

# Firm Investment and Stakeholder Choices: A Top-Down Theory of Capital Budgeting\*

Andres Almazan  
University of Texas

Zhaohui Chen  
University of Virginia

Sheridan Titman  
University of Texas and NBER

25 April 2010

## Abstract

This paper provides a model of the capital budgeting process of a firm whose top executives have private information. The success of the firm depends on the motivation and efforts of its lower level managers (and other stakeholders) who try to infer the information of the firm's top executives from their actions. In particular, higher levels of investment indicate lower level managers that the firm has promising prospects which induces them to take actions that contribute to the firm's success. In the presence of these effects, capital budgeting rules not only play their traditional allocative role but also affect the transmission of information from the top down. In this setting, there can arise a number of commonly observed investment distortions such as capital rationing, investment rigidities, overinvestment, and inflated discount rates.

---

\*We would like to thank Adolfo de Motta and David Dicks and seminar participants at DePaul University, UCLA, University of Essex, UNC-Chapel Hill, University of Notre Dame, University of Pennsylvania, University of Virginia and UT-Austin. Corresponding author: [andres.almazan@mcombs.utexas.edu](mailto:andres.almazan@mcombs.utexas.edu).

# 1 Introduction

The capital budgeting process has been described (i.e., Brealey, Myers and Allen 2008) as a combination of “bottom-up” procedures, where lower units solicit capital from headquarters, and “top-down” procedures where headquarters use their discretion to allocate capital downstream. An extensive literature has analyzed the incentive and information considerations that can emerge in “bottom-up” capital allocation processes in which headquarters allocate capital by reacting to solicitations from better informed lower units (e.g., Harris and Raviv 1996 or Bernardo et al. 2004). However, up to now, the literature has not focused on “top-down” processes in which capital allocation decisions by better informed headquarters convey information to lower units.<sup>1</sup>

In this paper, we model the firm’s capital budgeting process as a top-down procedure in which top executives have private information that is conveyed to lower level managers by the firm’s capital allocation. Our premise, that headquarters have relevant information not possessed by downstream units, can be motivated in several ways. For instance, by aggregating information from lower level managers, top-executives acquire broader information that is useful for their strategic planning. When this is the case, the amount of capital that the executives allocate to various projects conveys information to lower level managers, and the procedures that firms use to evaluate investments can influence how this information is interpreted. Alternatively, better information at the top can simply reflect access to better sources of information e.g., connections with other CEOs, consulting firms or politicians. Whatever the reason, as we emphasize, the information conveyed by top managers in their capital allocation choices is especially important when such information influences the decisions of lower level managers and other stakeholders of the firm.

To illustrate these strategic considerations, consider the capital allocation decision of a large integrated oil company, such as British Petroleum (BP). For a firm like this, a major investment in biofuels is likely to be viewed as an indication that the firm’s top management has favorable information about the future prospects of these alternative sources of energy or, alternatively, that it has serious concerns on the future of the more traditional sources of

---

<sup>1</sup>Supporting the importance of top-down information transmission, Grinstein and Tolowsky (2004) provide evidence that suggests that directors of S&P 500 firms act to “alleviate conflicts of interest between agents and principals and to communicate principals’ information to agents.”

energy.<sup>2</sup> A central premise of our paper is that information conveyed by investment choices can have an important influence on the stakeholders' choices. For instance, employees in BP's biofuel division will be encouraged by the implications of the higher investment and will work harder. Indeed, this incentive to exert effort will be especially strong if employees believe that capital is tight and the firm is using very high hurdle rates to evaluate their investments.

To better understand how capital budgeting procedures and lower unit perceptions interact we develop a simple model of an entrepreneurial firm whose production process requires a capital expenditure (i.e., an investment) by the firm's owner (i.e., the entrepreneur) as well as effort exerted by a lower level manager (i.e., an employee). Specifically, we consider a framework where the entrepreneur has private information about the firm's prospects, chooses the level of investment, and promises the employee a bonus based on the firm's output. In this setting both the firm's investment and the employee's effort are more productive when the future prospects of the firm are better. As a result, with full information, it is optimal for the firm to invest more and for the employee to exert more effort when the firm's future prospects are more favorable. This in turn implies that when the entrepreneur has private information, there is an incentive for the entrepreneur to overinvest relative to the full information case, because doing so elicits additional effort from the employee.

Most of the intuition of our model can be illustrated in a setting in which the firm's prospects are one of two types, high or low and in which the firm can commit to an investment policy before observing the private information about its prospects. Within this setting, two investment policies can emerge as optimal: (i) a separation policy in which a high prospect firm invests more than a low prospect firm and (ii) a pooling policy in which firms commit to a maximum investment level regardless of their types. With a separation policy, the firm will tend to over-invest because of the benefits associated with conveying favorable information. This cost of separation is offset by the benefits of investing more when the marginal productivity of capital is higher. Hence, ex-ante, the choice between the pooling and the separation policy is determined by a trade-off between the efficiency gain associated with having an investment policy which incorporates the entrepreneur's

---

<sup>2</sup>In June 2007, after BP announced plans to invest \$500M in biofuels, analysts at Global Insight stated that it was the "*first time BP has committed itself to such a large investment in biofuels research*". (See [www.ihsglobalinsight.com/SDA/SDADetail6147.htm](http://www.ihsglobalinsight.com/SDA/SDADetail6147.htm)).

information and the efficiency loss associated with over-investment.

Having established the incentive to overinvest and the potential benefits of committing to a fixed level of investment we enrich the previous framework by adding a third party to our model. In addition to the entrepreneur, who obtains information, and the employee who exerts effort, we consider a third party that sets capital budgeting policies and offers a compensation contract to the entrepreneur. One interpretation is that the third party is a venture capitalist that provides funding for the firm and retains some control of its operations. Another is that the informed party is the manager of the division of a conglomerate and the third party is either the executives at the firm's headquarters or the firm's board.

We investigate a number of additional issues within this three-party setting. First, we consider the possibility that the board can force the entrepreneur to commit to using a discount rate that may not equal the rate that would be used with symmetric information. This could be a single discount rate that is independent of the magnitude of the investment or it could be a function of the investment's magnitude. Second, we consider how the firm's capital structure can be designed by the board to reduce inefficiencies in the capital budgeting process.

The analysis of the three-party model demonstrate that several commonly observed capital budgeting practices can enhance firm value. Specifically, the imposition by the board of high discount rates can change the opportunity cost of capital for the entrepreneur and hence offset the incentive to overinvest that arises when the level of investment conveys information to stakeholders.<sup>3</sup> Similarly, by using more debt financing, a Myers (1977) debt overhang problem can arise that offsets the incentive to overinvest.

As we mentioned at the outset, our analysis of a "top-down" capital allocation process is in contrast of the analysis of "bottom-up" process that is the focus of the existing literature. Specifically, this literature examines the issue of how a firm may distort its capital budgeting practices in order to induce managers to exert proper effort (Bernardo, Cai, and Luo 2001, 2004, 2006), to curb managers' empire building tendencies (e.g., Harris and Raviv 1996, 1997, Marino and Matsusaka 2004, and Berkovitch and Israel 2004) or to reveal their private information. This literature has also considered the trade-offs that arise in the decision to

---

<sup>3</sup>See Poterba and Summers (1995) who document the use of high-discount rates on American corporations and Meier and Tarhan (2007) who provide evidence of what they refer to as a "hurdle rate puzzle" i.e., the use of a discount rate that substantially exceeds estimates of their cost of capital.

delegate capital budgeting decisions to a better informed agent (e.g., Aghion and Tirole, 1997 and Burkart, Gromb, and Panunzi, 1997).

