Optimal Contract for

Experimentation and Production

Fahad Khalil, Jacques Lawarrée and Alexander Rodivilov

Department of Economics
University of Washington

June 9, 2017

Abstract: Before embarking on a project, a principal must often rely on an agent to learn about its profitability. These situations are conveniently modeled as two-armed bandit problems highlighting a trade-off between learning (experimentation) and production (exploitation). We derive the optimal contract for both experimentation and production when the agent has private information about his skill or efficiency in experimentation. Private information in the experimentation stage can generate asymmetric information between the principal and agent about the expected profitability of production. The degree of asymmetric information is endogenously determined by the length of the experimentation stage. An optimal contract uses the timing of payments, the length of experimentation, and the output to screen the agent. To induce revelation during the experimentation, the principal utilizes the stochastic structure of asymmetric learning by agents with different skills. Both upward and downward incentive constraints can be binding. The relative probabilities of success and failure between agents of different skills imply that agents are rewarded for success or failure at the boundaries of the experimentation stages: an efficient agent is rewarded for early success and an inefficient agent for late success. When the experimentation stage is short, we show that rewarding failure may be optimal. The optimal contract may also feature excessive experimentation, and over- or under-production.

Keywords: Information gathering, optimal contracts, strategic experimentation.

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

1. Introduction

Before embarking on a project, it is important to learn about its profitability to determine how much resource to allocate to the project. When an oil company explores new areas for oil fields, it performs seismic surveys and exploration drills to figure out the amount of oil it can expect from the areas.[1] While the oil company experiments with different potential sites, it also diverts resources and delays the production of oil. This creates a trade-off between experimentation and production (exploitation) analogous to a two-armed bandit problem.[2]

An additional complexity arises if the experiments are performed by a manager (agent) who privately knows his skill or efficiency in experimentation. The experimentation process itself can then create asymmetric information about the profitability of the project. Exploration drills will demonstrate the profitability of the oil field. If the agent is not efficient at experimenting, a poor result from the exploration only provides weak evidence of low profitability. However, if the owner (principal) is misled into believing that the agent is highly efficient, she becomes more pessimistic than the agent. A new trade-off appears for the principal. More experiments may provide more information about the profitability of the well but can also increase asymmetric information about the expected profitability. Because of this asymmetry of information, when production ultimately starts, the principal may not allocate the right amount of resources to the exploitation of the field. Another example is a standard procurement problem where a government agency (principal) finds it important to motivate a supplier (agent) to acquire planning information before executing the project.

In this paper, we derive the optimal contract for an agent who conducts both experimentation and production. At the outset, the principal and agent are symmetrically informed that production cost can be high or low. Before production takes place, the principal asks the agent to gather additional information about the actual production cost. This is the experimentation stage. For most of the paper, we assume that the information gathering takes the form of looking for good news, i.e., whether cost of production is low.[3] When

[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.

[2] See Bolton and Harris (1999), Keller, Rady, and Cripps (2005), or Bergemann and Välimäki (2008).

[3] We also show that our key results extend to the case of bad news.

experimentation succeeds, it is publicly revealed that the cost is low, and production occurs under symmetric information.[4] We say that experimentation fails when the agent does not learn the actual cost by the end of the experimentation stage. Then, production occurs under expected cost, which can be different for the principal and the agent. This will lead to a rent for the agent as we explain next.

The agent is privately informed about his efficiency in experimentation, which is his type. He can be either efficient or inefficient. An efficient agent has a greater probability of finding low cost (when cost is indeed low). To see how experimentation can endogenously create asymmetric information between the principal and the agent, consider the case when an inefficient agent claims to be efficient, and experimentation fails. The principal is now relatively more pessimistic about the expected cost than the inefficient agent. Since the agent knows that he is not very efficient at experimenting, his failure is not so informative about cost being high. Thus, the lying inefficient agent will have a lower expected cost of production compared to the principal, who will overcompensate him in the production stage (mistakenly believing he is efficient). To deter lying, the principal must pay a rent to the inefficient agent.

A key contribution of our model is to study how incentives for production affects incentives to experiment and, conversely, how the asymmetric information generated in the experimentation stage impacts production. At the end of the experimentation stage, there is a non-trivial decision regarding the scale of output. This decision depends on what is learned during experimentation. Relative to the nascent literature on incentives for experimentation, the novelty of our approach is to study optimal contracts for *both* experimentation and production. As we will see, much can be learned even when experimentation fails, and this information would be lost in a model without a production stage. If the principal asks the agent to experiment longer, there is a greater chance to succeed and fine-tune the size of the project. However, experimentation is costly since it endogenously creates asymmetric information and production has to be postponed.

[4] We study the case where success can be hidden in Section 3.

The asymmetric information created by experimentation impacts the optimal contract. In particular, it affects (i) the length of the experimentation period, (ii) the information rent for the agent and the timing of its payment, and (iii) the output. We discuss them in turn.

(i) First, we show that the optimal length of the experimentation period could be longer than the first-best length while most models of experimentation find under experimentation relative to the first best. Under asymmetric information, the difference in expected costs between types, the driving force for the rent, determines the distortion in the length of the experimentation stage. When the agent lies about his type, his expected cost after failure is different from the principal's expected cost. As a result, the agent's informational rent is positively related to the difference in the expected cost of the two different types. We show that this difference is non-monotonic in time because the two types learn at different speeds. Thus, the principal may benefit from increasing the length of the experimentation stage since it may lower the difference in expected costs. Using time for screening the types turns out to be complex as it balances two countervailing forces: a more efficient experimenter learns good news more quickly but also becomes pessimistic more quickly after successive failures. Another consequence is that the optimal length of the experimentation phase varies for each type.

(ii) Second, it is possible that the efficient agent also gets a rent. The timing of payment is used to limit the rents paid to both types. The rent given to the inefficient agent can make it attractive for the efficient type to misreport. To illustrate this incentive, suppose the principal rewards early success for the inefficient agent. Such a scheme may look attractive for the efficient agent because he is more likely to succeed early in the experimentation stage. But an alternative scheme, such as rewarding failure by the inefficient agent, may also become attractive for the efficient agent during a long experimentation stage because successive failures convince the efficient agent that the project is high cost and experimentation is likely to fail.

As indicated by the above, the relative likelihoods of success and failure play a crucial role in determining the optimal timing of the payments. When the principal must pay a rent to the efficient agent, it will be as a reward for early success since he is more likely to succeed early. If the principal wants to reward an inefficient agent for success, it has to be late in the experimentation stage.

Remarkably, we also show that it may be optimal for the principal to reward an agent after *failure*.[5] When the optimal length of the experimentation period is short, the relative probability of failure for an inefficient agent is greater than his relative probability of success. Thus, rewarding the inefficient agent after failure becomes a useful tool to screen the types. One may wonder if this result depends on the assumption that the agent cannot hide success, and we show in an extension that it does not.[6]

Because we combine experimentation and production in our model, the efficient agent faces a gamble if he pretends to be inefficient. On the one hand, he can collect the rent offered to the inefficient type. On the other, if experimentation fails, he will be under-compensated at the production stage as he is relatively more pessimistic compared to the principal. These two parts are related to the difference in the expected costs after failure. Because the duration of experimentation can be different for each type, the gamble can be positive or negative.

To highlight the role of the different lengths of the experimentation stage for different types, we consider, in an extension, a case when the length of the experimentation stage must be identical for both types. Then, the difference in expected cost is identical for each type and the gamble is negative for the efficient type if he misreports. We find that the efficient type can no longer command a rent. Thus, it is the principal's choice to have different lengths of experimentation stage for each type that leads to a rent for the efficient type.

