Investing in Failure: Belief Manipulation and Strategic Experimentation[1]

Peter Buisseret

15/10/14

Abstract

A Principal hires an Agent to experiment with a technology of unknown quality. The Principal invests funds and the Agent supplies costly, hidden effort. Higher investment by the Principal raises the probability of success when the Agent works and the project is good, but increases the Agent's dynamic information rent from shirking. I show that the Principal may over-invest relative to the first best. This commits the Principal to sufficiently pessimistic beliefs upon initial failure that she discontinues the project, denying the Agent rents from the repeated interaction. When the Principal can write a long-term contract, over-investment may persist and inefficiency increase.

Keywords— strategic experimentation, learning, agency

[1] *Author affiliation and address*: Department of Economics, Warwick University, Coventry, CV4 7AL, UK. *Email*:p.buisseret@warwick.ac.uk. *Telephone*: +44 (0)24 7652 8248. *Fax*: +44(0)2476 523032.

Acknowledgements: I am indebted to Steven Callander for his generous and extremely helpful discussion of an early draft of this project. I also thank Avidit Acharya, Dan Bernhardt, Marina Halac, Niall Hughes, Navin Kartik, Motty Perry and Carlo Prato for suggestions and encouragement. I extend my deepest gratitude to Lora Appel for her advice and support during the early stages of the project. Wentao Fu provided outstanding research assistance. All errors and omissions are solely my responsibility.

1. Introduction

A fundamental problem of experimentation is knowing when to quit. I explore how this decision is affected by the presence of an Agent to whom a Principal must delegate a role in the experiment. Previous work highlights *under-investment* and early abandonment of productive projects (Bergemann and Hege (2005) or Hörner and Samuelson (2013)) as a consequence of experimentation in a principal-agent setting. This paper provides a rationale for why *over-investment* might occur relative to the efficient benchmark, whilst still inducing early project abandonment.

In my model, the Principal invests funds in a project which is either *good*, with probability v_0, or *bad*, with probability $1 - v_0$. At each date, she chooses investment and offers the Agent a share of the one-off surplus that may be generated from the project. The role of the Agent is to exert costly, hidden effort. If the project is good and the Agent works, the probability of success is proportional to the funds invested by the Principal at that date; otherwise, it fails for sure.

This technology induces the following agency problem: suppose that the Principal observes an initial failure and believes that the Agent worked on the project. Then, her posterior belief about its quality falls to $\hat{v} < v_0$, and the difference $v_0 - \hat{v}$ strictly increases in the initial investment. If the Principal wishes to continue for another period she will offer a contract that holds the Agent-type who shares her belief to her participation constraint. But suppose that, in reality, the Agent shirked, so that initial failure was inevitable. Then, the Agent's belief about the project is in fact v_0. But the compensation package designed for the Agent-type $\hat{v}$ gives rent to an Agent-type $v_0 > \hat{v}$. This creates a moral hazard problem which stems from the private information about the project quality created by the Agent's deviation.

The Principal faces a trade-off: choosing higher initial levels of investment at the outset of the project increases the probability of success, conditional on the project being good and the Agent working hard. But higher initial levels of investment also increase the divergence between the Principal's prior and posterior beliefs after observing failure, since failure is more informative when initial investment is larger. This, in turn, raises the continuation value to the Agent from shirking, since her information rent is increasing in the divergence of v_0 and $\hat{v}$.

Relative to the first-best, I show that the Principal may choose an inefficiently high level of initial investment. Given her belief that the Agent worked, the informativeness of failure under such a large initial investment leads her to discontinue the project, eliminating the Agent's opportunity to acquire dynamic information rent. When the Principal can only write a short-term contract (i.e. within-period commitment) this choice of investment partly serves as a *belief commitment* device for the Principal to ensure that she prefers to quit after early failure. Nonetheless, over-investment also occurs when the Principal can commit to a long-term contract (i.e. between-period commitment); in fact, there may be instances in which the Principal invests more under long-term commitment than *either* the first-best contract *or* short-term commitment. Moreover, the long-term contract may increase inefficiency since it allows the Principal to credibly destroy even more of the surplus associated with the two-date investment solution in order to deny rents to the Agent.

The agency problem that I study builds on Bhaskar (2012).[2] In his model, an Agent chooses privately observed effort, and the Principal's problem is solely to provide pecuniary incentives to induce effort. Bhaskar (2012) characterizes the Principal's optimal contract when the Agent can create private information in the same way as the present model, in a setting where the project is always profitable for any belief. I show that when there are beliefs about project quality for which the Principal prefers to abandon the project, and she is able to manipulate the distribution of the public signal - and thus the evolution of her posterior belief upon observing failure - she may do so as a means to soften the Agent's incentive constraint. The interpretation of this manipulation as a decision about how much to invest connects the model to a large literature on strategic experimentation in which the Principal chooses a level of funding and delegates the funds to the Agent (Bergemann and Hege (2005) or Hörner and Samuelson (2013)). These papers also characterize investment distortions as a consequence of the Agency problem, however the funding inefficiency in these models manifests as *under-investment*, or a delay in the release of funds. Inefficient over-investment is ruled out in Bergemann and Hege (2005) because the Principal's cost of funds is linear; in Hörner and Samuelson (2013) it is ruled out because the Principal's investment decision has binary support. As a result, early quitting in existing models arises despite the fact that it would be socially optimal to continue experimenting for additional periods. In my model, by contrast, the Principal's decision not to experiment in the second period is socially optimal *given her investment decision in the first period*: the initial choice of investment effectively destroys tomorrow's surplus in order to allow the Principal to withhold rents from the Agent, today.

2. The Model

There is a Principal and an Agent and two dates, $t \in \{t_0, t_1\}$. First, Nature selects a project quality $\omega \in \{0, 1\}$, according to $\Pr(\omega = 1) = v_0 \in (0, 1)$ which persists across both dates. This choice is not observed by the players. Next, the Principal chooses $\lambda_t \in [0, 1]$ at cost $\tau(\lambda_t)$ which is strictly increasing, strictly convex, twice differentiable and $\tau(0) = 0$. She then offers the Agent a contract, described in greater detail, below. The Agent subsequently chooses effort, $e_t \in \{0, 1\}$ at cost $c(e_t)$ satisfying $c(0) = 0$ and $c(1) = c > 0$. Only the Agent observes this choice. A public signal $s_t \in \{s_t^L, s_t^H\}$ is then realized according to the distribution:

$$\Pr(s_t^H | \lambda_t, e_t, \omega) = \lambda_t e_t \omega \tag{1}$$

The public signal s_t^H is interpreted as a 'success' and s_t^L a 'failure' at date t. If a success is realized, a one-off surplus of W is generated and the game ends at date zero.[3] Otherwise, no surplus is generated and, at date zero, the game proceeds to date one. The timing of the game

[2]See also Bhaskar (2014). There is a large and growing literature on strategic experimentation: in addition to work discussed explicitly, recent contributions include Halac, Kartik and Liu (2013), who study a model with adverse selection and moral hazard, and Manso (2011).