Our analysis also has implications that are similar to the agency literature that argues that since managers get private benefits from managing larger enterprises, firm may exhibit a tendency to overinvest (e.g., Jensen 1986, and Hart and Moore 1995). As we show, the tendency to overinvest, as well as procedures that curb this tendency, can also arise within a setting without managerial private benefits. Hence, in addition to offering a theory that can rationalize capital budgeting rigidities in large corporations, we offer an explanation for the high hurdle rates imposed by venture capitalists on the investment choices of young start-up firms.<sup>4</sup>

Although the setting is very different, our analysis is actually closest to Hermalin (1998) and Komai, Stegeman, and Hermalin (2007), who study the problem of leadership in organizations.<sup>5</sup> In these models, the leader is better informed than other team members and credibly communicates his favorable information, which motivates his subordinates to work harder, by exerting greater effort. There are, however, two key differences between our setting and the setting considered in the leadership papers. First, in our setting, the entrepreneur's costly action (i.e., the investment choice) is contractible, while the leader's effort in the leadership models is not. Indeed, the capital budgeting policies that we explore arise because of the contractibility of investment. Second, we introduce a third party that is not involved in the production process but which has the authority to set investment rules and to claim firm output. This allows us to depart from a team setting (where parties share output) and instead focus on a setting in which parties that contribute to the production process are not necessarily those claiming the output (i.e., a firm). As we show, the presence of a third party who acts as "budget breaker" can broaden the contract space, and thus provide the entrepreneur with incentives that lead to more efficient investment choices.<sup>6</sup>

---

<sup>4</sup>Indeed, private equity firms and venture capitalist take this tendency to evaluate investments with very high discount rates to an extreme, generally requiring "expected" internal rates of return on their new investments that exceed 25%. Gompers (1999) describes the "venture capital" valuation method, where discount rates of more than 50% per year are used.

<sup>5</sup>See also Benabou and Tirole (2003) who consider the signaling effect of providing explicit incentives to employees.

<sup>6</sup>Holmstrom (1982) first pointed out the value of third parties who by threatening to break the budget eliminate inefficiencies in team production. While in Holmstrom (1982), third parties act "off-equilibrium path" in our setting, the third party breaks the budget in equilibrium in order to improve entrepreneurial incentives to invest appropriately.

More generally, our paper belongs to the principal-agent literature with informed principals. Among other things, this literature considers the effects of the principal’s information on the optimal compensation contract (Beaudry 1994 and Inderst 2000), the value of the private information to the principal (Chade and Silvers 2004 and Karle 2009) and the incentives to disclose information (i.e., provide “advice”) to the agent (Strausz 2009). In contrast to these models, the principal in our model not only designs the agent compensation but also takes an action (i.e., makes an investment) that directly affects the firm’s production.

The rest of the paper is organized as follows. Section 2 presents the base model and Section 3 analyzes it. Section 4 considers a modified setting with a board of directors and considers a implementation of corporate budgeting practices and Section 5 presents our conclusions.

## 2 The model

We consider a firm that operates in a risk-neutral economy. The firm is run by an entrepreneur (“the principal”) and requires the input from a penniless lower-level manager (“the agent”) who is subject to limited liability and a zero reservation wage.<sup>7</sup> The technology of the firm is characterized by a (decreasing returns to scale) stochastic production function:

$$Q(e, \theta, k) = z(e, \theta)k - \frac{1}{2}k^2. \quad (1)$$

This function combines two inputs, a capital investment  $k$  (i.e., the firm’s scale) which is subject to a quadratic cost  $g(k) \equiv \frac{1}{2}k^2$  and the agent’s effort  $e$ , in order to produce the firm output  $q(e, \theta, k) = z(e, \theta)k$ .<sup>8</sup> Specifically, the firm’s random productivity

$$z(e, \theta) = \begin{cases} 0 & \text{with prob. } (1 - \theta e) \\ r > 0 & \text{with prob. } \theta e \end{cases} \quad (2)$$

features a probability of success that depends on (i) an exogenous random shock  $\theta$ , which is privately observed by the principal,

$$\theta = \begin{cases} 1 & \text{with prob. } (1 - \pi) \equiv p_1 \\ \beta > 1 & \text{with prob. } \pi \equiv p_\beta \end{cases}$$

---

<sup>7</sup>In this section we identify the entrepreneur (the principal) with the firm. In Section 4, however, we endogenize the entrepreneur’s payoffs in a setting in which the entrepreneur is compensated by the external owners of the firm.

<sup>8</sup>Alternatively, the firm’s production function can be expressed as  $Q(e, \theta, I) = z(e, \theta)\sqrt{2I} - I$ , i.e., the scale of the firm increases by a factor of  $\sqrt{2I}$  per unit of capital invested.

and (ii) the agent's privately exerted effort  $e \in [0, \frac{1}{\beta}]$ . The agent's cost of effort is given by:

$$h(e, k) = \frac{1}{2}ce^2k,$$

which is convex in  $e$  (i.e.,  $h_{ee} = c$ ) and increases linearly with the firm's scale (i.e.,  $h_k = \frac{1}{2}ce^2$ ).<sup>9</sup> In what follows, we implicitly assume that  $c$  is sufficiently large that the optimal effort choice lies within the bounds  $[0, \frac{1}{\beta}]$ .

The timing of events is as follows: At  $t = 0$ , before observing  $\theta$ , the principal offers the agent a compensation schedule contingent on the firm's scale and the realized output:  $w(k, q)$ . At  $t = 1$ , the principal privately observes  $\theta$  and chooses the scale  $k$ . At  $t = 2$ , the agent makes an unobservable effort choice  $e$ . At  $t = 3$ , firm output  $q$  is realized and contracts are settled. Figure 1 summarizes the timing of events.

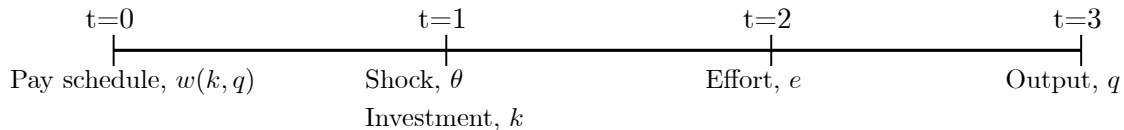


Figure 1: Timing of Events

We analyze the model under the assumption that the principal can commit to any policy at  $t = 0$  which is contingent on observable information afterwards. In other words, the principal can commit to any compensation schedule that is contingent on both the investment chosen at  $t = 1$  and realized output at  $t = 2$ . After we present the main results of the analysis, we consider the implications of relaxing this commitment assumption. That is, we consider how the results would change if the principal offers the agent a compensation contract after choosing the investment at  $t = 1$ .

