\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q_F - \delta^{T^L} P^H_{T^L}\Delta c_{T^L+1} q_F < 0$. However, $\delta^{T^H} P^L_{T^H}\Delta c_{T^H+1} q^H\left(c^H_{T^H+1}\right) >$ $\delta^{T^L} P^L_{T^L}\Delta c_{T^L+1} q_F$ implies that $\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q_F > \delta^{T^L} \frac{P^H_{T^H}}{P^L_{T^H}} P^L_{T^L}\Delta c_{T^L+1} q_F$. 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 $\delta^{T^H} P^H_{T^H}\Delta c_{T^H+1} q_F > \delta^{T^L} P^H_{T^L}\Delta c_{T^L+1} q_F$, which implies $x^L > 0$ and we have a contradiction. Thus, $\xi^H > 0$ and the high type gets rent only after success ($x^H = 0$). *Q.E.D.*

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

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

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

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

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

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

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

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 $1 \le t \le T^L$ and we conclude that $T^L \le \hat{T}^L$ and there can be one of two sub-cases:[42] (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): $T^L \le \hat{T}^L$, $\psi > 0$, $\frac{\xi^H}{\delta^{T^H}} < \upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L$, $\xi^L = 0$, $\alpha_t^L > 0$ for $1 \le t \le T^L$.

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

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

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

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

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

---

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

[42] 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 $t \le T^L$..

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

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

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

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

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

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

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

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

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

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

$y_1^H = q_F \frac{P_{T^L}^H\left(\delta^{T^L} P_{T^L}^L \Delta c_{T^L+1} - \delta^{T^H} P_{T^H}^L \Delta c_{T^H+1}\right)}{\beta_0 \delta P_{T^L}^L f_2(1,T^L)} \geq 0$; $x^L = \frac{\delta^{T^L}\lambda^L P_{T^L}^H \Delta c_{T^L+1} - \delta^{T^H}\lambda^H P_{T^H}^L \Delta c_{T^H+1}}{\delta^{T^L} P_{T^L}^L f_2(1,T^L)} > 0.$

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

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

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

Claim: 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 $1 \leq j \leq T^L$ such that $y_j^L > 0$ ($\alpha_j^L = 0$) it follows that

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $1 \leq t \neq s \leq T^H$ 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

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

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

Since Case B.2 is possible only if $\delta^{T^H}P^L_{T^H}\Delta c_{T^H+1}q^H\left(c^H_{T^H+1}\right) - \delta^{T^L}P^L_{T^L}\Delta c_{T^L+1}q^L\left(c^L_{T^L+1}\right) > 0$[43], 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:

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

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

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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Combining the two equations together, $\alpha^L P^H_{T^H}\big(f_1(k,T^H)f_2(m,T^L) - f_1(m,T^H)f_2(k,T^L)\big)$ $+\upsilon\lambda^H\Big((1-\lambda^H)^{k-1}f_2(m,T^L) - (1-\lambda^H)^{m-1}f_2(k,T^L)\Big) = 0$, which can be rewritten as follows[44]:

$$\frac{P^H_{T^L}}{P^L_{T^L}}\lambda^L\big((1-\lambda^H)^{k-1}(1-\lambda^L)^{m-1} - (1-\lambda^H)^{m-1}(1-\lambda^L)^{k-1}\big) = 0,$$

[44] After some algebra, one could verify that $f_1(k,T^H)f_2(m,T^L) - f_1(m,T^H)f_2(k,T^L)$ $= \psi\frac{\lambda^H\lambda^L}{P^H_{T^H}P^L_{T^L}}[(1-\lambda^H)^{m-1}(1-\lambda^L)^{k-1} - (1-\lambda^L)^{m-1}(1-\lambda^H)^{k-1}].$

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

In the same way, there exists *at most* one time period $1 \leq j \leq T^L$ 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 $1 \leq s \leq T^H$ 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 $1 \leq t \neq s \leq T^H$.