[3]The results are not affected by an alternative assumption that the project is implemented again at date one in the event of success at date zero. However, the assumption that experimentation ends after a breakthrough is more common in the literature: see, for example, Bergemann and Hege (2005) and Hörner and Samuelson (2013).

at date one is the same as at date zero, beginning with the Principal's investment decision. The common discount factor is $\beta \in (0,1)$.

Contracts, Histories and Equilibrium

A date one *history* is $h_1 = (\lambda_0, s_0, e_0)$. A date one *public history* is $h_1^p = (\lambda_0, s_0)$. The set of all date one public histories is H_1^p, and the set of all date one histories is H_1. The date zero history is the null history. Let v_1^p denote the belief of the Principal in period t_1 that the project is good.

The Principal's contract with the Agent specifies a bonus to be paid after a success or failure at each date. I initially assume that the Principal has short-term (intra-period) commitment. Her strategy specifies at each date t an investment choice $\lambda_t(h_t^p)$ as a function of the public history, and a bonus $b_t(h_t^p, \lambda_t, s_t)$ conditioned on the public history, the investment decision and the realization of the public signal at that date. I assume that all transfers from the Principal must be weakly positive, and the Agent's outside option is zero at each date. The Agent's (pure) strategy specifies at each date an effort choice as a function of the history, the Principal's investment and the bonuses offered at that date. The static surplus at date t when the agent chooses high effort and the Principal invests λ_t is:

$$f(\lambda_t; v_t^p) \equiv v_t^p \lambda_t W - \tau(\lambda_t) - c \tag{2}$$

I let $\lambda_t^*(v_t^p)$ denote the maximizer of $f(\lambda_t; v_t^p)$ and $S(v_1^p) = \max\{0, f(\lambda_1^*(v_1^p); v_1^p)\}$ is the surplus in the continuation game beginning at date one, after failure at date zero. I assume $\lambda_t^*(1) \in (0,1)$ and $S(1) > 0$.

3. Equilibrium Experimentation

I begin by characterizing properties of the Principal's investment at date one, as a function of her date zero investment decision. Throughout, I assume that if the Principal is indifferent between zero and positive investment, she chooses zero investment. I refer to the setting in which the Principal observes the Agent's effort as the 'first-best', and that in which she cannot directly observe the Agent's effort as the 'second-best'.

Lemma 1. In an equilibrium in the first-best or the second-best:

i. after any two date one public histories h_1^p and $\bar{h}_1^p$ with respective date zero investment levels λ_0 and $\bar{\lambda}_0$, and in which failure occurred at date zero, $\bar{\lambda}_0 > \lambda_0$ implies $\lambda_1^*(\bar{v}_1^p) \leq \lambda_1^*(v_1^p)$

ii. there exists a unique threshold $\underline{v} \in (0,1)$ such that if $v_1^p \leq \underline{v}$, $\lambda_1^*(v_1^p) = 0$ and for all $v_1^p > \underline{v}$, $\lambda_1^*(v_1^p) > 0$.

Point (i) states that the more the Principal invested at date zero, so long as the Agent is believed to have exerted effort, initial failure becomes increasingly bad news about project quality. And more pessimism induces a lower flow of funds at date one. According to (ii) if the Principal becomes sufficiently pessimistic after initial failure, she may choose to make no subsequent investment at date one, and instead abandon the project. For the rest of the paper, I assume

$v_0 > \underline{v}$, so that there is a rationale for strictly positive investment in at least one of the two dates.

We have obtained a threshold value of the posterior ($\underline{v}$) below which no further investment occurs at date two. It is useful to characterize for each date zero prior belief v_0 the level of investment at date zero which will induce a posterior $\underline{v}$ upon initial failure. This level of investment is:

$$\hat{\lambda}_0(v_0) = \frac{v_0 - \underline{v}}{v_0(1-\underline{v})} \tag{3}$$

A consequence of Lemma 1 is that if the Principal invests below this threshold at date zero, she will wish to continue with positive investment at date one. If, instead, she invests above this threshold at date zero, early failure will lead her to abandon the project and make no further investment at date one. Note that $\hat{\lambda}_0(v_0)$ is increasing and strictly concave in v_0. The crux of the subsequent analysis will be to understand how the Principal's choice of investment above or below this threshold is distorted under unobservable effort, relative to the benchmark of complete information.

I now consider the Principal's remuneration to the Agent at date one. It is easy to show that the Principal never offers the Agent a positive bonus after a failure at either date, at any solution in the first- and second-best. I therefore write $b_t(v_t^p, \lambda_t)$ as the bonus offered by the Principal at date t in the event that a success is realized. I first characterize optimal investment and bonuses at date one. Recall that $\lambda_t^*(v_t^p)$ is the (interior) maximizer of the static investment problem $f(\lambda_t; v_t^p)$ at date t, under the supposition that the Agent chooses high effort.

Lemma 2. At date one, if and only if $v_1^p \leq \underline{v}$, the Principal chooses zero investment. If $v_1^p > \underline{v}$ the Principal chooses $\lambda_1^*(v_1^p)$, and offers the Agent the bonus $b_1(v_1^p, \lambda_1) = \frac{c}{\lambda_1 v_1^p}$.

Proof. On the equilibrium path, the Principal's and the Agent's beliefs are $v_1^p = v_1$. For any choice $\lambda_1 > 0$, the Agent's incentive and participation constraints coincide, so the Principal satisfies both with $b_1(v_1^p, \lambda_1) = \frac{c}{\lambda_1 v_1^p}$ for $\lambda_1 > 0$ and $b_1(v_1^p, 0) = 0$. Thus, λ_1 solves $S(v_1^p)$, so that the Principal's investment problem coincides with the first-best, at that date. □

I now proceed to the Principal's contract in t_0. Under the first-best, her initial choice of investment generates the value:[4]

$$S(v_0) = \max_{\lambda_0 \in [0,1]} [f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta S(v_1^p)] \tag{4}$$

Let $\lambda_0^{FB}(v_0)$ denote the solution to this problem. To motivate the analysis for the case of unobserved effort, suppose that the Principal's belief about the state date one (v_1^p) is based on a presumption that the Agent worked hard at date zero. Suppose also that $v_1^p > \underline{v}$, so that the date one surplus is positive and thus the Principal will optimally continue to invest at date one. At that date, let the Agent's belief that the project is good be v_1. Using Lemma 2, we may write the Agent's continuation pay-off as a function of each player's belief:

$$V(v_1, v_1^p) = c\left(\frac{v_1}{v_1^p} - 1\right) \tag{5}$$

[4] Given $v_0 > \underline{v}$, it is easy to show that in the first best, the Principal will always choose $\lambda_0 > 0$. This is not guaranteed when effort is unobserved.