### 3 Model Analysis

#### 3.1 Observable productivity shock $\theta$

As a benchmark we analyze the case in which the productivity shock  $\theta$  is observed by the agent as well as the principal but there is still managerial moral hazard (i.e., the agent makes an unobservable effort choice). We denote as  $k = \{k_1, k_\beta\}$  and  $e = \{e_1, e_\beta\}$  the principal's investment and agent's effort levels when  $\theta = \{1, \beta\}$  is realized. Without loss of

<sup>9</sup>An alternative formulation in which effort costs are independent of the firm scale but in which production costs are cubic on firm scale, i.e.,  $g(k) \equiv k^3$  is less tractable but produces similar results.

generality, we restrict the analysis to finding compensation contracts that offer non-negative payments when the output is high ( $q = rk$ ) and a zero-payment when it is low ( $q = 0$ ), i.e.,  $w = \{w_1, w_\beta\}$  when  $\theta = \{1, \beta\}$  is realized.<sup>10</sup> Therefore, the principal's problem can be expressed as

$$\max_{k, w, e} V = \sum_{\theta=\{1, \beta\}} p_\theta \left[ (rk_\theta - w_\theta)\theta e_\theta - \frac{1}{2}k_\theta^2 \right] \quad (3)$$

$$\text{s.t.} \quad e_\theta = \arg \max_e \{w_\theta \theta e - \frac{1}{2}ce^2 k_\theta\}, \quad \text{for } \theta = \{1, \beta\} \quad (4)$$

$$w_\theta \theta e_\theta - \frac{1}{2}ce_\theta^2 k_\theta \geq 0, \quad \text{for } \theta = \{1, \beta\}. \quad (5)$$

For each realization of the shock,  $\theta = \{1, \beta\}$ , the principal maximizes firm value (3), subject to the corresponding (endogenous) effort (4), and individual rationality constraints (5). The problem can be simplified because the agent's limited liability requires  $w_\theta \geq 0$  and, since  $e = 0$  is feasible, constraints (5) always hold. As a result, we can solve this problem by substituting the first-order condition of (4), (i.e.,  $e_\theta = \frac{w_\theta \theta}{k_\theta c}$ ) in (3),

$$\max_{k, w} V = \sum_{\theta=\{1, \beta\}} p_\theta \left[ (rk_\theta - w_\theta)\theta \frac{w_\theta \theta}{k_\theta c} - \frac{1}{2}k_\theta^2 \right], \quad (6)$$

and then deriving the f.o.c. of problem (6) to obtain the following proposition:

**Proposition 1** *For  $\theta = \{1, \beta\}$  the optimal compensation, investment and effort are:*

$$w_\theta^* = \frac{r^3 \theta^2}{8c}, \quad k_\theta^* = \frac{r^2 \theta^2}{4c} \quad \text{and} \quad e_\theta^* = \frac{r\theta}{2c}. \quad (7)$$

From Proposition 1, it follows that optimal compensation, investment and effort are increasing in  $\theta$ , i.e.,  $w_\beta^* > w_1^*$ ,  $k_\beta^* > k_1^*$  and  $e_\beta^* > e_1^*$ . The intuition for this result is straightforward: Since effort is more valuable when the firm's prospects are better, the principal will invest more and offer higher compensation to induce greater managerial effort when the firm has better prospects. It is illustrative to also describe the equilibrium relations among these endogenous variables. Specifically, one can check that (i) effort increases with managerial compensation and decreases with investment, i.e.,  $e_\theta^* = \frac{\theta w_\theta^*}{ck_\theta^*}$ , and (ii) compensation and investment are proportional to each other, i.e.,  $w_\theta^* = \frac{r}{2}k_\theta^*$ . Alternatively expressed,

---

<sup>10</sup>In the appendix we show that this is indeed the case.



this last observation implies that the optimal compensation is a sharing rule independent of  $\theta$ , i.e.,  $\frac{w_\theta^*}{rk_\theta^*} = \frac{1}{2}$ . For future comparisons, we compute firm value in the case when  $\theta$  is observable which is:

$$V^* = \frac{1}{2} \left[ (1 - \pi)k_1^{*2} + \pi k_\beta^{*2} \right]. \quad (8)$$

### 3.2 Unobservable productivity shock $\theta$

The feasibility of the contractual arrangement described in Proposition 1 crucially hinges on the fact that the agent can observe  $\theta$ . When this is not the case, the principal's investment  $k$  plays not only its usual allocating role but also conveys information about  $\theta$  to the agent, which in turn affects the agent's effort choice.

In a setting like this (with an informed principal and an agent who makes a costly effort choice) the contract that maximizes firm value need not be a direct mechanism, i.e., the revelation principle does not apply. Even though the principal can commit to any policy as a function of the disclosed  $\theta$ , the agent cannot commit to a specific effort choice since effort is unobservable to anyone but the agent himself.<sup>11</sup>

In the appendix we show that the optimal contract consists of two investment-compensation pairs, i.e.,  $\{k_a, w_a\}$  and  $\{k_b, w_b\}$ . Within this set of contracts three possibilities arise. If the principal's choices depend on the observed type, (i.e., the principal chooses  $\{k_a, w_a\}$  after observing  $\theta = 1$  and  $\{k_b, w_b\} \neq \{k_a, w_a\}$  after observing  $\theta = \beta$ ), such choices communicate  $\theta$  to the agent. We refer to this case as *separation*. If, however,  $\{k_b, w_b\} = \{k_a, w_a\}$  then the principal's choices convey no information. We refer to this case as *pooling*. A third possibility would be *partial pooling* where the principal mixes between these pairs.<sup>12</sup> In what follows, we search for the optimal contract under separation and pooling and compare firm value when these contracts are implemented. As we show in the appendix, this is without loss of generality since firm value under partial pooling is lower than firm value when either the optimal separation or the optimal pooling policy is implemented.

---

<sup>11</sup>See Bester and Strausz (2001) for an analysis of the optimal contract in a dynamic adverse selection setting where the principal cannot commit to a specific allocation after the agent's first period choice and Strausz (2009) for a setting in which, similarly to ours, the revelation principle lacks bite due to the unobservability of the agent's effort.

<sup>12</sup>Partial pooling would require us to solve for the optimal level of information revelation by the principal. Formally, this requires defining  $\{\sigma_a, \sigma_b\}$  which are the probabilities with which the principal choose  $(k_a, w_a)$  and  $(k_b, w_b)$  after observing  $\{1, \beta\}$ .

### 3.3 Separation

With separation, the principal's choices consist of two distinct managerial compensations  $w^s = \{w_1^s, w_\beta^s\}$  and investment levels  $k^s = \{k_1^s, k_\beta^s\}$  which depend on  $\theta$ . Since  $\theta$  is unobservable to the agent, the principal's choices are subject to two incentive compatibility (IC) constraints. We define  $V_\theta^{\hat{\theta}}$  as the firm value when the principal observes  $\theta$  and chooses  $w_\theta^s$  and  $k_\theta^s$  for  $\hat{\theta}, \theta = \{1, \beta\}$ :

$$V_\theta^{\hat{\theta}} \equiv (rk_\theta^s - w_\theta^s)\theta e_\theta^s - \frac{1}{2}k_\theta^{s^2} \quad (9)$$

where  $e_\theta^s = \arg \max_e \{w_\theta^s \hat{\theta} e - k_\theta^s c e^2\}$  for  $\hat{\theta} = \{1, \beta\}$  and  $e^s = \{e_1^s, e_\beta^s\}$ . The principal's problem can be written as:

$$\max_{k^s, w^s, e^s} V^s = \sum_{\theta=\{1, \beta\}} p_\theta \left[ (rk_\theta^s - w_\theta^s)\theta e_\theta^s - \frac{1}{2}k_\theta^{s^2} \right] \quad (10)$$

s.t.:

$$e_\theta = \arg \max_e \{w_\theta \theta e - \frac{1}{2}c e^2 k_\theta\}, \quad \text{for } \theta = \{1, \beta\} \quad (11)$$

$$w_\theta \theta e_\theta - \frac{1}{2}c e_\theta^2 k_\theta \geq 0, \quad \text{for } \theta = \{1, \beta\} \quad (12)$$

$$V_\theta^\theta \geq V_\theta^{\hat{\theta}}, \quad \text{for } \theta, \hat{\theta} = \{1, \beta\} \text{ and } \hat{\theta} \neq \theta. \quad (13)$$