(iii) Third, unless experimentation succeeds, the principal will use the choice of output to screen the two types during the production stage. Since the inefficient agent always gets a rent, we expect and indeed find that the output of the efficient agent is distorted downward as in a standard static second best contract. However, when the efficient agent also commands a rent, the output of the inefficient agent is distorted *upward*. A higher output for the inefficient agent makes it more costly for the efficient agent to lie since a lying efficient agent, being relatively more pessimistic after failure, will be under-compensated in the production stage.

---

[5] See Manso (2011) for a similar result in a model with moral hazard.

[6] If the agent could hide success, he can guarantee apparent failure in the experimentation stage. In such a case, preventing the agent from hiding success introduces additional ex post moral hazard constraints that impose additional costs on the principal, but rewarding failure can still be optimal due to the same argument based on relative probabilities of success and failure.

Finally, we also consider an extension where the agent looks for bad news during the experimentation stage. Now it is the efficient agent who always gets rent as he is more pessimistic after the same amount of failures. We find results analogous to the case of good news. In particular, the timing of the payment is reminiscent of that from learning good news. However, distortions in output are reversed since now it is the efficient agent who always gets rent.

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.[7] It is typical in this literature to consider one shot models, where an agent exerts effort that increases the precision of a signal relevant to production. By modeling this effort as experimentation, we introduce a dynamic learning aspect, and especially the possibility of learning with asymmetric speeds. 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 there can be over or under-experimentation.

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 do not consider adverse selection.[8] Recent exceptions that introduce adverse selection are Gomes, Gottlieb and Maestri (2016) and Halac, Kartik and Liu (2016).[9] In 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

[7] 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.
[8] See also Horner and Samuelson (2013).
[9] See also and Gerardi and Maestri (2012) for another model where the agent is privately informed about the quality (prior probability) of the project.

about expertise in the experimentation stage to study how asymmetric learning by the high and inefficient agents affects the bonus that needs to be paid to induce the agent to work.[10]

We add to the literature by explicitly modeling a production stage following the experimentation stage, such that contracts provide incentive for production as well as experimentation. Furthermore, asymmetric information about skill in experimentation implies that production can occur under asymmetric information. This latter aspect would be missing in a model of incentives for experimentation without a production stage. Unlike the rest of the literature, we find that over-experimentation relative to the first best can be optimal because the difference in expected production cost, the source of information rent, can decrease in time after a succession of failures. Also, the agents are rewarded for success or failure at the boundaries of the experimentation stages: an efficient agent is rewarded for early success and an inefficient agent for late success or failure.

2. The Model (Learning good news)

A principal hires an agent to implement a project of a variable size. Both the principal and agent are risk neutral and have a common discount factor $\delta \in (0,1]$. It is common knowledge that the marginal cost 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* (exploitation), the agent can gather information regarding the production cost, which is called 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. In each period t , experimentation costs $\gamma > 0$, and we assume that this cost γ is paid by the principal at the end of each period. Thus, there is no moral hazard aspect in this model. We assume that it is always optimal to experiment at least once.

[10] 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.

In the base model, we also assume that information gathering takes the form of looking for good news (see section 5 for the case of bad news). If the cost is actually 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.[11] The experimentation stage then stops. If the agent fails to learn that the cost is low in a period $t < T$, the agent continues to experiment, but both the agent and the principal become more *pessimistic* about the likelihood of the cost being low. We say that experimentation has failed if the agent fails to learn that cost is low in all T periods.

We assume that the agent is privately informed about his experimentation skill or efficiency represented by λ. Therefore, the principal faces an adverse selection problem. As we will see next, this implies that 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 $0 < \lambda^L < \lambda^H < 1$.[12] If experimentation fails to reveal low cost in a period, agents with different types form different beliefs about the expected cost of the project. We denote by β_t^θ the updated belief of a θ-type agent that the cost is actually low at the beginning of period t given $t-1$ failures. For period $t > 1$, we have $\beta_t^\theta = \frac{\beta_{t-1}^\theta(1-\lambda^\theta)}{\beta_{t-1}^\theta(1-\lambda^\theta)+(1-\beta_{t-1}^\theta)}$, which in terms of β_0 is

$$\beta_t^\theta = \frac{\beta_0 (1-\lambda^\theta)^{t-1}}{\beta_0 (1-\lambda^\theta)^{t-1} + (1-\beta_0)}.$$

The θ-type agent's expected cost at the beginning of period t is then given by:

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

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

[11] We assume that the agent cannot hide the evidence of the cost being low. We will revisit this assumption in Section 3.2 below.

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

experimentation, the expected cost is higher as both c_t^H and c_t^L approach $\bar{c}$. An example of how the expected cost c_t^θ converges to $\bar{c}$ for each type is presented in Figure 1 below.

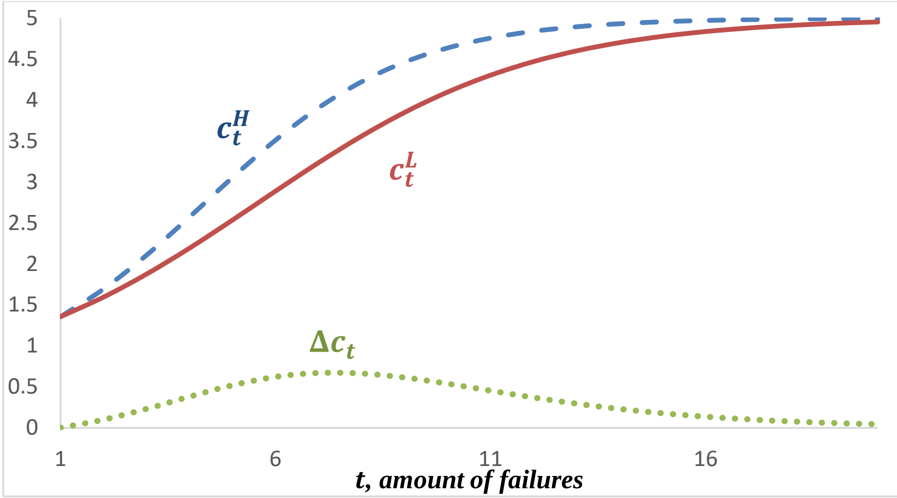

Figure 1. Expected cost with $\lambda^H = 0.35$, $\lambda^L = 0.2$, $\beta_0 = 0.7$, $\underline{c} = 0.5$, $\bar{c} = 5$.

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. This is due to the fact the two types learn at different speeds.[13]

The production (exploitation) stage

After the experimentation stage ends, production takes place in the production stage. The principal's value of the project is $V(q)$, where $q > 0$ is the size of the project. The function $V(\cdot)$ is strictly increasing, strictly concave, twice differentiable on $(0, +\infty)$, and satisfies the Inada conditions. 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, experimentation stops and production takes place based on $c = \underline{c}$. If experimentation fails, production occurs based on the expected cost in period T.[14]

The contract

Before the experimentation stage takes place, the principal offers the agent a menu of dynamic contracts. Relying on the revelation principle, we use a direct truthful mechanism,

---

[13] There exists a unique time period t_Δ such that Δc_t achieves the highest value at this time period, where $t_\Delta = arg \max_{1 \le t \le T} \frac{(1-\lambda^L)^t - (1-\lambda^H)^t}{\left(1-\beta_0+\beta_0(1-\lambda^H)^t\right)\left(1-\beta_0+\beta_0(1-\lambda^L)^t\right)}$.