On the equilibrium path, $V(v_1, v_1^p) = V(v_1^p, v_1^p) = 0$. Suppose, however, that the Agent initially deviated to no effort, unbeknownst to the Principal. Then, $v_1 = v_0$, and:

$$V(v_0, v_1^p) = c\frac{\lambda_0}{1-\lambda_0}(1 - v_0) \tag{6}$$

which is strictly increasing in λ_0: more initial investment by the Principal induces ever greater pessimism in the light of initial failure. Therefore, so long as $v_1^p > \underline{v}$, greater pessimism induces a larger bonus in order to satisfy the type-v_1^p Agent's participation constraint. This creates a powerful temptation on the part of the Agent to shirk at date zero. Henceforth, I write $V(\lambda_0)$ as the Agent's continuation pay-off from shirking when the Principal believes she exerts effort and invests λ_0. We have:

$$V(\lambda_0) = \begin{cases} c\frac{\lambda_0}{1-\lambda_0}(1 - v_0) & \text{if } 0 < \lambda_0 < \hat{\lambda}_0(v_0)) \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Equipped with the Agent's payoff from shirking when the Principal conjectures that she chooses high effort, it is straightforward to derive the bonus that the Principal must offer the Agent to induce her to work.

Lemma 3. At date zero in the second-best, if the Principal chooses a strictly positive level of investment, her bonus offer to the Agent is:

$$b_0(\lambda_0) = \begin{cases} c\lambda_0^{-1}\left(v_0^{-1} + \beta\left((v_1^p)^{-1} - v_0^{-1}\right)\right) & \text{if } \lambda_0 \in \left(0, \hat{\lambda}_0(v_0)\right) \\ c(\lambda_0 v_0)^{-1} & \text{if } \lambda_0 \in \left[\hat{\lambda}_0(v_0), 1\right] \end{cases} \tag{8}$$

The term in brackets in the middle line is a convex combination of the inverse of two beliefs: the prior (v_0) and the posterior conditioned on the public history in the event of failure (v_1^p). The convex weight is the discount factor, β: as the Agent cares more about the future, the value of holding out for the more favorable contract by shirking today increases. The difference $(v_1^p)^{-1} - v_0^{-1}$, can also be written as the likelihood ratio $\frac{1-v_1^p}{v_1^p}\lambda_0$, which increases in the Principal's initial investment, since on the equilibrium path higher initial investment makes failure more informative.

Let $g(\lambda_0; v_0)$ denote the Principal's problem at date zero under unobserved effort:

$$g(\lambda_0; v_0) \equiv v_0\lambda_0 W - \tau(\lambda_0) - \mathbf{1}[\lambda_0 > 0]c + (1 - v_0\lambda_0)\beta S(v_1^p) - \beta V(\lambda_0) \tag{9}$$

and let $G(v_0) = \max_{\lambda_0 \in [0,1]} g(\lambda_0; v_0)$. If the Principal's contract induces positive experimentation at date zero, then her pay-off $G(v_0)$ coincides with the corresponding pay-off that the Principal receives under the first-best with one important exception: when the Principal switches to an investment which foments a second attempt after initial failure, she must provide the Agent with an additional date zero incentive - her discounted continuation pay-off from shirking - to discourage her from manipulating the Principal's beliefs. This continuation payoff is itself an endogenous consequence of the Principal's investment. Let $\lambda_0^{SB}(v_0)$ denote the maximizer of

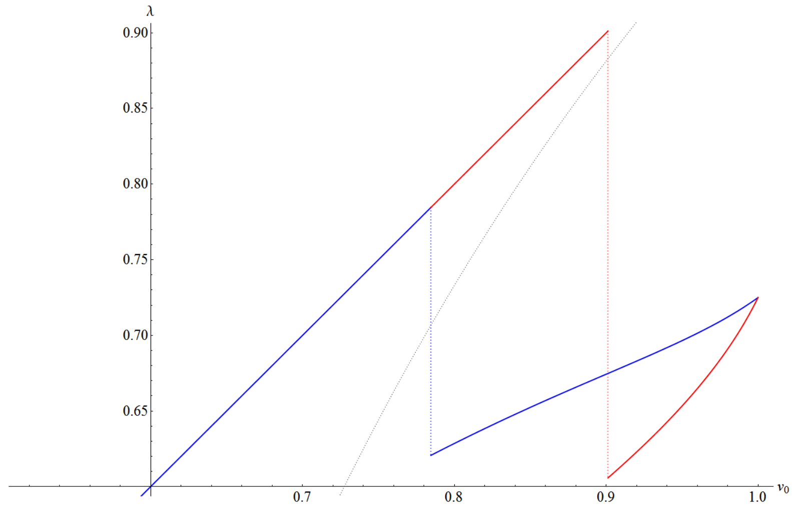

Figure 1: Illustration of the Principal's solution for λ_0 under *observed* (blue) and *unobserved* (red) effort, with $\tau(\lambda) = \frac{\lambda^2}{2}$. The dashed ray is $\hat{\lambda}_0(v_0)$. Parameter values are $\beta = \frac{3}{4}$, $W = 1$, $c = \frac{2}{15}$. Inefficient over-investment arises for $v_0 \in \Phi = (.785, .901]$. For $v_0 \leq .785$, the solutions coincide.

$g(\lambda_0; v_0)$. I refer to $\lambda_0^{SB}(v_0)$ as the second-best level of investment at date zero under short-term contracts.[5] My main result is:

Proposition 1. Whenever $\lambda_0^{FB}(v_0) \geq \hat{\lambda}_0(v_0)$, $\lambda_0^{SB}(v_0) = \lambda_0^{FB}(v_0)$. Moreover, there exists a non-empty set of prior beliefs about project quality, $\Phi \subset [0,1]$, such that $v_0 \in \Phi$ implies:

$$\lambda_0^{SB}(v_0) \geq \hat{\lambda}_0(v_0) > \lambda_0^{FB}(v_0) \tag{10}$$

For $v_0 \in \Phi$, the Principal invests too much relative the first-best, and destroys surplus by doing so. Nonetheless, such over-investment is superior for the Principal since her subsequent decision to discontinue the experiment after initial failure denies the Agent rents from the repeated interaction. In effect, she destroys surplus at *both* dates in order to capture a larger share at date zero.

This model is the first to propose over-investment relative to the efficient benchmark as a way to soften the Agent's incentive constraint - by contrast, existing work emphasizes under-investment as the second-best solution. The latter prescription is a necessity of a framework in which the Principal's investment costs are linear, such as Bergemann and Hege (2005). But Proposition 1 offers a second important contrast with existing work. In Bergemann and Hege (2005) and Hörner and Samuelson (2013), the Principal stops funding the experiment too soon, despite the fact that there is a positive surplus from continued investment.[6] In this model, by

[5]Notice that the Principal may prefer to invest nothing at date zero, then choose positive investment at date one, yielding a payoff $\beta f(\lambda_0^*(v_0); v_0)$. This possibility arises because the pre-specified end after date one in the game form provides some artificial commitment power to the Principal to quit at the end of date one. It does not play an important role in the analysis.