Formally, this problem is identical to the benchmark problem (3)-(5) where  $\theta$  is observable with the addition of the truthtelling constraints (13). To solve it, we ignore the IC constraint of the high productivity firm (i.e.,  $V_\beta^\beta \geq V_\beta^1$ ) and assume that the IC of the low productivity firm binds (i.e.,  $V_1^1 = V_1^\beta$ ). We then substitute in the objective function (10), derive the corresponding first order conditions and get:<sup>13</sup>

**Proposition 2** *The optimal managerial compensation and firm investment are:*

$$w^{s*} = \{w_1^*, \Delta w_\beta^*\} \quad \text{and} \quad k^{s*} = \{k_1^*, \Delta k_\beta^*\} \quad (14)$$

where  $\Delta \equiv \max \left\{ \frac{1 + \sqrt{1 - 1/\beta}}{\beta}, 1 \right\}$ . With these choices, the manager's effort is  $e_\theta^{s*} = e_\theta^* = \frac{\theta r}{2c}$  and the principal's payoff is

$$V^{s*} = \frac{1}{2} \left[ (1 - \pi)k_1^{*2} + \pi\Delta(2 - \Delta)k_\beta^{*2} \right]. \quad (15)$$

---

<sup>13</sup>As we show in the appendix ignoring (as we do) compensation schemes that offer positive rewards when the output is low ( $q = 0$ ) is without loss of generality

The intuitions behind Proposition 2 can be illustrated by comparing its results with those results obtained in the benchmark case described in Proposition 1. In comparison with the observable  $\theta$  case, there is no change in compensation, investment or effort when  $\theta = 1$ , i.e.,  $w_1^* = w_1^{s*}$ ;  $k_1^* = k_1^{s*}$  and  $e_1^* = e_1^{s*}$  but when  $\theta = \beta$ , both compensation and investment are increased by the factor  $\Delta$  ( $k_\beta^{s*} = \Delta k_\beta^*$  and  $w_\beta^{s*} = \Delta w_\beta^*$ ). Such an increase in investment and compensation offsets any effect on effort which equals the effort of the benchmark case ( $e_\beta^* = e_\beta^{s*}$ ). That is, firms increase their scale (and managerial compensation) in order to convey that their prospects are promising. The scale increase  $\Delta$  depends on the difference of productivity among firms' types. In the relevant range (i.e.,  $1 < \beta < \beta^* \cong 1.84$ ),  $\Delta$  is a concave function of  $\beta$  which reach its maximum at  $\hat{\beta} = 1.15$ . In other words, when  $\beta \geq \beta^*$  there is no distortion in compensation or investment which implies that the solution of the problem when  $\theta$  is observable is also implementable when the agent cannot observe  $\theta$ .

It is worth noting that, in this setting, investment distortions provide a relatively efficient signal relative to other potential signaling possibilities such as relying solely on a distorted compensation policy. This is so because a type  $\theta = 1$  firm finds it relatively more costly to overinvest than a type  $\theta = \beta$  firm since, given its lower expected productivity, the marginal benefits of a high investment level are lower for the  $\theta = 1$  than for the  $\theta = \beta$  firm.<sup>14</sup> In contrast, signaling solely through the compensation does not work because, for incentive reasons, compensation needs to be contingent on success and the probability of success is higher when  $\theta = \beta$ . In these conditions, an increase in  $w_\theta^s$  (i.e., if  $w_1^s$  and  $w_\beta^s$  increase by a constant), is relatively more costly for a type  $\beta$  firm than for a type 1 firm which implies that signaling by distorting compensation is ineffective.<sup>15</sup> Alternatively, the firm could burn money (i.e., engage in wealth destruction regardless of the firm scale or output success) in order to signal its type. Such a device, however, is also suboptimal relative to signaling by distorting the firm scale. This is so because money burning is equally costly to both types of firms and hence less efficient than signaling by overinvesting which is relatively more costly for the low productivity firm.

---

<sup>14</sup>Since agent's effort costs increase with the firm's scale, the overinvestment distortion is accompanied by a corresponding distortion in compensation in order to restore the proper effort incentives to the agent.

<sup>15</sup>In the appendix we provide a more extensive analytical argument to show that other distortions in compensation (e.g., increasing compensation in case of failure) are suboptimal as well.

### 3.4 Pooling

Alternatively, the principal can commit to make choices that are independent of  $\theta$  and then commit not to disclose his information on  $\theta$ . In this case, we refer to  $\bar{w}$  and  $\bar{k}$  as the compensation in case of success and the investment and to  $\bar{\theta} \equiv \pi\beta + (1 - \pi)1$  as the average productivity shock. Therefore, the principal's problem is:

$$\max_{\bar{k}, \bar{w}, \bar{e}} V_p = (r\bar{k} - \bar{w})\bar{\theta}\bar{e} - \frac{1}{2}\bar{k}^2 \quad (16)$$

$$\text{s.t.}:$$

$$\bar{e} = \arg \max_e \{\bar{w}\bar{\theta}e - \frac{1}{2}ce^2\bar{k}\} \quad (17)$$

Since the firm's investment does not convey information as (17) shows, the effort decision is made based on the average type of the firm. The solution to this problem is given by:

**Proposition 3** *The optimal compensation and firm investment under pooling are:*

$$\bar{w}^* = \frac{r^3\bar{\theta}^2}{8c} \quad \text{and} \quad \bar{k}^* = \frac{r^2\bar{\theta}^2}{4c}. \quad (18)$$

*With these choices, the agent's effort is  $\bar{e}^* = \frac{r\bar{\theta}}{2c}$  and the principal's payoff is  $V_p^* = \frac{1}{2}\bar{k}^{*2}$ .*

Under pooling, the agent's effort is the average of the two efforts levels under the benchmark case when  $\theta$  is observable and the compensation is a sharing rule as in separation on the observable  $\theta$  case i.e.,  $\frac{\bar{w}^*}{r\bar{k}^*} = \frac{1}{2}$ . In contrast to separation, however, the firm does not incorporate information on the realized  $\theta$  for their investment choices which means that the firm overinvests when  $\theta = 1$  and underinvests when  $\theta = \beta$ .

Economically, the case of pooling is interesting since it shows that rigidities in investment and compensation can be part of an optimal capital budgeting policy. In a strict sense, our results show that setting a fixed investment level can be the optimal investment policy. However, these results can be more liberally interpreted as specifying a limit on the amount of capital that can be used by the firm. After observing the shock, an unconstrained high productivity firm would like to invest above the limit imposed by the investment policy, while the low productivity firm would just invest the investment limit only to mimic the high productivity principal's investment choice. This interpretation of pooling as an investment limit is consistent with a number of studies that show that firms may exhibit a tendency toward capital rationing when they allocate capital in different business units (e.g., Ross

1986 documents that from twelve firms examined in detail, six of them employ fixed capital budgets).