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

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

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 $1 \leq t \neq s \leq T^H$ 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 $1 \leq t \neq s \leq T^H$ and it must be that $s = 1$. In a similar way, for $1 \leq j \leq T^L$ such that $y_j^L > 0$ it must be that either $j = 1$ or $j = T^L > \hat{T}^L$. *Q.E.D.*

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

II. Optimal length of experimentation

(Proposition 2)

Since T^L and T^H affect the information rents, U^L and U^H, there will be a distortion in the duration of the experimentation stage for both types:

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

The exact values of U^H and U^L depend 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, by Claim A.1, the low type's rent $\delta^{T^H} P_{T^H}^L \Delta c_{T^H+1} q_F$ is not affected by T^L. Therefore, the F.O.C. with respect to T^L is identical to that under first best: $\frac{\partial \mathcal{L}}{\partial T^L} = \frac{\partial E_\theta\, \Omega^\theta(\varpi^\theta)}{\partial T^L} = 0$, or, equivalently, $T_{SB}^L = T_{FB}^L$ when $(IC^{H,L})$ is not binding. However, since the low type's information rent depends on T^H, there will be a distortion in the duration of the experimentation stage for the high type:

$$\frac{\partial \mathcal{L}}{\partial T^H} = \frac{\partial\left(E_\theta\, \Omega^\theta(\varpi^\theta) - (1-\upsilon)\, \delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q_F\right)}{\partial T^H} = 0.$$

Since the informational rent of the low-type agent, $\delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q_F$, is non-monotonic in T^H, it is possible, in general, to have $T^H_{SB} > T^H_{FB}$ or $T^H_{SB} < T^H_{FB}$.

In Case B, the exact values of U^H and U^L depend on whether $T^L < \hat{T}^L$ (Lemma 4) or $T^L > \hat{T}^L$ (Lemma 5), but in each case $U^L > 0$ and $U^H \geq 0$. It is possible, in general, to have $T^H_{SB} > T^H_{FB}$ or $T^H_{SB} < T^H_{FB}$ and $T^L_{SB} > T^L_{FB}$ or $T^L_{SB} < T^L_{FB}$.

We provide sufficient conditions for over and under experimentation next.

Sufficient conditions for over/under experimentation

***Case A*:** If $(IC^{H,L})$ is not binding, $T^H_{FB} < T^H_{SB}$ if $\lambda^H > \overline{\lambda}^H$ and $T^H_{FB} > T^H_{SB}$ if $\lambda^H < \underline{\lambda}^H$.

Proof: If $(IC^{H,L})$ is not binding, the rent to the low type is $U^L = \delta^{T^H} P^L_{T^H} \Delta c_{T^H+1} q_F$. Given that $\Delta c_t = (\overline{c} - \underline{c})(\beta^L_t - \beta^H_t)$, the rent becomes $U^L = \delta^{T^H} P^L_{T^H}\left(\beta^L_{T^H+1} - \beta^H_{T^H+1}\right)(\overline{c} - \underline{c}) q_F$.

Recall the function $\zeta(t) \equiv \delta^t P^L_t (\beta^L_{t+1} - \beta^H_{t+1})$ from the proof for Sufficient Conditions for $(IC^{H,L})$ to be binding. Rent to the low type can be rewritten as $U^L = \zeta(T^H)(\overline{c} - \underline{c}) q_F$. In proving sufficient condition for $(IC^{H,L})$ to be binding we proved that for any t, $\frac{d\zeta(t)}{dt} < 0$ for $\lambda^H > \overline{\lambda}^H$, and $\frac{d\zeta(t)}{dt} > 0$ for $\lambda^H < \underline{\lambda}^H$. Therefore, if $\lambda^H > \overline{\lambda}^H$ $(< \underline{\lambda}^H)$ then $\frac{dU^L}{dT^H} < 0$ (> 0) and we have over (under) experimentation. *Q.E.D.*

***Case B*:** If both $(IC^{H,L})$ and $(IC^{L,H})$ are binding, $T^H_{FB} < T^H_{SB}$ if $\lambda^H > \overline{\lambda}^H$ and $T^H_{FB} > T^H_{SB}$ if $\lambda^H < \underline{\lambda}^H$, and $T^L_{FB} < T^L_{SB}$ if $\lambda^H < \underline{\underline{\lambda}}^H$.