[6]Hörner and Samuelson (2013) also study non-Markov equilibria of their model, in which the Principal front-loads the Agent's effort in early periods until some threshold belief is attained, then switches to the worst equilibrium which is either funding with delay, or stopping experimentation. However, in the initial period in which front-loading takes place, it is the same investment that would arise in the first-best.

contrast, the Principal's decision not to provide funds at date one is socially optimal, conditional on her choice of investment at date zero. The inefficiency arises from the fact that this initial choice of investment destroys future surplus in order to avoid making transfers to the Agent at date zero.

Figure 1 illustrates an example; it shows optimal experimentation at date zero under the first-best (blue) and moral hazard (red).[7] The dashed black ray is $\hat{\lambda}_0(v_0)$. In the first-best, the Principal's optimal plan depends on her initial optimism. When v_0 is relatively low, it is sufficiently 'cheap' for her to experiment above the threshold which would induce her to quit, in the event of failure. When v_0 is sufficiently large, however, the Principal is better off spreading the cost of investment over both dates: it is too expensive to initially invest the funds required to generate a sufficiently informative failure that would lead her to quit after the initial round.

Note that in the second-best, when the Principal experiments at date zero in such a way as to induce a positive continuation pay-off from further experimentation, she nonetheless chooses an initial investment that is too low, relative to the first-best. In this case, she trades off the benefits of early discovery with the incentive cost that arises from giving the Agent a higher continuation pay-off from shirking - both of which are increasing in her initial investment. This inefficient under-investment is similar to Bergemann and Hege (2005).

4. Overinvesting with Commitment

For $v_0 \in \Phi$, the Principal's investment induces sufficiently pessimistic beliefs upon initial failure that she can credibly commit to undertake no further experimentation. This allows her to extract all of the surplus from the one-shot investment, despite the inefficiency. I now show that when the Principal can write long-term contracts, over-investment still arises - possibly in situations where it would not under short-term contracts - and inefficiency may increase.

I assume that the Principal can write a two-date contract specifying both investment and bonus decisions at both dates, at the start of date zero ("long-term commitment"). I let λ_t^{CO} and b_t^{CO} denote her optimal choice of investment and bonuses at each date under the contract with long-term commitment. First, I characterize the Principal's solution under long-term contracts. This characterization shows that under a range of circumstances, the Principal's investments and bonuses under the long- and short-term contracts coincide. Recall that $G(v_0)$ is the principal's value under the second-best short term contracts.

Lemma 4. The Principal's long-term contract always induces strictly positive investment at date zero. Moreover:

(i.) if $G(v_0) \geq f(\lambda_0^*(v_0); v_0)$: investments and bonuses in the short- and long-term contracts coincide.

(ii.) If $G(v_0) < f(\lambda_0^*(v_0); v_0)$: $\lambda_0^{CO}(v_0) = \lambda_0^*(v_0)$, $\lambda_1^{CO} = 0$, $b_0^{CO} = c\left(v_0\lambda_0^{CO}\right)^{-1}$, $b_1^{CO} = 0$.

[7] In this, and all other examples, all approximations are to three decimal places. Note that a closed-form expression for $\lambda_0^{SB}(v_0)$ cannot be obtained even with quadratic costs.

When (i) applies, the Principal's choices replicate those made under the sequence of short-term contracts. Intuitively, if she invests only at date zero under the short-term contract and her unconstrained solution already induces her to quit after initial failure ($\lambda_0^*(v_0) \geq \hat{\lambda}_0(v_0)$), the Principal can do no better than replicate this choice under the long term contract, if she also wishes to make a one-off investment. If, instead, she invests at both dates under the short-term contract, her date one bonus already holds the Agent to her participation constraint, and raising this or the date zero bonus under long-term commitment brings her no additional benefit in incentives. Since bonuses are the same, the Principal's investment problem is also the same.

Only when (ii) applies will the Principal's solutions differ according to whether short- or long-term contracts can be written. In this case, the one-off investment solution induces a positive continuation payoff from further investment at date one:

$$\lambda_0^*(v_0) < \hat{\lambda}_0(v_0) \tag{11}$$

but the Principal would prefer to make this one-off investment and quit the project after failure, rather than invest at both dates and offer the Agent the associated incentives:

$$\max_{\lambda_0 \leq \hat{\lambda}_0(v_0)} [f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta S(v_1^p) - \beta V(\lambda_0)] < f(\lambda_0^*(v_0); v_0) \tag{12}$$

Under short-term commitment, this plan cannot be implemented and instead her date zero investment solution satisfies:

$$\lambda_0^{SB}(v_0) \in \left\{ 0, \hat{\lambda}_0(v_0), \arg \max_{\lambda_0 \in [0, \hat{\lambda}_0(v_0)]} f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta S(v_1^p) - \beta V(\lambda_0) \right\} \tag{13}$$

With commitment, by contrast, the Principal can implement $\lambda_0^{CO} = \lambda_0^*(v_0)$, $\lambda_1^{CO} = 0$ and hold the Agent to her date zero participation constraint. So under condition (ii), if the Principal chooses $\lambda_0^{SB}(v_0) = \hat{\lambda}_0(v_0)$ in the absence of formal commitment power, she is using this investment explicitly as a form of belief commitment. This clarifies the role that investment plays solely as a form of belief commitment in the setting with short-term contracts.

Let Ψ denote the set of prior beliefs about project quality for which $\lambda_0^*(v_0) \geq \hat{\lambda}_0(v_0)$. We obtain the following result.

Proposition 2. If $v_0 \in \Phi \cap \Psi$, $\lambda_0^{CO}(v_0) \geq \hat{\lambda}_0(v_0) > \lambda_0^{FB}(v_0)$.

$\Phi \cap \Psi$ is the subset of prior beliefs for which *both* inefficient over-investment arises under the short-term contract *and* for which the unconstrained one-shot investment solution $\lambda_0^*(v_0)$ lies above $\hat{\lambda}_0(v_0)$. Since the short- and long-term contracts coincide for these prior beliefs, over-investment necessarily arises under the latter for these parameters. However, $\Phi \cap \Psi \neq \emptyset$ is not necessary for over-investment to arise. To see this, suppose the unconstrained one shot investment decision induces a positive continuation payoff, i.e. $\lambda_0^*(v_0) < \hat{\lambda}_0(v_0)$. If $\lambda_0^*(v_0) > \lambda_0^{FB}(v_0)$, and the Principal chooses to invest $\lambda_0^*(v_0)$ with commitment, inefficient over-investment arises with commitment even though it leads to a positive continuation payoff from experimenting. In fact, it is possible (and holds in an example that I present, shortly) that for some prior beliefs, we can have the ordering:

$$\hat{\lambda}_0(v_0) > \lambda_0^{CO}(v_0) = \lambda_0^*(v_0) > \lambda_0^{FB}(v_0) > \lambda_0^{SB}(v_0) \tag{14}$$

That is: the Principal *underinvests* in the second best and *overinvests* with long-term commitment but in both cases generates a positive continuation payoff in the event of failure.