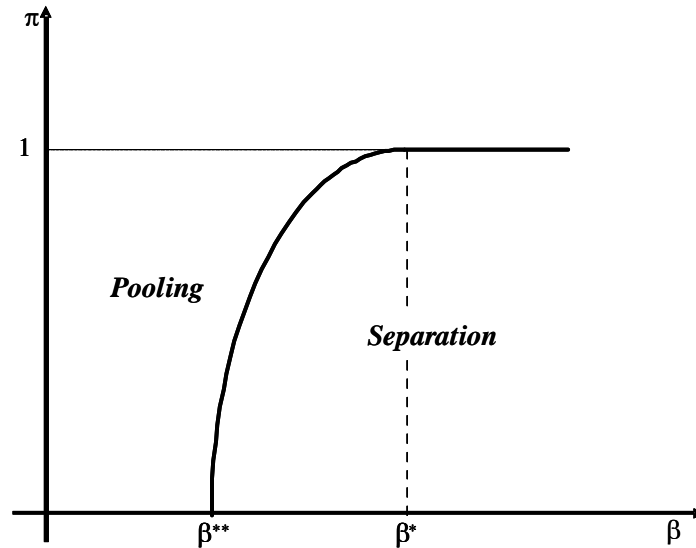
### 3.5 Separation or Pooling?

Compared with the separation case, the benefit to the firm with the pooling contract is that it does not have to overinvest nor overcompensate managers to signal its type. The cost, however, is that the firm ignores its information and cannot tailor investment to marginal productivity. The optimal choice between pooling and separation depends on the trade-off between these costs and benefits as the next proposition shows.

**Proposition 4** (*Comparative statics*) *Separation is more likely to be the optimal policy:*

- (i) *the lower the likelihood of a high productivity project (i.e.,  $\pi$ ) and*
- (ii) *the larger the difference in productivity among projects (i.e.,  $\beta$ ).*

The following figure illustrates the regions in the space  $(\beta, \pi)$  where either policy is optimal:



It is worth noting that, as the previous graph illustrates, pooling is locally optimal, that is, when productivity levels are sufficiently close (i.e.,  $\beta \rightarrow 1$ ) pooling is the optimal investment policy. To the extent that firms can exhibit a wide range of productivity levels, this observation suggests that investment rigidities (i.e., discrete levels of investment) are likely

to be part of any optimal capital budgeting policy. This observation will have important implications when we examine the continuum type version of the model discussed below.

To understand part i), it is helpful to consider cases where  $\pi$  is close to 1. When  $\pi$  is large, the expected over investment cost for the type  $\beta$  is large under separation. On the other hand, under pooling, because the equilibrium investment is close to the type  $\beta$  full-information investment level, the inefficiency is small if the firm is of type  $\beta$ . Even though for type 1 the investment inefficiency may be large, the expected cost is small because it is unlikely for the firm to become type 1. Therefore the trade-off favors pooling when  $\pi$  is large. To understand part ii), it is helpful to consider cases where  $\beta$  is large. On the one hand, the efficiency loss due to pooling is large because the average investment is far from both type 1 and type  $\beta$ 's full information level. On the other hand, the over investment cost may not be large when  $\beta$  is large. For example, when  $\beta$  is close to  $\beta^*$ , type  $\beta$  need only to over invest a little bit in order to deter type 1 mimicking because the full information investment level for type  $\beta$  is already very high for a type 1 firm. Therefore the trade-off favors separation when  $\beta$  is large.

## 4 A three-party model: Corporate capital budgeting

Having established the incentive to overinvest and the potential benefits of committing to specifying a maximum (fixed) investment level we extend the model to consider the interplay of three parties in the capital budgeting process. In addition to the entrepreneur (to whom we now refer to as a “CEO”) who obtains information and determines investment expenditures, and the employee who exerts effort, we introduce a third party (i.e., a “board”) that sets specific capital budgeting policies and offers compensation contracts to the CEO. One interpretation is that the third party is a venture capitalist that provides funding for the firm and retains some control of its operations. Another is that the informed party is the manager of the division of a conglomerate and the third party includes executives at the firm’s headquarters or are members of the board. In either interpretation, the third party is an entity that has the authority to set the firm policy.

The following figure considers the timing of the three-party model:

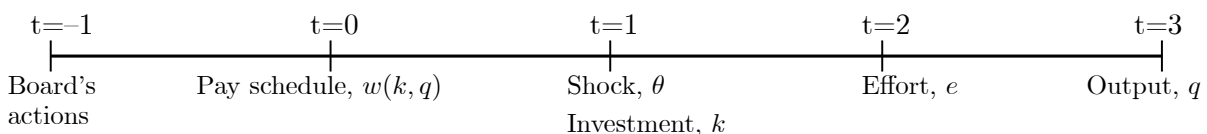


Figure 3: <sup>13</sup>Timing of Events

As shown in Figure 3, the three-party model adds period  $t = -1$  where the board sets the policy that determines the CEO's payoffs (i.e., the principal's objective function).

#### 4.1 Optimal discount rate policy

We consider first the case where at  $t = -1$ , the board chooses the discount rate,  $i$ , or equivalently, a discount factor,  $\delta \equiv \frac{1}{1+i}$  that the CEO uses to evaluate the firm's investment choices. We focus this analysis on the case where separation is the optimal firm policy, and consider the case of pooling briefly below. As a matter of practical implementation, we assume that the board can offer the CEO a compensation contract with "EVA-like" features that induces the CEO to evaluate investment with the discount factor  $\delta$ . Specifically, the board will offer the CEO a share,  $\alpha$ , of the firm's net value at  $t = 3$ , where net firm value is calculated as revenues minus wages and investment expenditures divided by the factor  $\delta$ .<sup>16</sup>

Consider the problem encountered by the CEO at  $t = 0$  who faces a discount factor  $\delta$  imposed by the board to evaluate the firm's investments. We denote by  $k^\delta = \{k_1^\delta, k_\beta^\delta\}$ ,  $e^\delta = \{e_1^\delta, e_\beta^\delta\}$  and  $w^\delta = \{w_1^\delta, w_\beta^\delta\}$  the investment, bonus and effort exerted after  $\theta = \{1, \beta\}$  respectively. In addition, we refer to  $V_{\theta, \delta}^{\hat{\theta}}$  as the CEO's payoff when, after observing  $\theta$ , offers  $w_\theta^\delta$  and invests  $k_\theta^\delta$  for  $\hat{\theta}, \theta = \{1, \beta\}$ :

$$V_{\theta, \delta}^{\hat{\theta}} \equiv (rk_\theta^\delta - w_\theta^\delta)\theta e_\theta^\delta - \frac{1}{2\delta}k_\theta^{\delta^2}. \quad (19)$$

Formally, the CEO solves the following maximization problem:

$$\max_{k^\delta, w^\delta, e^\delta} V_\delta = \sum_{\theta=\{1, \beta\}} p_\theta [(rk_\theta^\delta - w_\theta^\delta)\theta e_\theta^\delta - \frac{1}{2\delta}k_\theta^{\delta^2}] \quad (20)$$

s.t.:

$$e_\theta^\delta = \arg \max_e \{w_\theta^\delta \theta e - k_\theta^\delta c e^2\} \quad \text{for } \theta = \{1, \beta\} \quad (21)$$

$$V_{\theta, \delta}^\theta \geq V_{\theta, \delta}^{\hat{\theta}} \quad \text{for } \theta, \hat{\theta} = \{1, \beta\} \text{ and } \hat{\theta} \neq \theta. \quad (22)$$

---

<sup>16</sup>If  $\delta > 1$  the firm would be subsidizing the use of capital and if  $\delta < 1$  the firm would be taxing the use of capital.