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

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 $c^{\hat{\theta}}_{T^{\hat{\theta}}}$.[15] A contract is defined formally by

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

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

An agent of type θ, announcing his type as $\hat{\theta}$, receives expected utility $U^{\theta}(\varpi^{\hat{\theta}})$ at time zero from a contract $\varpi^{\hat{\theta}}$:

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

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

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 $\left(1-\lambda^{\theta}\right)^{t-1}\lambda^{\theta}$. If the agent succeeds, he will produce $q_t^{\hat{\theta}}(\underline{c})$ and will be paid $w_t^{\hat{\theta}}(\underline{c})$ by the principal. In addition, it is possible that experimentation fails. This is the case either if the cost is actually 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 $\beta_0(1-\lambda^{\theta})^{T^{\hat{\theta}}}$. In this case, the agent produces $q^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}}\right)$ based on expected cost and is paid $w^{\hat{\theta}}\left(c^{\hat{\theta}}_{T^{\hat{\theta}}}\right)$.

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

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

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

We also assume that the agent must be paid his expected production costs whether experimentation succeeds or fails. Therefore the individual rationality constraints have to be satisfied *ex post* (i.e., after experimentation):[16]

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

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

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

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

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

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:

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

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 stage starts with output and transfers as specified in the contract.

[16] We discuss later the case of ex ante participation constraints.

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.

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

subject to

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

$(IRF_{T^\theta}^\theta)$ $\qquad w^\theta\left(c_{T^\theta}^\theta\right) - c_{T^\theta}^\theta q^\theta\left(c_{T^\theta}^\theta\right) \geq 0.$

The individual rationality constraints are binding. If the agent succeeds, the efficient output will be produced such that $V'\left(q_{t^\theta}^\theta(\underline{c})\right) = \underline{c}$ for any t^θ and the transfers cover the actual cost with no rent given to the agent. In case the agent fails, the efficient output based on the current *expected* cost, such that $V'\left(q^\theta\left(c_{t^\theta}^\theta\right)\right) = c_{t^\theta}^\theta$ for any t^θ. The transfer covers the expected cost and no expected rent is given to the agent.

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

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

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. This non-monotonicity is a result of two countervailing forces. 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 $\beta_0\left(1-\lambda^\theta\right)^{t-1}\lambda^\theta$, 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.

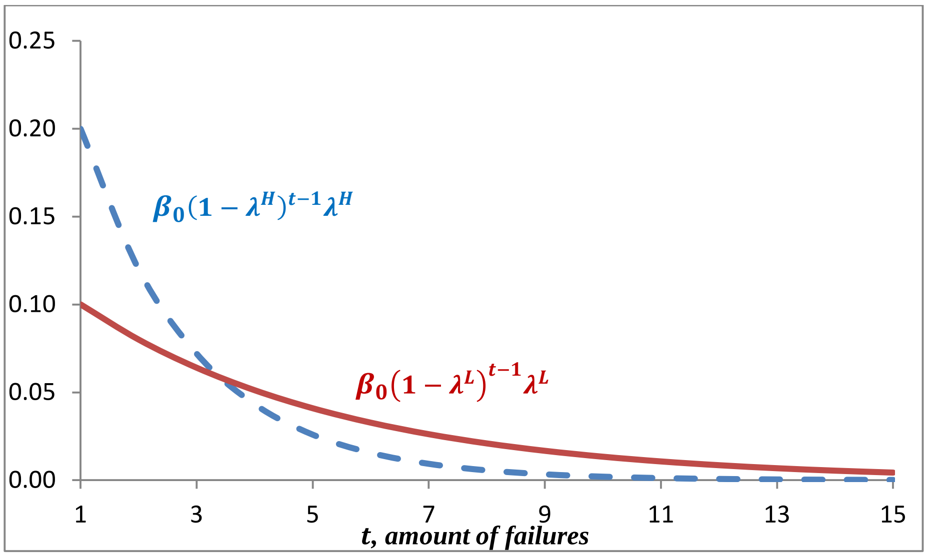

Figure 2. Probability of success with $\lambda^H = 0.4$, $\lambda^L = 0.2$, $\beta_0 = 0.5$.

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, $T_{FB}^H = 4 > T_{FB}^L = 3$.

2.2 Asymmetric information

Assume now that the agent privately knows this type. To understand the role of beliefs in generating rent in the production stage, we start with a benchmark without an experimentation stage but with asymmetric information about expected cost of production. The principal can only screen the agents with the output and payments. We obtain a standard second best contract, where the hidden parameter is the *expected* marginal cost (e.g., Baron-Myerson (1982), Laffont-

Tirole (1986)). Suppose in this case, a type θ agent's belief is denoted by β^{θ}, which implies that the expected cost at the production stage is $c^{\theta} = \beta^{\theta}\underline{c} + (1-\beta^{\theta})\overline{c}$. Suppose that the high-type is more pessimistic than the low type about the cost being low: $\beta^{H} < \beta^{L}$. This implies that $c^{H} > c^{L}$. As a result, the principal's optimization problem now is:

$$\max_{q(c^H),w(c^H),q(c^L),w(c^L)} \upsilon\left[V\left(q(c^H)\right) - w(c^H)\right] + (1-\upsilon)\left[V\left(q(c^L)\right) - w(c^L)\right]$$

(IR^{H}) $w(c^{H}) - c^{H}q(c^{H}) \geq 0$,

(IR^{L}) $w(c^{L}) - c^{L}q(c^{L}) \geq 0$,

$(IC^{H,L})$ $w(c^{H}) - c^{H}q(c^{H}) \geq w(c^{L}) - c^{H}q(c^{L})$,

$(IC^{L,H})$ $w(c^{L}) - c^{L}q(c^{L}) \geq w(c^{H}) - c^{L}q(c^{H})$.

It can be easily shown that the optimal contract resembles a standard second-best contract with adverse selection. In particular, (IR^{H}) and $(IC^{L,H})$ constraints are binding, the low type gets a positive informational rent and produces the first-best output: $V'\left(q(c^{L})\right) = c^{L}$. The high type gets zero rent and his output is distorted as follows: $V'\left(q(c^{H})\right) = c^{H} + \frac{(1-\upsilon)}{\upsilon}(c^{H} - c^{L})$. As in a standard adverse selection model $q^{SB}(c^{H}) < q^{FB}(c^{H}) < q^{FB}(c^{L}) = q^{SB}(c^{L})$.

We now return to our main case where an experimentation stage precedes production. Recall that 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. If an agent experiences success before the termination date, T^{θ}, the true cost $c = \underline{c}$ is revealed.

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:

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

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

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

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

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

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

Note that, if the principal only had to satisfy an ex ante participation constraint $U^\theta(\varpi^\theta) \geq 0$, she could use the fact that the high type is relatively more likely to succeed (conditional on $c = \underline{c}$) to screen the agent without distorting the duration of the experimentation stage. In other words, since success during the experimentation stage is a random event that is correlated with the agent's type, we can apply well-known ideas from mechanisms à la Crémer-McLean (1985) that says the principal can still receive the first best profit.[17]

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:

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

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

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

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

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

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

In this problem, it is the low type who has an incentive to claim to be a high type, and $(IC^{L,H})$ is binding: since a high type must be given his expected cost following failure, a low

[17] To implement the first best, the principal has to counter the incentive of the low type to pretend to be a high type. Relative to the first best payments, the principal can change the payments to the high type only. She can increase the payment in case of success and lower it otherwise, while keeping the high type at zero expected rent, his level of utility in the first best contract. This payment scheme will lower the rent of the *low* type who is less likely to succeed in the experimentation stage. By choosing the appropriate transfers, the principal can obtain the first best level of profit. While this scheme ensures a zero expected rent for each type, it also implies a large positive ex post rent in case of success and a large negative ex post rent in case of failure. When such schemes are allowed, the first best can be reached. See Theorem 1 in Halac, Kartik, and Liu (2016) for a formal proof in a case without production and a fixed up-front payment.

type will have to be given a rent to truthfully report his type as his expected cost is lower, that is, $c^L_{T^H} < c^H_{T^H}$. We will see that it is also possible that the $(IC^{H,L})$ becomes binding making both (IC) binding simultaneously.