The question arises as to the welfare consequences of moving from short- to long-term contracts. Clearly, the Principal is better off under long- than short-term commitment. Moreover, whenever $v_0 \in \Phi \setminus \Psi$, the surplus must also increase under the long-term contract, since the difference in surplus between the long- and short-run contracts is $f(\lambda^*(v_0); v_0) - f(\hat{\lambda}_0(v_0); v_0) > 0$.

Suppose, however, that under short-term contracts, at some prior belief, the Principal's date zero investment generates a positive continuation payoff at date one. That is, her investment decision is:

$$\lambda_0^{SB}(v_0) = \arg \max_{\lambda_0 \leq \hat{\lambda}_0(v_0)} \{f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta S(v_1^p(\lambda_0)) - \beta V(\lambda_0)\} \tag{15}$$

and the surplus is:

$$f(\lambda_0^{SB}(v_0); v_0) + (1 - v_0\lambda_0^{SB})\beta S(v_1^p(\lambda_0^{SB})) \tag{16}$$

since $\beta V(\lambda_0^{SB})$ is a transfer from the Principal to the Agent. If, under long-term commitment, the Principal chooses to invest $\lambda_0^{CO} = \lambda_0^*(v_0) < \hat{\lambda}_0(v_0)$ and $\lambda_1^{CO} = 0$, the surplus is instead:

$$f(\lambda_0^*(v_0); v_0) \tag{17}$$

It is easy to construct examples (see below) in which (16) exceeds (17). Intuitively: the Principal is able to appropriate the full share of instantaneous surplus created by her one-off investment decision under long-term commitment - that is: she fully appropriates (17). But this solution now leaves a positive continuation payoff on the table which is claimed by neither player. If, instead, the Principal invests over both dates, she hands a portion of the date zero surplus to the Agent, but appropriates all of the date one surplus. If the difference of (16) and (17) is positive, but less than $\beta V(\lambda_0^{SB})$, long-term commitment power for the Principal worsens the inefficiency associated with the agency problem. This is because formal commitment power allows the Principal to credibly destroy even more of the surplus associated with the two-date investment solution in order to deny rents to the Agent at date zero.

The following example illustrates this effect with quadratic investment costs, and also shows that over-investment may occur under the long-term contract for parameter values where under-investment occurs under the short-run contract, i.e. parameters for which (14) holds.

Example 1. Let $c = \frac{1}{15}$, $W = 1$, $\beta = \frac{1}{2}$ and $\tau(\lambda) = \frac{1}{2}\lambda^2$. We have $\underline{v} = \sqrt{\frac{2}{15}}$.

(i.) In the first-best, the Principal invests $\lambda_0^*(v_0) \geq \hat{\lambda}_0(v_0)$ if $v_0 \in (\underline{v}, .538]$, otherwise she invests positively at both dates when $v_0 \in (.538, 1]$.

(ii.) In the second-best under short-term commitment, the Principal invests $\lambda_0^*(v_0) \geq \hat{\lambda}_0(v_0)$ if $v_0 \in (\underline{v}, .575]$. On the interval $[.575, .736]$, $\lambda_0^*(v_0) < \hat{\lambda}_0(v_0)$, but the Principal prefers to invest $\lambda_0^{SB}(v_0) = \hat{\lambda}_0(v_0)$ in order to buy commitment to quit, rather than invest over both periods and hand the agent $\beta V(\lambda_0)$. Finally, on the interval $(.736, 1]$ she invests over both periods.

(iii) In the second-best under long-term commitment, the Principal's investment decisions are the same as (ii) except for $v_0 \in [.575, .766]$, on which she selects $\lambda_0^{CO} = \lambda^*(v_0) < \hat{\lambda}_0(v_0)$ and quits after initial failure.

When effort is not observed, surplus is higher under the long-term commitment contract when $v_0 \in [.575, .736]$ and under the short-term commitment contracts when $v_0 \in [.736, .766]$. Notice also that on the interval $[.736, .766]$, the Principal over-invests under long-term commitment when she under-invests with short-term commitment; in both cases, she generates a positive surplus from experimentation at date one, since $\lambda_0^*(v_0) < \hat{\lambda}_0(v_0)$, but with long-term commitment this surplus is left unclaimed. This corresponds to the ordering given in expression (14).

5. Concluding Observations

Incentives to over-invest arise from two features of the model. First, there exist beliefs about project quality for which the Principal would prefer to abandon the project. This gives her the ability to make initial investments which imply no further experimentation upon early failure. Second, the Agent's rents are inter-temporal: even when effort is unobserved, the Principal can hold the Agent to her participation constraint when she invests at date zero and quits after initial failure. This gives the Principal a motive to distort investments in order to capture the surplus.[8]

The results have application to project evaluation: observing that a project was initially given a large budget but that it was discontinued at an early stage without success may constitute evidence that the project was poorly conceived or badly managed. However, this analysis suggests another interpretation: the early release of large volume of funds followed by early abandonment may represent the solution of an incentive problem arising from the Agent's desire to prolong the life of the contract, in order to capture information rents.

In my analysis, the Principal has *either* intra-period *or* inter-period commitments with respect to *both* investment *and* bonus offers. In the Appendix, I study an 'intermediate' commitment environment in which, at date zero, the Principal can commit to offer the Agent a fixed (across dates) bonus which is awarded at whichever date a success is generated, but in which she possesses only intra-period commitment with respect to her investment decisions.[9] I show that over-investment must still arise, relative to the first-best. A partial intuition for this result is that since the date zero bonus can no longer be adjusted upward at date one, it must instead be adjusted upward at date zero in anticipation of the subsequent participation constraint faced by the Agent who is more pessimistic in the future. This is in addition to the incentive constraints for not shirking at date zero. If, on the other hand, the Principal invests in such a way as to induce discontinuation after failure at date zero, she need only compensate the Agent for her date zero participation, as in the earlier analysis.

[8] Over-investment can also arise when the Principal's investment is not observed by the Agent, so that there is scope for the former to manipulate the latter's beliefs.

[9] I am grateful to Dan Bernhardt for encouraging me to examine this setting. The assumption of intra-period commitment on investment decisions is not important to the result.