Relative to the two-party case, the agent's problem remains unchanged and the CEO faces the same problem as the principal in the basic model except that the investment costs are given by  $\frac{1}{2\delta}k_\theta^{\delta^2}$  rather than  $\frac{1}{2}k_\theta^{\delta^2}$ . (Notice that since  $\alpha > 0$  is a constant it cancels out in the previous problem, i.e., it does not have any effect on the CEO's choices.) Following the analysis in the last section the solution can be described as follows:

**Lemma 1** *The solution of the CEO problem with separation scales investment and wage by  $\delta$  and  $\delta^2$ , respectively (i.e.,  $k_\theta^{\delta^*} = \delta k_\theta^{s*}$ ,  $w_\theta^{\delta^*} = \delta^2 w_\theta^{s*}$ ). The agent's effort is unaffected:  $e_\theta^{\delta^*} = e_\theta^{s*}$ .*

At  $t = -1$  the board takes into account the distortion that the discount factor  $\delta$  produces on the CEO's behavior at  $t = 0$  and, consequently, chooses  $\delta$  to maximize firm value:

$$\max_{\delta} V^\delta = \sum_{\theta=\{1,\beta\}} p_\theta [(rk_\theta^{\delta^*} - w_\theta^{\delta^*})\theta e_\theta^{\delta^*} - \frac{1}{2}k_\theta^{\delta^{*2}}]. \quad (23)$$

The solution to the board's problem is described in the following proposition.

**Proposition 5** *If separation with overinvestment characterizes the CEO's investment policy ( $\beta < \beta^*$ ) then the board distorts the discount rate upwards ( $\delta^* < 1$ ). If separation requires no overinvestment ( $\beta \geq \beta^*$ ) then the board does not distort the discount rate ( $\delta^* = 1$ ).*

To better explain the intuition behind the previous proposition it is useful to notice that in problem (23) the board maximizes firm value (i.e., the value determined by discounting cashflows with  $\delta = 1$ , the "correct" discount factor) by imposing  $\delta < 1$  on the CEO when  $\beta < \beta^*$ . By maximizing a distorted measure of firm value, one in which cashflows are discounted by  $\delta < 1$ , the CEO reduces the firm's investment expenditures. Intuitively, the benefits of distorting  $\delta$  can be explained as follows. A lower  $\delta$  reduces the incentives of a CEO who observes  $\theta = 1$  to mimic the behavior of a CEO who observes  $\theta = \beta$ . A relaxed IC constraint for  $\theta = 1$  reduces  $\beta$ 's cost to overinvest but creates some underinvestment in  $\theta = 1$ . In particular, an infinitesimal reduction from  $\delta = 1$  simultaneously diminishes  $w_\theta^\delta$ , and  $k_\theta^\delta$ . Since compensation and investment for the low type ( $w_1^\delta$  and  $k_1^\delta$ ) are at their optimal levels, such a reduction has a second order effect on firm value, while the reduction for the high type (a lower  $w_\beta^\delta$  and  $k_\beta^\delta$ ), which are set at local optima, has a first order effect on firm value. (When separation entails no distortion,  $\beta \geq \beta^*$ , the optimal discount factor



is  $\delta = 1$ , which keeps the investment and compensation policies undistorted.) We provide the formal proof of these arguments in the appendix.

It should be noted that the ability to set a discount rate does not increase firm value in the case of pooling. In this case, if the board sets  $\delta \neq 1$ , the CEO distorts his choices away from the optimal pooling choices (i.e.,  $\bar{k}^*$  and  $\bar{w}^*$ ). This reduces firm value since, by construction, the optimal pooling choice maximizes firm value relative to the set of choices that convey no information to the agent.<sup>17</sup> As a result, separation is more likely to be optimal when the hurdle rate is set by the board since a modified hurdle rate may improve the separation policy but does not help the pooling policy.

It should be emphasized that the board creates value by altering the CEO's objective function, which in turn affects the set of CEO's actions that are incentive compatible for a given  $\theta$ . Without the board, the CEO can also take actions that affect his ex-post investment incentives. For instance, the CEO can change the IC conditions by committing to transfer output to the agent (or to burn money) when his investment expenditures are high. These commitments by the CEO are costly and as the analysis of the two party model shows undesirable for the CEO. In contrast, when the board is the residual claimant, a CEO can effectively encourage or prevent certain investment levels without burning nor transferring output to any external party.

As we just discussed, imposing a higher discount rate (i.e.,  $\delta < 1$ ) can increase firm value by reducing the overinvestment costs of separation when  $\theta = \beta$ . However, within this setting the firm fails to achieve  $V^*$  (i.e., the firm value when  $\theta$  is observable) since it leads to underinvestment when  $\theta = 1$ . We conclude this section by considering whether a policy of multiple hurdle rates (i.e., rates set as a function of the amount of capital invested) will solve the problem. Proposition 6 confirms that this is indeed the case:

**Proposition 6** *For each  $\beta$  there exists  $\delta_\beta^*$  such that the firm follows the optimal investment policy:  $k_1^{1*} = k_1^{s*}$  and  $k_\beta^{\delta^*} = k_\beta^{s*}$  and its value reaches  $V^*$ .*

Proposition 6 implies that a policy that imposes higher hurdle rates for larger investments eliminates overinvestment when  $\theta = \beta$  without inducing underinvestment when  $\theta = 1$ .

---

<sup>17</sup>In separation, a distorted  $\delta$  may create value by enlarging the set of incentive compatible CEO policies. Since the pooling choices are made by the CEO before observing the type, in pooling distorting  $\delta$  does not have the potential to increase firm value.

The proposition is thus consistent with the evidence presented in Ross (1986); firms appear to have hurdle rates which are increasing in the size of the investment project. Nevertheless other practical considerations outside this model could make a multiple rate policy hard to implement. For instance, in the presence of multiple projects, multiple divisions or projects that require staged investments having a rate which depends on the amount invested could lead the CEO to take actions that “game” the nonlinearity inherent in the multiple hurdle rate policy.

## 4.2 Optimal debt policy

In this section we relax the assumption that the board can directly impose a hurdle rate on the CEO and consider instead how the board can use the firm’s financial policy to influence the CEO’s behavior. Specifically, we consider a CEO who faces a debt level  $d$  chosen by the board at  $t = -1$ . We denote by  $k^d = \{k_1^d, k_\beta^d\}$ ,  $e^d = \{e_1^d, e_\beta^d\}$  and  $w^d = \{w_1^d, w_\beta^d\}$  the investment, bonus and effort exerted after  $\theta = \{1, \beta\}$  respectively and, analogously to (19), define  $V_{\theta,d}^{\hat{\theta}} \equiv (rk_{\hat{\theta}}^d - w_{\hat{\theta}}^d - d)\theta e_{\hat{\theta}}^d - \frac{1}{2}k_{\hat{\theta}}^{d^2}$  as the CEO’s payoff when, after observing  $\theta$ , offers  $w_{\hat{\theta}}^d$  and invests  $k_{\hat{\theta}}^d$  for  $\hat{\theta}, \theta = \{1, \beta\}$ . Formally, the CEO’s problem is:

$$\max_{k^d, w^d, e^d} V^d = \sum_{\theta=\{1,\beta\}} p_\theta [(rk_{\hat{\theta}}^d - w_{\hat{\theta}}^d - d)\theta e_{\hat{\theta}}^d - \frac{1}{2}k_{\hat{\theta}}^{d^2}] \quad (24)$$

$$\text{s.t.:} \quad e_{\hat{\theta}}^d = \arg \max_e \{w_{\hat{\theta}}^d \theta e - k_{\hat{\theta}}^d f(e)\} \quad \text{for } \theta = \{1, \beta\} \quad (25)$$

$$V_{\theta,d}^\theta \geq V_{\theta,d}^{\hat{\theta}} \quad \text{for } \theta, \hat{\theta} = \{1, \beta\} \text{ and } \hat{\theta} \neq \theta. \quad (26)$$