Using our notation, the principal maximizes the following objective function subject to (IRS^L_t), $\left(IRF^L_{T^L}\right)$, (IRS^H_t), $\left(IRF^H_{T^H}\right)$, $(IC^{L,H})$, and $(IC^{H,L})$

$$E_\theta \left\{ \begin{matrix} \beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta \left[V\left(q^\theta_t(\underline{c})\right) - \underline{c} q^\theta_t(\underline{c}) - \Gamma_t\right] + \delta^{T^\theta} P^\theta_{T^\theta} \left[V\left(q^\theta\left(c^\theta_{T^\theta}\right)\right) - c^\theta_{T^\theta} q^\theta\left(c^\theta_{T^\theta}\right) - \Gamma_{T^\theta}\right] \\ -\beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1-\lambda^\theta\right)^{t-1} \lambda^\theta y^\theta_t - \delta^{T^\theta} P^\theta_{T^\theta} x^\theta \end{matrix} \right\}$$

The optimal contract is presented in Proposition 1 and derived in Appendix A. The principal has three tools to screen the agent: the length of the experimentation period, the timing of the payments and the output for each type. We examine them below.

Proposition 1

(i) *In the optimal contract, each type may under-experiment or over-experiment relative to the first best.*

(ii) *Both $(IC^{L,H})$ and $(IC^{H,L})$ can be binding simultaneously. In this case, the principal must reward the low type after failure when the experimentation stage is relatively short and after late success (last period) when the experimentation stage is long enough; the principal must reward the high-type for early success (in the very first period). When only $(IC^{L,H})$ is binding, there is no restriction on when to reward the low type.*

(iii) *After failure, the high type under-produces relative to the first best output. The low type over-produces if the high type receives a rent and produces at the first best level otherwise. After success, each type produces at the first best level.*

Proof: See Appendix A.

2.2.1 The length of the experimentation period: over- or under-experimentation

We already know from the first best contract that the length of the experimentation period is non-monotonic in types. So, in general, the low type or high type may experiment longer. This result extends to the second best and, in general, T^L can be larger or smaller than T^H. Moreover, we find that each type may under- or over-experiment compared to the first best. The reason is that the difference in the expected cost, $\Delta c_t = c^H_t - c^L_t > 0$, is a non-monotonic function of time (see Figure 1). We explain this next.

The driving factor for the rent, and therefore the length of the experimentation stage, is the difference in expected costs between the types after experimentation fails. When the agent lies about his type, his expected cost after failure is different from principal's expected cost. As a result, the agent's informational rent is positively related to the difference in the expected cost different types have. Since Δc_t is a non-monotonic in t, the principal will sometimes benefit from increasing the length of the experimentation stage to reduce rent.[18]

One important aspect of the length of the experimentation period is whether T^L is larger or smaller than a critical value $\hat{T}^L$. This critical value 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. We now turn to this issue and provide the precise derivation of $\hat{T}^L$.

2.2.2 The timing of the payments: rewarding success or failure

Recall that the low type receives a strictly positive rent and $(IC^{L,H})$ is binding.[19] The principal has to determine when to pay this rent to the low type, and the timing of payment can lead to both types of agents getting a rent.

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: there is no restriction on how the principal pays the rent to the low type as long as $\beta_0 \sum_{t=1}^{T^L} \delta^t (1-\lambda^L)^{t-1} \lambda^L y_t^L + \delta^{T^L} P_{T^L}^L x^L = \delta^{T^H} P_{T^H}^L \Delta c_{T^H} q^H\left(c_{T^H}^H\right)$. Therefore, the principal can reward either success, failure, or both.

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 of Appendix A. This is the more interesting case and it allows us to characterize the

---

[18] Consider an example with $V(q) = 3.5\sqrt{q}$, $\beta_0 = 0.7$, $\underline{c} = 0.1$, $\overline{c} = 10$, $\delta = 0.9$, $\lambda^L = 0.15$, $\lambda^H = 0.24$. The first-best termination dates are $T_{FB}^H = 10$ and $T_{FB}^L = 13$. In the second-best case, the principal optimally chooses $T^H = 11$, which is over-experimentation. Consider another example for under-experimentation: $\lambda^L = 0.2$, $\lambda^H = 0.35$. The first-best termination dates are $T_{FB}^H = 8$ and $T_{FB}^L = 11$. In the second-best case, the principal optimally chooses $T^H = 7$, which is under-experimentation.

[19] In Appendix A, we begin the proof of Proposition 1 by proving this result.

stochastic structure of the dynamic problem. In particular, we will explain the derivation of the critical cut-off period $\hat{T}^L$ mentioned above.

As the relative likelihood of reaching different periods is different for each type, paying the rent to the low type by rewarding him early or late has different incentive effects. The goal is to discourage the high type from pretending to be low in order to claim the low-type's rent. We explain below that misreporting his type is a gamble for the high type: he has a chance to obtain the low-type's rent, but he will incur an expected loss in the production stage if he fails during the experimentation stage. The high type has an incentive to pretend to be the low type only if the gamble is positive. We analyze next the details of this gamble, which is the RHS of $(IC^{H,L})$.

We first simplify the analysis by showing in Appendix A that if the principal rewards success, it will be in at most one period.[20] This means that $y_j^L > 0$ for at most one period j. We denote by U^L the rent to the low type, i.e., the LHS of the $(IC^{L,H})$. The principal can pay U^L by rewarding success in some period j ($y_j^L > 0$), or rewarding failure ($x^L > 0$). [21]

If she rewards success in some period j, such that $1 \le j \le T^L$, then

$$y_j^L = \frac{U^L}{\beta_0 \delta^j (1-\lambda^L)^{j-1} \lambda^L} > 0, \tag{1}$$

and $y_t^L = 0$ for $t \neq j$, and $x^L = 0$. If she rewards failure, then

$$x^L = \frac{U^L}{\delta^{T^L} P_{T^L}^L} > 0, \tag{2}$$

and $y_t^L = 0$ for all $t \le T^L$, where $P_t^\theta = 1 - \beta_0 + \beta_0 \left(1-\lambda^\theta\right)^t$.

To understand the nature of restrictions on the timing of rewards, we compare the relative incentive effects on the *high* type of rewarding the *low* type after success or after failure.

(a) If the principal rewards the low type only after success in some period j, with $1 \le j \le T^L$, the high type's expected utility from misreporting (i.e., the RHS of the $IC^{H,L}$ constraint) is:

[20] As we can see in Lemmas 2 and B.2.2 in Appendix A, this is due to the strictly monotonic relative likelihood of success of the two types.