6. Appendix

Proof of Lemma 1

To prove (i), I show $\lambda_1^*(v_1^p)$ increases in v_1^p and v_1^p decreases in λ_0. Fix $v_1^p = v \in (0,1)$ and $\lambda_1^*(v)$, which I shorthand λ. Take $v' > v$ and associated solution $\lambda_1^*(v') = \lambda'$. Since λ' is optimal for v' and λ is optimal for v:

$$v'(\lambda' - \lambda)W + \tau(\lambda) - \tau(\lambda') \geq 0 \tag{18}$$

$$v(\lambda - \lambda')W + \tau(\lambda') - \tau(\lambda) \geq 0 \tag{19}$$

and combining yields:

$$(v - v')(\lambda - \lambda') \geq 0$$

so $v < v'$ implies $\lambda \leq \lambda'$. Next, $\lambda_0 > 0$ if and only if $e_0 = 1$, in a solution, so $v_1^p(h_1^p) = \frac{v_0(1-\lambda_0)}{1-v_0\lambda_0}$ which strictly decreases in λ_0. This completes the proof of (i).

I now prove (ii). Note that $S(v_1^p) = \max\{0, f(\lambda_1^*(v_1^p); v_1^p)\}$. It is easy to show $f(\lambda_1^*(v_1^p); v_1^p)$ is strictly increasing in v_1^p. Thus, $S(v_1^p) \geq 0$ implies $S(\hat{v}_1^p) > S(v_1^p)$ for $\hat{v}_1^p > v_1^p$. Next, $f(\lambda_1^*(0); 0) < 0$ and $f(\lambda_1^*(1); 1) > 0$. So, there is a unique $v_1^p \in (0,1)$ such that (a) $S(v_1^p) = f(\lambda_1^*(v_1^p); v_1^p) = 0$ and (b) for any $(v_1^p)' > v_1^p$, $S((v_1^p)') > 0$, which I call $\underline{v}$.

Proof of Proposition 1

Define $\Theta = \{v_0 : \lambda_0^{FB}(v_0) \geq \hat{\lambda}_0(v_0)\}$. First, I show $\Theta \neq \emptyset$. We have $f(\lambda_0^*(v_0); v_0)$ is continuous and strictly increasing in v_0. By construction, $f(\lambda_0^*(\underline{v}); \underline{v}) = 0$, $\hat{\lambda}_0(\underline{v}) = 0$ and for all $\lambda_0 \geq 0$:

$$f(\lambda_0; \underline{v}) + \beta(1 - \underline{v}\lambda_0) f(\lambda_1^*(v_1^p); v_1^p) < 0 \tag{20}$$

Finally, $\lambda_0^*(\underline{v}) > 0 = \hat{\lambda}_0(\underline{v})$. Continuity and strict monotonicity of $\lambda_0^*(v_0) > \hat{\lambda}_0(v_0)$ implies that there exists $\delta > 0$ and $B_+^\delta(v_0) \equiv \{v_0' : 0 < v_0' - v_0 < \delta\}$ such that $v_0' \in B_+^\delta(v_0)$ implies $f(\lambda_0^*(v_0'); v_0') > 0$, $\lambda_0^*(v_0') > \hat{\lambda}_0(v_0')$ and $\max_{\lambda_0} [f(\lambda_0; v_0') + \beta(1 - v_0'\lambda_0) f(\lambda_1^*(v_1^p); v_1^p)] < f(\lambda_0^*(v_0'); v_0')$. Thus, Θ is non-empty.

Next, I show $v_0 \in \Theta$ implies $\lambda_0^{SB}(v_0) = \lambda_0^{FB}(v_0)$. Suppose $v_0 \in \Theta$. Then

$$\begin{aligned} f(\lambda_0^*(v_0); v_0) &\geq \max_{\lambda_0 \in [0, \hat{\lambda}_0(v_0)]} [f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta f(\lambda_1^*(v_1^p); v_1^p)] \\ &> \max_{\lambda_0 \in [0, \hat{\lambda}_0(v_0)]} [f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta f(\lambda_1^*(v_1^p); v_1^p) - \beta V(\lambda_0)] \end{aligned} \tag{21}$$

Since $\lambda_0^*(v_0) \geq \hat{\lambda}_0(v_0)$ for $v_0 \in \Theta$, the Principal can attain the payoff $f(\lambda_0^*(v_0); v_0)$ in the second-best. Thus, $\lambda_0^{FB}(v_0) = \lambda_0^{SB}(v_0)$. So, whenever the Principal's investment at date zero satisfies $\lambda_0^{FB}(v_0) = \lambda_0^*(v_0) \geq \hat{\lambda}_0(v_0)$, we have $\lambda_0^{SB}(v_0) = \lambda_0^{FB}(v_0)$.

Next, I show $\Phi \equiv \{v_0 : \lambda_0^{SB}(v_0) \geq \hat{\lambda}_0(v_0) > \lambda_0^{FB}(v_0)\}$ is non-empty. Observe that $\hat{\lambda}_0(v_0)$ is strictly concave, and satisfies $\hat{\lambda}_0(\underline{v}) = 0$ and $\hat{\lambda}_0(1) = 1$; moreover, $\lambda_0^*(v_0)$ is continuous, strictly increasing and satisfies $\lambda_0(\underline{v}) > 0$ and $\lambda_0^*(1) < 1$. Thus, there exists a largest v_0 such that $\lambda_0^*(v_0)$ induces $f(\lambda_1^*(v_1^p); v_1^p) \leq 0$. In turn, there exists a largest v_0 such that $\lambda_0^*(v_0) \geq \hat{\lambda}_0(v_0)$

and $f(\lambda_0^*(v_0); v_0) \geq \max_{\lambda_0 \leq \hat{\lambda}_0(v_0)} [f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta f(\lambda_1^*(v_1^p); v_1^p)]$, which I denote v_0^A. For all $v_0 > v_0^A$, $\lambda_0^{FB}(v_0) < \hat{\lambda}_0(v_0)$. Suppose, first, $\lambda_0^*(v_0^A) > \hat{\lambda}_0(v_0^A)$; then:

$$f(\lambda_0^*(v_0^A); v_0^A) > \max_{\lambda_0 \in [0, \hat{\lambda}_0(v_0)]} \left[f(\lambda_0; v_0^A) + (1 - v_0^A\lambda_0)\beta f(\lambda_1^*(v_1^p); v_1^p) - \beta V(\lambda_0)\right] \tag{22}$$

and since the LHS is continuous and increasing, and $\lambda_0^*(v_0)$ is continuous, there exists $\delta > 0$ such that $v_0' \in B_+^\delta(v_0^A)$ implies $\lambda_0^{FB}(v_0') < \hat{\lambda}_0(v_0') < \arg\max_{\lambda_0 \in [\hat{\lambda}_0(v_0'), 1]} f(\lambda_0; v_0') = \lambda_0^{SB}(v_0')$. Suppose, instead, $\lambda_0^*(v_0^A) = \hat{\lambda}_0(v_0^A)$. Define:

$$\epsilon_1(v_0) \equiv f(\hat{\lambda}_0(v_0); v_0) - \max_{\lambda_0 \in [0, \hat{\lambda}_0(v_0)]} [f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta f(\lambda_1^*(v_1^p); v_1^p) - \beta V(\lambda_0)] \tag{23}$$

and $\epsilon_2(v_0) \equiv \max_{\lambda_0 \geq \hat{\lambda}_0(v_0)} f(\lambda_0; v_0) - \beta \max_{\lambda_0 \in [0,1]} f(\lambda_0; v_0)$ and note that $\epsilon_1(v_0^A) > 0$ by construction and $\epsilon_2(v_0^A) > 0$ for any $\beta < 1$, since $\hat{\lambda}_0(v_0^A) = \lambda_0^*(v_0^A)$ by supposition and so $f(\hat{\lambda}_0(v_0^A); v_0^A) - f(\lambda_0^*(v_0^A); v_0^A) = 0$. So, there exists $\delta' > 0$ such that $v_0' \in B_+^{\delta'}(v_0^A) \equiv \{v_0' : 0 < v_0' - v_0 < \delta\}$ implies $\epsilon_1(v_0') > 0$, and $\epsilon_2(v_0') > 0$. Thus, $\lambda_0^{SB}(v_0') \geq \hat{\lambda}_0(v_0') > \lambda_0^{FB}(v_0')$. So, Φ is non-empty, as was to be shown.

Proof of Lemma 4

First, I show $\lambda_0^{CO} > 0$. Suppose, to the contrary, $\lambda_0^{CO} = 0$. Since $v_0 > \underline{v}$, we must have $\lambda_1^{CO} = \lambda_1^*(v_0)$, $b_1 = \frac{c}{v_0\lambda_1^*(v_0)}$, and the Principal's payoff at date zero is $\beta f(\lambda_0^*(v_0); v_0)$. But by bringing forward both this investment choice and bonus to date zero, and pre-committing to $\lambda_1 = 0$ and $b_1 = 0$, the Principal can attain $f(\lambda_0^*(v_0); v_0)$. Thus, $\lambda_0^{CO} > 0$.

Suppose the Principal chooses a long-term contract having the property $\lambda_0 > 0$ and $\lambda_1 = 0$. Then, her optimal contract is plainly $\lambda_0^{CO} = \lambda_0^*(v_0)$, $\lambda_1^{CO} = 0$, $b_0^{CO} = c(v_0\lambda_0^{CO})^{-1}$ and $b_1^{CO} = 0$, since $\lambda_0^*(v_0) = \arg\max_{\lambda \in [0,1]} f(\lambda_0; v_0)$, and $v_0 > \underline{v}$ implies $f(\lambda_0^*(v_0); v_0) > 0$.

Suppose, instead, the Principal chooses a long-term contract satisfying $\lambda_0 > 0$ and $\lambda_1 > 0$. This contract solves:

$$\max_{\lambda_0, \lambda_1, b_0, b_1} [v_0\lambda_0(W - b_0) - \tau(\lambda_0) + (1 - v_0\lambda_0)\beta(v_1^p\lambda_1(W - b_1) - \tau(\lambda_1))] \tag{24}$$

subject to:

$$b_1 \geq c(v_1^p\lambda_1)^{-1} \tag{25}$$

$$v_0\lambda_0 b_0 - c + (1 - v_0\lambda_0)\beta(v_1\lambda_1 b_1 - c) \geq \beta\,(v_0\lambda_1 b_1 - c) \tag{26}$$

We can re-write the date two constraint as $b_1 = c(v_1^p\lambda_1)^{-1} + \delta_1$ subject to $\delta_1 \geq 0$ and thus the date zero IC constraint may be written:

$$b_0 = \frac{c}{v_0\lambda_0} + \beta\frac{c(1 - v_0)}{v_0\lambda_0(1 - \lambda_0)} + \beta\delta_1\lambda_1 + \delta_0 \tag{27}$$

subject to $\delta_0 \geq 0$. Substituting into the Principal's objective function, we obtain:

$$\max_{\lambda_0, \lambda_1, \delta_0, \delta_1} [f(\lambda_0; v_0) + \beta(1 - v_0\lambda_0)f(\lambda_1; v_1^p) - \beta V(\lambda_0) - v_0(\delta_0\lambda_0 + \beta\delta_1\lambda_1)] \tag{28}$$

subject to $\delta_0, \delta_1 \geq 0$, and where $V(\lambda_0)$ is defined, earlier. Thus: $\delta_0 = \delta_1 = 0$ and her investment decision under the supposition $\lambda_0 > 0$, $\lambda_1 > 0$ solves the same problem as the second-best with short-term commitment, and thus $\lambda_1^{CO} = \lambda_1^*(v_1^p)$ and $\lambda_0^{CO} = \arg\max_{\lambda \in [0,1]} f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta S(v_1^p) - \beta V(\lambda_0)$. Thus, the Principal's value from the contract under long-term commitment is:

$$L(v_0) \equiv \max \left\{ \max_{\lambda_0 \in [0,1]} f(\lambda_0; v_0), \max_{\lambda_0 \in [0, \hat{\lambda}_0(v_0)]} [f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta f(\lambda_1^*(v_1^p); v_1^p) - \beta V(\lambda_0)] \right\} \tag{29}$$

whereas her value from the contract under short-term commitment is:

$$G(v_0) \equiv \max \left\{ \max_{\lambda_0 \in [0,1]} \beta f(\lambda_0; v_0), \max_{\lambda_0 \in [\hat{\lambda}_0(v_0), 1]} f(\lambda_0; v_0), \right. \tag{30}$$

$$\left. \max_{\lambda_0 \in [0, \hat{\lambda}_0(v_0)]} [f(\lambda_0; v_0) + (1 - v_0\lambda_0)\beta f(\lambda_1^*(v_1^p); v_1^p) - \beta V(\lambda_0)] \right\} \tag{31}$$

Suppose, then, $G(v_0) \geq f(\lambda_0^*(v_0); v_0)$. Then, $G(v_0) = L(v_0)$ and the long-term and short-term contracts coincide. Suppose, instead, $G(v_0) < f(\lambda_0^*(v_0); v_0)$. Then, $G(v_0) < L(v_0) = f(\lambda_0^*(v_0); v_0)$. This implies $\lambda_0^{CO} = \lambda_0^*(v_0)$, $\lambda_1^{CO} = 0$, $b_0^{CO} = c(v_0\lambda_1^{CO})^{-1}$ and $b_1^{CO} = 0$, as was to be shown.

Proof of Proposition 2

This is an immediate consequence of the previous Lemma.

Time Independent Fixed Bonuses

This section expands on the final paragraph of the Concluding Comments. Suppose that the Principal can commit only within each date to investment choices, but from the start of the game can commit to offer the Agent a single bonus, $b_0 = b_1 = b$, which is awarded to the Agent at whichever (if any) date the surplus is generated. The Principal continues to make her investment decisions at the start of each date.