Notice that, in contrast to problem (20) the cashflows are discounted at the appropriate discount rate but the CEO is maximizing equity value rather than total firm value. Following a similar approach as in previous cases (see the appendix for details), and focusing as before on the case of separation, the solution to the CEO problem for the levered firm can be described as follows:<sup>18</sup>

**Lemma 2** *The solution of the CEO problem with separation in the presence of a levered capital structure reduces the wage, effort, and investment relative to the unlevered case:*

$$w_{\hat{\theta}}^{d*} < w_{\hat{\theta}}^{s*}, \quad e_{\hat{\theta}}^{d*} < e_{\hat{\theta}}^{s*}, \quad \text{and} \quad k_{\hat{\theta}}^{d*} < k_{\hat{\theta}}^{s*}.$$

<sup>18</sup>It is easy to show that positive leverage reduces firm value in the case of pooling, where the optimal leverage is nil, i.e.,  $d^* = 0$ .

At  $t = -1$  the board chooses debt to solve the following problem:

$$\max_d V^d = \sum_{\theta=\{1,\beta\}} p_\theta [(rk_\theta^{d*} - w_\theta^{d*})\theta e_\theta^{d*} - \frac{1}{2}k_\theta^{d*2}]. \quad (27)$$

The solution to the board’s problem is described in the following proposition.

**Proposition 7** *If separation with overinvestment characterizes the CEO’s investment policy ( $\beta < \beta^*$ ) then the board chooses some debt to fund the project ( $d^* > 0$ ). If separation implies no overinvestment ( $\beta \geq \beta^*$ ) then the board funds the project entirely with equity ( $d^* = 0$ ).*

As with the distortion of hurdle rates, debt induces underinvestment to  $\theta = 1$  and reduces overinvestment to  $\theta = \beta$ .<sup>19</sup> Intuitively, leverage increases the mimicking costs of low quality firms to look like high quality ones. When  $d > 0$  the marginal cost of increasing  $k$  is borne by the equityholders but benefits (in part) the bondholders and, since the chance of success is lower for  $\theta = 1$ , mimicking  $k_\beta^{d*}$  becomes less attractive for  $\theta = 1$ . In short, in this setting, debt overhang is a mechanism to create disincentives to “mimic”.

While this section focuses on the case where the board sets the financing policy, it is useful to note that a similar logic would apply to the two-party model of Section 3. In such a setting, a principal-owner may find it useful to fund operations with a positive amount of debt in order to ameliorate his own overinvestment incentives even if the principal has sufficient funds to make the investments. As the logic of Proposition 7 suggests, the use of debt reduces the incentives of low productivity firms to make investments that partially benefit the debtholders of the firm.

## 5 Concluding remarks

In this paper we have proposed a “top-down” theory of capital budgeting where top manager’s actions motivate lower managers and the firm’s stakeholders to exert efforts on the firm’s behalf. One of the main purposes of our approach is to help reconcile several disconnects between academic theory and industry practice about capital budgeting practices. For example, while the NPV rule proposes using expected cash-flow estimates and expected

---

<sup>19</sup>In (27), the board chooses  $d$  to maximize firm value by affecting the incentives of the equity-value maximizer CEO.

return on investments with equivalent risk to discount those expected cash-flows, in practice there is a tendency to “inflate” cash-flow estimates and “inflated” hurdle rates.

A central premise of our paper is that the communication process between managers, headquarters and other firm’s constituencies has a direct influence on the capital budgeting process because the information conveyed by investment choices can have an important influence on the stakeholders’ choices. Specifically, the stakeholders’ incentives to exercise effort are affected by how costly stakeholders feel that capital is for the firm.

To explore the interaction between firm investments and stakeholder incentives, we have examined two related models. First, we considered an entrepreneurial firm whose production process requires a capital expenditure by the firm and effort by the employee. In this setting, in which the principal obtains information that can be more efficiently signaled with the firm’s investment policy, we show that there is a natural tendency for firms toward overinvestment and a natural demand for capital budgeting modifications of the NPV rule. Our main trade-off is between the efficiency gain associated with having an investment policy which incorporates the entrepreneur’s information and the efficiency loss associated with over-investment.

We then modified the analysis by introducing a third party with the authority to set specific capital budgeting policies and to offer compensation contracts to the entrepreneur. Within this setting we investigate the value of using a number of commonly observed capital budgeting practices. We find that high hurdle rates, the use of debt financing and the use of EVA managerial compensation can help to offset the overinvestment tendencies that emanate from firms with private information about their prospects.

Our theory provides an alternative explanation for the observed overinvestment in firms. While managerial private benefits are likely to provide impetus for excessive investment, in this paper we show that overinvestment can be a second best response of a manager who needs to communicate information to their subordinates or other stakeholders. While our model provides a unique rationale for overinvestment in firms in which agency problems are not a consideration (i.e., firms in which ownership and control are not separated), more generally, the empirical relevance of our theory vis-a-vis a theory based on managerial private benefits is an open question that we leave for future research.

## 6 Appendix: Proofs and other technical derivations

[To be completed]

### Proof of Proposition 1

We denote as  $w_0 = \{w_{1,0}, w_{\beta,0}\}$  the wage when  $z = 0$  and  $\theta = 1, \beta$  respectively. (In the main text, we denote as  $w = \{w_1, w_\beta\}$  when  $z = r$  and  $\theta = 1, \beta$ .) Therefore the principal's problem can be expressed as:

$$\max_{k, w, w_0, e} V = \sum_{\theta=\{1, \beta\}} p_\theta \left[ (rk_\theta - w_\theta)\theta e_\theta - w_{\theta,0}(1 - \theta e_\theta) - \frac{1}{2}k_\theta^2 \right] \quad (28)$$

$$\text{s.t.} \\ e_\theta = \arg \max_e \{w_\theta \theta e + w_{\theta,0}(1 - \theta e) - \frac{1}{2}k_\theta c e^2\}, \quad \text{for } \theta = \{1, \beta\} \quad (29)$$

$$w_\theta \theta e_\theta + w_{\theta,0}(1 - \theta e_\theta) - \frac{1}{2}k_\theta c e_\theta^2 \geq 0, \quad \text{for } \theta = \{1, \beta\}. \quad (30)$$

We prove that  $w_{\theta,0} = 0$  by contradiction. If  $w_{\theta,0} > 0$  and  $w_\theta = 0$  then setting  $w_{\theta,0} = 0$  would increase firm value and induce a higher level of agent's effort. If  $w_{\theta,0} > 0$  and  $w_\theta > 0$  then, by reducing  $w_{\theta,0}$  while keeping  $(w_\theta - w_{\theta,0})$  constant, firm value would increase without affecting the incentives to exert effort. The previous reasoning implies that  $w_{\theta,0} = 0$ .