[21] In the knife-edge case where $T^L = \hat{T}^L$, the principal is free to use a combination of x^L and $y_{T^L}^L$ to pay the low-type's rent.

$$\beta_0\delta^j(1-\lambda^H)^{j-1}\lambda^H y_j^L - \delta^{T^L}P^H_{T^L}\Delta c_{T^L}q^L(c^L_{T^L}),$$

which can be rewritten using (1) as,

$$= \frac{\beta_0\delta^j(1-\lambda^H)^{j-1}\lambda^H}{\beta_0\delta^j(1-\lambda^L)^{j-1}\lambda^L}U^L - \delta^{T^L}P^H_{T^L}\Delta c_{T^L}q^L(c^L_{T^L}).$$

(b) If the principal rewards the low type only after failure, the high type's expected utility from misreporting (i.e., the RHS of the $(IC^{H,L})$ constraint) is:

$$\delta^{T^L}P^H_{T^L}[x^L - \Delta c_{T^L}q^L(c^L_{T^L})],$$

which can be re-written using (2) as,

$$= \frac{\delta^{T^L}P^H_{T^L}}{\delta^{T^L}P^L_{T^L}}U^L - \delta^{T^L}P^H_{T^L}\Delta c_{T^L}q^L(c^L_{T^L}).$$

(c) Comparing these two expressions, we can study the nature of the gamble for the high type when he misreports and derive the optimal payment schemes. The choice to reward the low type after success or failure depends on which scheme yields a lower RHS of $(IC^{H,L})$. The first term in each expression is the expected gain from misreporting, and the second is the possibility of incurring a loss if experimentation fails. The second term is identical in the two expressions so we focus on the first term, the expected gain from misreporting.

If the low type is rewarded for success in period j, the relative probability of success of a high type in period j is:

$$\frac{\beta_0\delta^j(1-\lambda^H)^{j-1}\lambda^H}{\beta_0\delta^j(1-\lambda^L)^{j-1}\lambda^L}\ .$$

If the low type is rewarded only after failure, the relative probability of failure is:

$$\frac{\delta^{T^L}P^H_{T^L}}{\delta^{T^L}P^L_{T^L}}\ .$$

The relative probability of success for a high type decreases with j, while the relative probability of failure for a high type is constant given T^L. We show in Appendix A that there is a j such that the RHS of $(IC^{H,L})$ under success or failure equal each other.[22] This is achieved when the two relative probabilities are equal to each other. We can now formally define $\hat{T}^L(=j)$ by setting the two coefficients equal:

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

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

Thus, if the principal wants to reward the low type after success, it will only be optimal if the experimentation stage lasts long enough. If $T_L < \hat{T}^L$, then, $\frac{(1-\lambda^H)^{j-1}\lambda^H}{(1-\lambda^L)^{j-1}\lambda^L} > \frac{P^H_{T^L}}{P^L_{T^L}}$ for all j and the high type will have an advantage over the low type in obtaining any reward given after success. To provide the rent to the low type, the principal will have to reward failure. If $T_L > \hat{T}^L$, the principal can reward success instead. Indeed, it is optimal to pay the reward after success in the last period as $\frac{\beta_0\delta^j(1-\lambda^H)^{j-1}\lambda^H}{\beta_0\delta^j(1-\lambda^L)^{j-1}\lambda^L}$ is declining over time and the rent is the smallest when $j = T^L$.

If the high type gets positive rent, we show in the Appendix A, that the principal will reward him for success in the first period only. Intuitively, this is the period when success is most likely to come from a high type than a low type.

2.2.3 The output: under- or over- production

Finally, the principal can use the choice of output to screen the types and limit the rent to both types. We derive the formal output scheme in Appendix A but present the intuition here. If experimentation was successful, there is no asymmetric information and no reason to distort the output. Both types produce the first best output. If experimentation failed 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.

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.

Asymmetric speed of learning creates a possibility that the high type's incentive constraint also binds. To limit the rent of the high type, the principal will then *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 has two components. The first component is the rent promised to the low type after failure in the experimentation stage. The second component is negative and comes from the higher expected cost of producing the output required from the low type $q^L(c^L_{T^L})$. By making this output higher, the principal can strengthen the negative component and lower the rent of the high type.

To conclude this section, we provide an example to illustrate the interaction between the length of experimentation and the agent's rent, which will also lead us to the following section of the paper. Consider $V(q) = 3.5\sqrt{q}$, $\beta_0 = 0.7$, $\underline{c} = 0.1$, $\overline{c} = 10$, $\delta = 0.9$, $\gamma = 1$, $\lambda^L = 0.14$, and $\lambda^H = 0.35$, then the first-best termination date for the high type agent is $T^H_{FB} = 9$, whereas $T^L_{FB} = 11$. In the optimal contract, $\hat{T}^L = 5$ and the principal optimally chooses $T^H = 8$, $T^L = 11$ and grants rent only to the low type. Note the by asking the high type to under experiment ($T^H = 8 < T^H_{FB} = 9$), the principal mitigates the low type's rent which depends on Δc_{T^H} and Δc_t achieves the highest value at $t_\Delta = 9$.

Suppose now we increase only the parameter $\lambda^H = 0.82$, then the first-best termination date for the high type agent becomes $T^H_{FB} = 3$, whereas $T^L_{FB} = 11$ remains the same. Now the principal optimally chooses $T^H = 3$, $T^L = 10$ and grants rent to both types. Intuitively, when the high type is very efficient in learning the true cost ($\lambda^H = 0.82$), the loss from distorting T^H is high and instead of mitigating the rent to the low type, the principal finds it optimal to give rent to both types.

3. Extensions

3.1 Identical length of experimentation stage for *both* types

As a special case of our model we consider an environment where it is not feasible to screen the agent with the duration of the experimentation stage.[23] That is, the principal must choose an identical length of the experimentation stage for both types. We prove that $(IC^{H,L})$ is never binding. Therefore, the main message of this section is that it is the principal's desire to have different lengths of the experimentation stage that resulted in $(IC^{H,L})$ being binding in the main model.

A contract is now defined formally by

[23] 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.

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

where $\tilde{T}$ is the maximum duration of the experimentation stage *regardless* of the announced type. In Appendix B, we present the first best case when the agent's type θ is common knowledge *before* the principal offers the contract. Recall that when the principal can choose different termination dates, the first-best termination date of the experimentation stage T_{FB}^{θ} is a *non-monotonic* function of the agent's type. Since the expected cost is rising as long as success is not obtained for both types, we immediately conclude that

$$\min_{\theta} T_{FB}^{\theta} \leq \tilde{T}_{FB} \leq \max_{\theta} T_{FB}^{\theta}.$$

This implies that, when the principal is restricted to set the same duration of the experimentation for both agents even if the type of the agent was known, one type will over experiment whereas the other will under experiment.

The principal's optimization problem when the type of the agent is *not known* is to choose contracts ϖ^H and ϖ^L to

$$\max_{\varpi^H, \varpi^L \in \varpi} \upsilon\pi^H(\varpi^H) + (1-\upsilon)\pi^L(\varpi^L) - \upsilon\left\{\beta_0 \sum_{t=1}^{\tilde{T}} \delta^t (1-\lambda^H)^{t-1}\lambda^H y_t^H + \delta^{\tilde{T}} P_{\tilde{T}}^H x^H\right\}$$

$$-(1-\upsilon)\left\{\beta_0 \sum_{t=1}^{\tilde{T}} \delta^t (1-\lambda^L)^{t-1}\lambda^L y_t^L + \delta^{\tilde{T}} P_{\tilde{T}}^L x^L\right\} \text{ subject to}$$

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

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

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

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

$\left(IRS_t^{\theta}\right)$ $y_t^{\theta} \geq 0$ for $t \leq T^{\theta}$, and

$\left(IRF_{T^{\theta}}^{\theta}\right)$ $x^{\theta} \geq 0$ for $\theta = H, L$.