Proof: If both $(IC^{H,L})$ and $(IC^{L,H})$ are binding, rent paid to the low and high type is

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

$$U^H = q_F \frac{\delta^{T^H} P^L_{T^H}\Delta c_{T^H+1}(1-\lambda^H)^{T^L-1}\lambda^H - \delta^{T^L}\Delta c_{T^L+1}(1-\lambda^L)^{T^L-1}\lambda^L}{\lambda^L\left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)}, \text{ respectively.}$$

Given that $\Delta c_t = (\overline{c} - \underline{c})(\beta^L_t - \beta^H_t)$, we have

$$U^L = q_F(\overline{c} - \underline{c}) \frac{\left(\delta^{T^H}\lambda^H P^L_{T^H}\left(\beta^L_{T^H+1} - \beta^H_{T^H+1}\right) - \delta^{T^L}\lambda^L P^H_{T^L}\left(\beta^L_{T^L+1} - \beta^H_{T^L+1}\right)\right)}{\lambda^H\left(1 - \left(\frac{1-\lambda^H}{1-\lambda^L}\right)^{T^L-1}\right)}$$

$$\text{and } U^H = q_F(\overline{c} - \underline{c}) \frac{\delta^{T^H} P^L_{T^H}\left(\beta^L_{T^H+1} - \beta^H_{T^H+1}\right)\left(\frac{1-\lambda^H}{1-\lambda^L}\right)^{T^L-1}\lambda^H - \delta^{T^L} P^H_{T^L}\Delta c_{T^L+1}\lambda^L}{\lambda^L\left(1 - \left(\frac{1-\lambda^H}{1-\lambda^L}\right)^{T^L-1}\right)}.$$

Recall function $\zeta(t) \equiv \delta^t P^L_t(\beta^L_{t+1} - \beta^H_{t+1})$. Then U^L and U^H can be rewritten as

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

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

Consider first T^H. Both U^L and U^H are increasing in $\zeta(T^H)$. In proving sufficient condition for $(IC^{H,L})$ to be binding we proved that $\frac{d\zeta(T^H)}{dT^H} < 0$ for $\lambda^H > \overline{\lambda}^H$, and $\frac{d\zeta(T^H)}{dT^H} < 0$ for $\lambda^H < \underline{\lambda}^H$. Therefore, it is optimal to let the high type *over* experiment ($T^H_{FB} < T^H_{SB}$) if $\lambda^H > \overline{\lambda}^H$ and *under* experiment ($T^H_{FB} > T^H_{SB}$) if $\lambda^H < \underline{\lambda}^H$.

Consider now T^L. Since $\left(\frac{1-\lambda^H}{1-\lambda^L}\right)^{T^L-1}$ is decreasing in T^L, both U^L and U^H will be decreasing in T^L if $\frac{\zeta(T^L) P^H_{T^L}}{P^L_{T^L}}$ is increasing in T^L. We next prove that $\frac{\zeta(T^L) P^H_{T^L}}{P^L_{T^L}}$ is increasing in T^L for small values of λ^H.

To simplify, $\frac{\zeta(T^L) P^H_{T^L}}{P^L_{T^L}} = \frac{\delta^{T^L} P^L_{T^L}\left(\beta^L_{T^L+1} - \beta^H_{T^L+1}\right) P^H_{T^L}}{P^L_{T^L}} = \delta^{T^L} \frac{\beta_0(1-\beta_0)\left((1-\lambda^L)^{T^L} - (1-\lambda^H)^{T^L}\right)}{P^L_{T^L}}$.