Suppose, first, the Principal's contract induces her to select $\lambda_0 > 0$ and $\lambda_1 > 0$. Then, her date one problem, given bonus $b > 0$, solves:

$$\max_{\lambda_1 \in [0,1]} v_1^p \lambda_1 (W - b) - \tau(\lambda_1) \tag{32}$$

Let $\tilde{\lambda}_1(v_1^p, b)$ denote the solution to this problem. We next obtain constraints on the Principal's offer b at date zero, given investment decision λ_0 and the induced subsequent investment decision $\tilde{\lambda}_1(v_1^p, b)$. First, to induce effort at date one, we require $b \geq \frac{c}{v_1^p \tilde{\lambda}_1(v_1^p, b)}$. Second, to induce effort at date zero, we need:

$$v_0\lambda_0 b - c + (1 - v_0\lambda_0)\beta(v_1^p \tilde{\lambda}_1(v_1^p, b)b - c) \geq \beta(v_0\tilde{\lambda}_1(v_1^p, b)b - c) \tag{33}$$

and together, this yields the constraint on the bonus, b:

$$b \geq \max \left\{ \frac{c}{v_1^p \tilde{\lambda}_1(v_1^p, b)}, c\frac{(v_0\lambda_0)^{-1} - \beta}{1 - \beta\tilde{\lambda}_1(v_1^p, b)} \right\} \equiv b_{min}(v_0, \lambda_0, \tilde{\lambda}_1(v_1^p, b)) \tag{34}$$

First, I claim that there exists $\check{\lambda}_0(v_0)$ such that the conjecture $\lambda_1 > 0$ is true only if $\lambda_0(v_0) < \check{\lambda}_0(v_0)$ for any $b \geq b_{min}$. This follows from the fact that the Principal's continuation payoff after failure at date zero and given the bonus b specified at date 0 is:

$$H(v_1^p) \equiv \max\left\{0, v_1^p\tilde{\lambda}_1(v_1^p, b)W - \tau(\tilde{\lambda}_1(v_1^p, b)) - v_1^p\tilde{\lambda}_1(v_1^p, b)b\right\} \leq \max\{0, f(\lambda_1^*(v_1^p); v_1^p)\} \quad (35)$$

where the second inequality follows from the fact that $b \geq b_{min}$ implies $v_1^p\tilde{\lambda}_1(v_1^p, b)b \geq c$. Thus, $S(v_1^p) = 0$ implies $H(v_1^p) = 0$. Since $\lambda_0 \geq \hat{\lambda}_0(v_0)$ implies $S(v_1^p) = 0$, there exists $\check{\lambda}_0(v_0) \leq \hat{\lambda}_0(v_0)$ such that $\lambda_0 \geq \check{\lambda}_0(v_0)$ implies $H(v_1^p) = 0$. Then, the Principal's value from investing $\lambda_0 \in (0, \check{\lambda}_0(v_0))$ (which follows by the supposition $\lambda_1 > 0$, $\lambda_0 > 0$) is:

$$H(v_0) \equiv \max_{\lambda_0 \in [0, \check{\lambda}_0(v_0)], b \geq b_{min}} \{v_0\lambda_0 W - \tau(\lambda_0) - v_0\lambda_0 b + (1 - v_0\lambda_0)\beta H(v_1^p)\} \quad (36)$$

Next, I claim:

$$v_0\lambda_0 b = c \max\left\{\frac{1 - \beta v_0\lambda_0}{1 - \beta\tilde{\lambda}_1(v_1^p)}, \frac{v_0\lambda_0}{v_1^p\tilde{\lambda}_1(v_1^p)}\right\} > c \quad (37)$$

We have $\frac{v_0\lambda_0}{v_1^p\tilde{\lambda}_1(v_1^p)} > 1$ if $\lambda_0 v_0 > v_1^p\tilde{\lambda}_1(v_1^p)$; if, instead, $\lambda_0 v_0 \leq v_1^p\tilde{\lambda}_1(v_1^p)$, then $v_0\lambda_0 < \tilde{\lambda}_1(v_1^p)$ since $\lambda_0 > 0$ implies $v_1^p < v_0 \leq 1$; thus $\frac{1-\beta v_0\lambda_0}{1-\beta\tilde{\lambda}_1(v_1^p)} > 1$. So:

$$H(v_0) < \max_{\lambda_0 \in [0, \check{\lambda}_0(v_0)]} \{v_0\lambda_0 W - \tau(\lambda_0) - c + (1 - v_0\lambda_0)\beta S(v_1^p)\} \quad (38)$$

$$\leq \max_{\lambda_0 \in [0, \hat{\lambda}_0(v_0)]} \{v_0\lambda_0 W - \tau(\lambda_0) - c + (1 - v_0\lambda_0)\beta S(v_1^p)\} \quad (39)$$

since the domain restriction for the first problem is $\lambda_0 \leq \check{\lambda}_0(v_0)$ and for the second is $\lambda_0 \leq \hat{\lambda}_0(v_0)$ and $\check{\lambda}_0(v_0) \leq \hat{\lambda}_0(v_0)$. So, the Principal's value at date zero is:

$$\begin{cases} f(\lambda_0^*(v_0); v_0) & \text{if } \lambda_0 \geq \check{\lambda}_0(v_0) \\ H(v_0) & \text{if } \lambda \in (0, \check{\lambda}_0(v_0)) \\ \beta f(\lambda_0^*(v_0); v_0) & \text{if } \lambda_0 = 0 \end{cases} \quad (40)$$

Let $\lambda_0^{TI}(v_0)$ denote the Principal's solution. Analogous arguments to the proof of Proposition 1 then imply (i) whenever $\lambda_0^*(v_0) = \lambda_0^{FB}(v_0) \geq \hat{\lambda}_0(v_0)$, $\lambda_0^{TI}(v_0) = \lambda_0^{FB}(v_0) \geq \hat{\lambda}_0(v_0) \geq \check{\lambda}_0(v_0)$, and there exists $\Lambda \in [0, 1]$ such that $\lambda_0^{FB}(v_0) < \hat{\lambda}_0(v_0) < \lambda_0^{TI}(v_0)$ for $v_0 \in \Lambda$.

References

Bergemann, Dirk and Ulrich Hege. 2005. "The financing of innovation: learning and stopping." *RAND Journal of Economics* pp. 719–752.

Bhaskar, V. 2012. "Dynamic Moral Hazard, Learning and Belief Manipulation." Working Paper.

Halac, Marina, Navin Kartik and Qingmin Liu. 2013. "Optimal Contracts for Experimentation." Working Paper.

Hörner, Johannes and Larry Samuelson. 2013. "Incentives for experimenting agents." *The RAND Journal of Economics* 44(4):632–663.

Manso, Gustavo. 2011. "Motivating innovation." *The Journal of Finance* 66(5):1823–1860.