Proposition 2. *If the duration of the experimentation stage must be chosen independently of the announced type, the high type gets no rent. The principal can choose any combinations of payments to the low type such that* $U^L = \delta^{\tilde{T}} P_{\tilde{T}}^L \Delta c_{\tilde{T}} q^H\left(c_{\tilde{T}}^H\right)$. *Each type may under-experiment or over-experiment relative to the first best. The high type under-produces relative to the first best output. The low type produces at the first best level.*

Proof: See Appendix B.

Based on Proposition 2, 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 in Proposition 1. We found that the termination date of the experimentation stage for each type determines the difference in the expected cost when production occurs after failure, Δc_{T^L} and Δc_{T^H}. These terms play a critical role in the gamble faced by a lying efficient agent. If $\tilde{T}$ is the same for both types, the difference in the expected cost are the same, and the gamble is determined by the difference in optimal quantities. Then, the gamble is negative and the $(IC^{H,L})$ is not binding. Indeed, using $(IC^{H,L})$, the high type pretending to be the low type receives

$$\delta^{\tilde{T}} P_{\tilde{T}}^{H} \Delta c_{\tilde{T}} \left(q^{H}\left(c_{\tilde{T}}^{H}\right) - q^{L}\left(c_{\tilde{T}}^{L}\right)\right).$$

We show in Appendix B that we have $q^{H}\left(c_{\tilde{T}}^{H}\right) < q_{FB}^{H}\left(c_{\tilde{T}}^{H}\right) < q^{L}\left(c_{\tilde{T}}^{L}\right)$. As a result, the high type receives a strictly negative utility if he mimics a low type. [24]

As in the main model, there is no distortion in the output relative to the first-best level after success. After failure, only the high type output is distorted. There is underproduction by the high type when experimentation fails: $q_{SB}^{H}\left(c_{\tilde{T}}^{H}\right) < q_{FB}^{H}\left(c_{\tilde{T}}^{H}\right)$.

3.2. Success might be hidden: ex post moral hazard

A notable result from the main section was that the principal may want to reward failure, or wait until later periods to reward success. The implementation of schemes with such properties relies on our assumption that the outcome of experiments in each period is publicly observable. If the agent were able to suppress a finding of success, he would gain by hiding success, or postponing the revelation of success. In this subsection, we allow the agent to engage in ex post moral hazard by hiding success when it occurs.[25] The dynamic optimization problem for the principal when success is privately observed by the agent becomes more complex as the principal has to deal with both adverse selection and ex post moral hazard problems simultaneously. We find that the agent will obtain a moral-hazard based rent in each period, but the principal's incentive to delay reward to the low-type to screen the high type agent remains

[24] In this case, the intuition from the static second best contract applies as in our benchmark case without experimentation but with asymmetric information about expected cost. We found that the $(IC^{H,L})$ is not binding since the high type also produces less than the low type.

[25] See, e.g., Rosenberg, Salomon, and Vieille (2013).

intact. In particular, we show that our finding of rewarding the agent after failure does not depend on success being observed publicly.

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 $c^{\theta}_{T^{\theta}}$ 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}) $y^{\theta}_{T^{\theta}} \geq x^{\theta} + (c^{\theta}_{T^{\theta}} - \underline{c})q^{\theta}(c^{\theta}_{T^{\theta}})$ for $\theta = H, L$, and

(EMP^{θ}_{t}) $y^{\theta}_{t} \geq \delta y^{\theta}_{t+1}$ 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 Appendix C, both $(IC^{H,L})$ and $(IC^{L,H})$ may be slack, and either or both may be binding.[26] 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.

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

Proposition 3. *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 Appendix C.

While details and the formal proof are in the Appendix C, we now provide some intuition why rewarding failure remains 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 1, we again find that it is optimal to reward failure when the low type 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 grants rent only to the low type by rewarding failure since $\hat{T}^L = 3$.

While we have focused on how the ex post moral hazard affects the benefit of rewarding failure, it is clear that 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 (which will decrease $c^\theta_{T^\theta}$) or $q^\theta\left(c^\theta_{T^\theta}\right)$. 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 $c = \overline{c}$. For instance, stage 1 of a drug trial looks for bad news by testing the safety of the drug. Following the literature on experimentation we call an event of observing $c = \overline{c}$ by the agent "success" although this is a 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.

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 again 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, $c_{T^L}^H < c_{T^L}^L$. The details of the optimization problem mirror the case for good news of Proposition 1, and the results are similar.

Proposition 4:

(i) *In the optimal contract, each type may under-experiment or over-experiment relative to the first best.*

(ii) *Both $(IC^{L,H})$ and $(IC^{H,L})$ can be binding simultaneously. In this case, the principal must reward the high-type for early success (in the very first period.) If the low type receives a rent, he is rewarded for failure when the experimentation stage is relatively short and for late success (last period) when the experimentation stage is long enough. When only $(IC^{H,L})$ is binding, there is no restriction on when to reward the high type.*

(iii) *After failure, the high type under-produces relative to the first best output. The low type over-produces if the high type receives a rent and produces at the first best level otherwise. After success, each type produces at the first best level.*

Proof: See Appendix D.

We find similar restrictions when both (IC) constraints bind as in Proposition 1. The type of news, however, determines the optimal production and length of experimentation decisions. Regarding the distortion in the output, it is similar to Proposition 1 except that the distortions in output are switched for the two types. 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. This interaction affects the optimal project scale. While success in experimentation typically resolves uncertainty in a two-armed bandit model, learning still occurs after successive failures, and it determines the scale of the project. While there has been much recent attention on studying incentives for experimentation in two-armed bandit settings, details of the production stage are typically suppressed to focus on incentives for exploration. In reality, each stage impacts 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 also related it to the length of the learning stage.

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 in order 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 not long enough, 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 and experimenting for 'good news', we show that our main insights survive if the agent can hide success or if we changed to model to learn 'bad news'. Without a production stage with a scalable project size after failure, there would be under experimentation relative to the first best. With a scalable project size, we find a new result that over-experimentation can be also optimal. Over production can occur in the production stage.

References

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.

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

Cremer J., Khalil F. and Rochet J.-C. (1992). Contracts and productive information gathering. *Games and Economic Behavior* 25(2), 174-193.

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.

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

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

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.

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

Appendix A (Proof of Proposition 1)

We first characterize the optimal payment structure, x_L, $\{y_t^L\}_{t=1}^{T^L}$, x_H and $\{y_t^H\}_{t=1}^{T^H}$ (part (*ii*) of Proposition 1), then the optimal length of experimentation, T^L and T^H(part (*i*) of Proposition 1), and finally the optimal outputs $\left\{q_t^H\left(\underline{c}\right)\right\}_{t=1}^{T^H}$, $q^H\left(c_{T^H}^H\right)$, $\left\{q_t^L\left(\underline{c}\right)\right\}_{t=1}^{T^L}$ and $q^L\left(c_{T^L}^L\right)$ (part (*iii*) of Proposition 1).