$$\frac{d\left[\delta^{T^L}\frac{\left((1-\lambda^L)^{T^L} - (1-\lambda^H)^{T^L}\right)}{P^L_{T^L}}\right]}{dT^L} =$$

$$\delta^{T^L} \frac{\left((1-\lambda^L)^{T^L} ln(1-\lambda^L) - (1-\lambda^H)^{T^L} ln(1-\lambda^H)\right) P^L_{T^L} - \beta_0 (1-\lambda^L)^{T^L} ln(1-\lambda^L)\left((1-\lambda^L)^{T^L} - (1-\lambda^H)^{T^L}\right)}{\left(P^L_{T^L}\right)^2}$$

$$+\delta^{T^L} \ln \delta \frac{P^L_{T^L}\left((1-\lambda^L)^{T^L} - (1-\lambda^H)^{T^L}\right)}{\left(P^L_{T^L}\right)^2}$$

$$= \delta^{T^L} \frac{\ln(1-\lambda^L)(1-\lambda^L)^{T^L} P^H_{T^L} - (1-\lambda^H)^{T^L} \ln(1-\lambda^H) P^L_{T^L}}{\left(P^L_{T^L}\right)^2} + \delta^{T^L} \ln \delta \frac{P^L_{T^L}\left((1-\lambda^L)^{T^L} - (1-\lambda^H)^{T^L}\right)}{\left(P^L_{T^L}\right)^2}$$

$$= \delta^{T^L} \frac{(1-\lambda^L)^{T^L}\left[P^H_{T^L} \ln(1-\lambda^L) + P^L_{T^L} \ln \delta\right] - (1-\lambda^H)^{T^L} P^L_{T^L} \ln[\delta(1-\lambda^H)]}{\left(P^L_{T^L}\right)^2}.$$

$\delta^{T^L} \frac{\left((1-\lambda^L)^{T^L} - (1-\lambda^H)^{T^L}\right)}{P^L_{T^L}}$ increases with T^L if and only if $\kappa(\lambda^H) > 0$,

$$\kappa(\lambda^H) = (1-\lambda^L)^{T^L}\left[P^H_{T^L} \ln(1-\lambda^L) + P^L_{T^L} \ln \delta\right] - (1-\lambda^H)^{T^L} P^L_{T^L} \ln[\delta(1-\lambda^H)].$$

We prove next that $\kappa(\lambda^H) > 0$ for any t if λ^H is sufficiently low.

Since both $(1-\lambda^H)^{T^L} P^L_{T^L}$ and $(1-\lambda^L)^{T^L}\left[P^H_{T^L} \ln(1-\lambda^L) + P^L_{T^L} \ln \delta\right]$ are increasing in t, we have $\kappa(\lambda^H) > (1-\lambda^L)[P^H_1 \ln(1-\lambda^L) + P^L_1 \ln \delta] - (1-\lambda^H)^{\overline{T}} P^L_{\overline{T}} \ln[\delta(1-\lambda^H)]$,

where $\overline{T} = \max\{T^H, T^L\}$.

Next, $(1-\lambda^L)[P^H_1 \ln(1-\lambda^L) + P^L_1 \ln \delta] > (1-\lambda^L) P^L_1 \ln[\delta(1-\lambda^L)]$ and $\frac{\ln[\delta(1-\lambda^H)]}{\ln[\delta(1-\lambda^L)]} > 1$. Therefore, $\frac{(1-\lambda^L)P^L_1}{(1-\lambda^H)^{\overline{T}} P^L_{\overline{T}}} < 1 \Rightarrow \kappa(\lambda^H) > 0$.

Rearranging the above, we have $\kappa(\lambda^H) > 0$ for any t if $\lambda^H < 1 - \left(\frac{(1-\lambda^L)P^L_1}{P^L_{\overline{T}}}\right)^{\frac{1}{\overline{T}}}$.

Denote

$$\underline{\underline{\lambda}}^H \equiv 1 - \left(\frac{(1-\lambda^L)P^L_1}{P^L_{\overline{T}}}\right)^{\frac{1}{\overline{T}}}.$$

Therefore, for any t, $\kappa(\lambda^H) > 0$ if $\lambda^H < \underline{\underline{\lambda}}^H$ and, consequently, both U^L and U^H are decreasing in T^L. Consequently, it is optimal to let the low type *over* experiment ($T^L_{FB} < T^L_{SB}$) if $\lambda^H < \underline{\underline{\lambda}}^H$. *Q.E.D.*