Denote the expected surplus net of costs for $\theta = H, L$ by $\Omega^\theta\left(\varpi^\theta\right) = \beta_0 \sum_{t=1}^{T^\theta} \delta^t \left(1 - \lambda^\theta\right)^{t-1} \lambda^\theta \left[V\left(q_t^\theta\left(\underline{c}\right)\right) - \underline{c} q_t^\theta\left(\underline{c}\right) - \Gamma_t\right] + \delta^{T^\theta} P_{T^\theta}^\theta \left[V\left(q^\theta\left(c_{T^\theta}^\theta\right)\right) - c_{T^\theta}^\theta q^\theta\left(c_{T^\theta}^\theta\right) - \Gamma_{T^\theta}\right]$. The principal's optimization problem then is to choose contracts ϖ^H and ϖ^L to maximize the expected net surplus minus rent of the agent, subject to the respective IC and IR constraints given below:

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

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

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

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

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

(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 $\Delta c_{T^H} q^H\left(c_{T^H}^H\right) > 0$. *Q.E.D.*

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

First we show that if the high type claims to be the low type, the high type is relatively more likely to succeed if experimentation stage is smaller than a threshold level, $\hat{T}^L$.

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 probability that the high type does not succeed to probability that the low type does not succeed for T^L periods during the experimentation stage. At the same time, $\beta_0(1-\lambda^\theta)^{t-1}\lambda^\theta$ is probability that the agent of type θ succeeds at period $t \le T^L$ of the experimentation stage and $\frac{\beta_0(1-\lambda^H)^{t-1}\lambda^H}{\beta_0(1-\lambda^L)^{t-1}\lambda^L} = \frac{(1-\lambda^H)^{t-1}\lambda^H}{(1-\lambda^L)^{t-1}\lambda^L}$ is a ratio of 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

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

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 $\frac{\lambda^H}{\lambda^L} > 1 > \frac{P_{T^L}^H}{P_{T^L}^L}$.[27] 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.*

---

[27] 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 $\frac{P_{T^L}^H}{P_{T^L}^L} = \frac{(1-\lambda^H)^{\hat{T}^L-1}\lambda^H}{(1-\lambda^L)^{\hat{T}^L-1}\lambda^L}$.

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 $y_t^H = 0$ for $1 \le t \le T^H$, which implies that the rent of the low type in this case becomes $\delta^{T^H} P_{T^H}^L \Delta c_{T^H} q^H(c_{T^H}^H)$. Replacing x_L in the objective function, the principal's optimization problem is to choose T^H, $\{q_t^H(\underline{c})\}_{t=1}^{T^H}$, $q^H(c_{T^H}^H)$, T^L, $\{y_t^L\}_{t=1}^{T^L}$, $\{q_t^L(\underline{c})\}_{t=1}^{T^L}$ and $q^L(c_{T^L}^L)$ to

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

(IRS_t^L) $y_t^L \ge 0$ for $t \le T^L$,

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

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

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

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

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

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

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:

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

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

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

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

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

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

First, we consider the case when $\psi \neq 0$. This is when the likelihood ratio of reaching the last period of the experimentation stage is different for both types i.e., when $\frac{P_{T^H}^H}{P_{T^H}^L} \neq \frac{P_{T^L}^H}{P_{T^L}^L}$ (Case B.1). We show 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 $T^L > \hat{T}^L$. We also show that the high type may may be rewarded only for the very first success. Finally, we analyze the case when $\frac{P_{T^H}^H}{P_{T^H}^L} = \frac{P_{T^L}^H}{P_{T^L}^L}$ (Case B.2). In this case, the likelihood ratio of reaching the last period of the experimentation stage is the same for both types and x^H and x^L cannot be used as screening variables. Therefore, the principal must reward both types for success and she chooses $T^L > \hat{T}^L$.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $1 \leq k, m \leq T^L$ such that $k \neq m$ and $y_k^L, y_m^L > 0$. Then from the Kuhn-Tucker conditions $(A1)$ and $(A2)$ it follows that

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

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

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

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

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

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

For later use, we prove the following claim:

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

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

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

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

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 $1 \leq t \leq T^L$, which leads to a contradiction since then $x^L = y_t^L = 0$ for $1 \leq t \leq T^L$ 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 $1 \leq t \leq T^H$.

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 $(IRF_{T^H}^H)$ and $(IRF_{T^L}^L)$ constraints are not binding.

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

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

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

Since $f_1(t,T^H)$ is strictly positive for all $t < \hat{T}^H$ from $P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L\right] = -\frac{\alpha_t^H\psi}{\beta_0\delta^t}$ it must be that $\alpha_t^H > 0$ for all $t < \hat{T}^H$ and $\psi < 0$. In addition, since $f_2(t,T^L)$ is strictly negative for $t < \hat{T}^L$ from $P_{T^L}^L f_2(t,T^L)\left[\upsilon P_{T^H}^H + (1-\upsilon)P_{T^H}^L\right] = -\frac{\alpha_t^L\psi}{\beta_0\delta^t}$ it must be that $\alpha_t^L > 0$ for $t < \hat{T}^L$ and $\psi > 0$, which leads to a contradiction[28].

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

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

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

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

If $\alpha_s^H = 0$ for some $1 \le s \le T^H$ then $P_{T^H}^H f_1(s,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L - \frac{\xi^L}{\delta^{T^L}}\right] = 0$, which implies that $\frac{\xi^L}{\delta^{T^L}} = \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$[29]. 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 $1 \le t \le T^H$.

---

[28] 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.

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

Finally, from $P_{T^H}^H f_1(t,T^H)\left[\upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L - \frac{\xi^L}{\delta^{T^L}}\right] = -\frac{\alpha_t^H \psi}{\beta_0 \delta^t}$ we conclude that $T^H \leq \hat{T}^H$ and there can be one of two sub-cases:[30] (b.1) $\psi > 0$ and $\frac{\xi^L}{\delta^{T^L}} > \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$, or (b.2) $\psi < 0$ and $\frac{\xi^L}{\delta^{T^L}} < \upsilon P_{T^L}^H + (1-\upsilon)P_{T^L}^L$. We consider each sub-case next.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

For both cases we just considered, we have

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

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

Note that Case B.2 is possible only if $\delta^{T^H} P^L_{T^H} \Delta c_{T^H} q^H\left(c^H_{T^H}\right) - \delta^{T^L} P^L_{T^L} \Delta c_{T^L} q^L\left(c^L_{T^L}\right) > 0$[31]. 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 $\delta^{T^H} P^H_{T^H} \Delta c_{T^H} q^H\left(c^H_{T^H}\right) - \delta^{T^L} P^H_{T^L} \Delta c_{T^L} q^L\left(c^L_{T^L}\right) < 0$. However, $\delta^{T^H} P^L_{T^H} \Delta c_{T^H} q^H\left(c^H_{T^H}\right) > \delta^{T^L} P^L_{T^L} \Delta c_{T^L} q^L\left(c^L_{T^L}\right)$ implies that $\delta^{T^H} P^H_{T^H} \Delta c_{T^H} q^H\left(c^H_{T^H}\right) > \delta^{T^L} \frac{P^H_{T^H}}{P^L_{T^H}} P^L_{T^L} \Delta c_{T^L} q^L\left(c^L_{T^L}\right)$. Note that $P^H_{T^H} P^L_{T^L} - P^H_{T^L} P^L_{T^H} > 0$ implies $\frac{P^H_{T^H}}{P^L_{T^H}} P^L_{T^L} > P^H_{T^L}$, and then $\delta^{T^H} P^H_{T^H} \Delta c_{T^H} q^H\left(c^H_{T^H}\right) > \delta^{T^L} P^H_{T^L} \Delta c_{T^L} q^L\left(c^L_{T^L}\right)$, which implies $x^L > 0$ and we have a contradiction. Thus, $\xi^H > 0$ and the high type gets rent only after success ($x^H = 0$). *Q.E.D.*

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

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 $y^H_1 > 0$).

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

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

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

If $\alpha^L_s = 0$ for some $1 \le s \le T^L$ then $P^L_{T^L} f_2(t, T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon) P^L_{T^H} - \frac{\xi^H}{\delta^{T^H}}\right] = 0$, which implies that $\frac{\xi^H}{\delta^{T^H}} = \upsilon P^H_{T^H} + (1-\upsilon) P^L_{T^H}$[32]. Since we already rule out this possibility it immediately follows that $\alpha^L_t > 0$ for all $1 \le t \le T^L$ which implies that $y^L_t = 0$ for $1 \le t \le T^L$.

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

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

Finally, $P^L_{T^L} f_2(t,T^L)\left[\upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H} - \frac{\xi^H}{\delta^{T^H}}\right] = -\frac{\alpha^L_t \psi}{\beta_0 \delta^t}$ 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:[33] (a) $\psi > 0$ and $\frac{\xi^H}{\delta^{T^H}} < \upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}$, or (b) $\psi < 0$ and $\frac{\xi^H}{\delta^{T^H}} > \upsilon P^H_{T^H} + (1-\upsilon)P^L_{T^H}$. We consider each sub-case next.

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

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

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

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

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 $f_1(t,T^H) f_2(j,T^L) < f_1(j,T^H) f_2(t,T^L)$

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

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

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

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

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

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

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

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

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

$$y_1^H=\frac{P_{T^L}^H\left(\delta^{T^L}P_{T^L}^L\Delta c_{T^L}q^L\left(c_{T^L}^L\right)-\delta^{T^H}P_{T^H}^L\Delta c_{T^H}q^H\left(c_{T^H}^H\right)\right)}{\beta_0\delta P_{T^L}^L f_2(1,T^L)}\ge 0;$$

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

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

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

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

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

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

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

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

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

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 $1\le t\neq j\le T^L$.

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 $1\le t\neq j\le T^L$.

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 $1\le t\neq j\le T^L$ 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 $1\le t\ne s\le T^H$ which implies that $s=T^H>\hat{T}^H$. *Q.E.D.*

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

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

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

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

$$x^H = \frac{\beta_0 \delta^{T^H} P_{T^L}^L f_2(T^H,T^L) y_{T^H}^H - \beta_0 \delta P_{T^L}^L f_2(1,T^L) y_1^L}{\delta^{T^H}\psi} + \frac{P_{T^L}^H\left(\delta^{T^H} P_{T^H}^L \Delta c_{T^H} q^H\left(c_{T^H}^H\right) - \delta^{T^L} P_{T^L}^L \Delta c_{T^L} q^L\left(c_{T^L}^L\right)\right)}{\delta^{T^H}\psi} = 0;$$

$$x^L = \frac{\beta_0 P_{T^H}^H \delta f_1(1,T^H) y_1^L - \beta_0 \delta^{T^H} P_{T^H}^H f_1(T^H,T^H) y_{T^H}^H}{\delta^{T^L}\psi} + \frac{P_{T^H}^L\left(\delta^{T^H} P_{T^H}^H \Delta c_{T^H} q^H\left(c_{T^H}^H\right) - \delta^{T^L} P_{T^L}^H \Delta c_{T^L} q^L\left(c_{T^L}^L\right)\right)}{\delta^{T^L}\psi} = 0.$$

Since Case B.2 is possible only if $\delta^{T^H} P_{T^H}^L \Delta c_{T^H} q^H\left(c_{T^H}^H\right) - \delta^{T^L} P_{T^L}^L \Delta c_{T^L} q^L\left(c_{T^L}^L\right) > 0$[34], 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:

$$y_1^H = \frac{\delta^{T^H} P_{T^H}^L \Delta c_{T^H} q^H\left(c_{T^H}^H\right)(1-\lambda^H)^{T^L-1}\lambda^H - \delta^{T^L} P_{T^L}^L \Delta c_{T^L} q^L\left(c_{T^L}^L\right)(1-\lambda^L)^{T^L-1}\lambda^L}{\beta_0 \delta \lambda^L \lambda^H\left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)} \geq 0;$$

$$y_{T^L}^L = \frac{\left(\delta^{T^H}\lambda^H P_{T^H}^L \Delta c_{T^H} q^H\left(c_{T^H}^H\right) - \delta^{T^L}\lambda^L P_{T^L}^H \Delta c_{T^L} q^L\left(c_{T^L}^L\right)\right)}{\beta_0 \delta^{T^L}\lambda^L\lambda^H\left((1-\lambda^L)^{T^L-1} - (1-\lambda^H)^{T^L-1}\right)} > 0. \qquad Q.E.D.$$

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

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

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

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

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

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

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

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. $P_{T^H}^H P_{T^L}^L - P_{T^L}^H P_{T^H}^L = 0 \Rightarrow T^L > \hat{T}^L$, $\xi^H > 0$, $\xi^L > 0$, $\alpha_t^H > 0$ for $t > 1$ and $\alpha_t^L > 0$ for $t < T^L$ (it is optimal to set $x^L = x^H = 0$, $y_t^H = 0$ for $t > 1$ and $y_t^L = 0$ for $t < T^L$).

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

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

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

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

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

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

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

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

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

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

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

Combining the two equations together, $\alpha^L P_{T^H}^H\big(f_1(k,T^H) f_2(m,T^L) - f_1(m,T^H) f_2(k,T^L)\big)$ $+\upsilon\lambda^H\Big((1-\lambda^H)^{k-1} f_2(m,T^L) - (1-\lambda^H)^{m-1} f_2(k,T^L)\Big) = 0$, which can be rewritten as follows[35]:

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

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

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

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

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

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

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

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

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

(part (*i*) of Proposition 1)

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 $\delta^{T^H}P_{T^H}^L\Delta c_{T^H}q^H(c_{T^H}^H)$ 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:

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

Since the informational rent of the low-type agent, $\delta^{T^H}P_{T^H}^L\Delta c_{T^H}q^H(c_{T^H}^H)$, is non-monotonic in T^H, it is possible, in general, to have $T_{SB}^H>T_{FB}^H$ or $T_{SB}^H<T_{FB}^H$.

In Case B, the exact values of 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 $U^H\ge 0$. It is possible, in general, to have $T_{SB}^H>T_{FB}^H$ or $T_{SB}^H<T_{FB}^H$ and $T_{SB}^L>T_{FB}^L$ or $T_{SB}^L<T_{FB}^L$. See footnote 18 in the main text for an example.

III. Optimal outputs

(part (*iii*) of Proposition 1)

The optimal $q_t^\theta(\underline{c})$ is efficient as it chosen to maximize $E_\theta\,\Omega^\theta(\varpi^\theta)$ in the Lagrangian.

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

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

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

$$V'\left(q^L\left(c^L_{T^L}\right)\right)-c^L_{T^L}=0.$$

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

We will prove that there is over production for the low type ($q^L_{SB}\left(c^L_{T^L}\right)>q^L_{FB}\left(c^L_{T^L}\right)$) and under production for the high type ($q^H_{SB}\left(c^H_{T^H}\right)<q^H_{FB}\left(c^H_{T^H}\right)$) after failure. We start with the main case when $\psi\neq 0$, and consider cases when $T^L\leq\hat{T}^L$ and $T^L>\hat{T}^L$ separately.

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

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

$$\upsilon P^H_{T^H}\left(V'\left(q^H\left(c^H_{T^H}\right)\right)-c^H_{T^H}\right)=\frac{P^L_{T^H}\Delta c_{T^H}E_\theta\left\{P^\theta_{T^L}\right\}}{P^L_{T^L}\lambda^H-P^H_{T^L}\lambda^L}.$$

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

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

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

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

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

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