Imposing  $w_{\theta,0} = 0$ , we can obtain the optimal  $k_\theta$ ,  $w_\theta$  and  $e_\theta$  by substituting the first order condition of the effort choice (29) and then solving in the first order conditions obtained by deriving with respect to  $w_\theta$  and  $k_\theta$ .

### Proof of Proposition 2

We distinguish two cases:

Case 1:  $\beta \geq \beta^*$ . In this case the unconstrained solution described in (7) satisfies  $V_\beta^\beta \geq V_\beta^1$  (i.e.,  $IC_\beta$ ) and  $V_1^1 \geq V_1^\beta$  (i.e.,  $IC_1$ ). (Notice that  $IC_1$  can be expressed as  $\beta^3(2 - \beta) \geq 1$  or simply as  $\beta \geq \beta^*$ .)

Case 2:  $\beta < \beta^*$ . In this case the unconstrained solution does not satisfy  $IC_1$  which forces us to solve the problem in the general case in which  $w_{\theta,0}^s \geq 0$ . (We denote as  $w_0^s = \{w_{1,0}^s, w_{\beta,0}^s\}$  the wage when  $z = 0$  when  $\theta = 1, \beta$  respectively.) Ignoring  $IC_\beta$  (we will check it later) we get the following program:

$$\max_{w_0^s, w^s, k^s, e^s} (1 - \pi)V_1^1 + \pi V_\beta^\beta \quad (31)$$

s.t.:

$$e_\theta = \arg \max_e \{w_\theta \theta e + w_{\theta,0}^s (1 - \theta e) - k_\theta c e^2\}, \quad \text{for } \theta = \{1, \beta\} \quad (32)$$

$$w_\theta \theta e_\theta - k_\theta c e_\theta^2 \geq 0, \quad \text{for } \theta = \{1, \beta\} \quad (33)$$

$$V_1^1 \geq V_1^\beta, \quad (34)$$

where

$$V_\theta^{\theta'} \equiv r k_\theta^s \theta e_{\theta'}^s - [w_{\theta',0} + (w_{\theta'}^s - w_{\theta',0}) \theta e_{\theta'}^s] - \frac{1}{2} k_{\theta'}^{s^2}.$$

In the previous problem,  $\theta = 1$ 's payoff is maximized as in the full information solution because any deviation from such solution (i.e.,  $w_{1,0}^s \neq 0$ ,  $w_1^s \neq \frac{r}{2} k_1^s$ , or  $k_1^s \neq k_1^*$ ) reduces  $V_1^1$  without helping with  $IC_1$ . This implies that in the optimal solution:  $w_{1,0}^s = 0$ ,  $w_1^s = \frac{r}{2} k_1^*$ , and  $k_1^s = k_1^*$ . We impose such values and solve for  $\theta = \beta$ 's optimal values:

$$\max_{w_{\beta,0}^s, w_\beta^s, k_\beta^s, e_\beta^s} (1 - \pi) V_1^1 + \pi V_\beta^\beta$$

$$\text{s.t.} \quad e_\beta^s \in \arg \max_e \{w_{\beta,0} + (w_\beta^s - w_{\beta,0}) \beta e - k_\beta^s c e^2\}, \quad (35)$$

$$V_1^1 \geq V_1^\beta. \quad (36)$$

Expression (36) can be written as:

$$V_1^1 = \frac{1}{2} k_1^{*2} \geq \{r k_\beta^s - (w_\beta^s - w_{\beta,0})\} e_\beta^s - \frac{1}{2} k_\beta^{s^2} - w_{\beta,0} = V_1^\beta. \quad (37)$$

We define  $\alpha_\beta \equiv \frac{w_\beta^s - w_{\beta,0}}{r k_\beta^s}$  and express the first order condition of (35) as  $e_\beta^s = \frac{\alpha_\beta r \beta}{c}$ . We then substitute it into the objective function and get

$$\max_{w_\beta^s, k_\beta^s, \alpha_\beta} (1 - \pi) \frac{k_1^{*2}}{2} + \pi \{r k_\beta^s (1 - \alpha_\beta) \beta \frac{\alpha_\beta r \beta}{c} - \frac{1}{2} k_\beta^{s^2} - w_{\beta,0}\},$$

whose Lagrangian is:

$$L = (1 - \pi) \frac{k_1^{*2}}{2} + \pi \{r k_\beta^s (1 - \alpha_\beta) \beta \frac{\alpha_\beta r \beta}{c} - \frac{1}{2} k_\beta^{s^2} - w_{\beta,0}\} - \lambda \{r k_\beta^s (1 - \alpha_\beta) \frac{\alpha_\beta r \beta}{c} - \frac{1}{2} k_\beta^{s^2} - w_{\beta,0} - \frac{1}{2} k_1^{*2}\} + \mu w_{\beta,0} + v k_\beta^s.$$

By Kuhn-Tucker Theorem, the FOCs with respect to  $w_{\beta,0}$ ,  $\alpha_\beta$ , and  $k_\beta^s$  are:

$$\lambda - \pi + \mu = 0$$

$$(\pi \beta - \lambda)(1 - 2\alpha_\beta) = 0 \quad (38)$$

$$(\pi \beta - \lambda) r (1 - \alpha_\beta) \frac{\alpha_\beta r \beta}{c} - (\pi - \lambda) k_\beta^s + v = 0 \quad (39)$$

Notice that since by (38),  $\pi\beta - \lambda = \mu + (\beta - 1)\pi > 0$ , we must have  $(\pi - \lambda)k_\beta^s > 0$ . That is  $\pi - \lambda = \mu > 0$  which implies that  $w_{\beta,0} \geq 0$  is binding. Equation (39) also implies that  $k_\beta^* = r(1 - \alpha_\beta)\frac{\alpha_\beta r\beta}{c} < k_\beta^s$ . That is, type  $\beta$  must overinvest. (38) implies that  $\alpha_\beta = \frac{1}{2}$ . Solving for  $k_\beta^s$  in the binding (37) after substituting  $\alpha_\beta = \frac{1}{2}$  and  $w_{\beta,0} = 0$ , we have

$$rk_\beta^s \frac{r\beta}{4c} - \frac{1}{2}k_\beta^{s2} - \frac{k_1^{*2}}{2} = 0 \quad (40)$$

which has two real roots. But since  $V_\beta^\beta = \beta rk_\beta^s \frac{r\beta}{4c} - \frac{1}{2}k_\beta^{s2} = rk_\beta^s \frac{r\beta}{4c}(\beta - 1) + rk_\beta^s \frac{r\beta}{4c} - \frac{1}{2}k_\beta^{s2}$ , we get:

$$V_\beta^\beta = rk_\beta^s \frac{r\beta}{4c}(\beta - 1) + \frac{k_1^{*2}}{2} \quad (41)$$

where (41) follows because of (40).  $V_\beta^\beta$  is increasing in  $k_\beta^s$ . As a result, the bigger root is optimal. So we have  $k_\beta^{s*} = \frac{k_\beta^* + \sqrt{k_\beta^{*2} - k_1^{*2}\beta^2}}{\beta}$ . Substituting  $k_\theta^* = \frac{r^2\theta^2}{4c}$ , we get  $k_\beta^{s*} = \Delta k_\beta^*$ .  $\Delta = 1$  has two real roots, 1 and  $\beta^*$ . Because  $\beta^*$  is the largest real root of  $\Delta = 1$  and the fact that for  $\beta \rightarrow \infty$   $\Delta < 1$ , it follows that  $\Delta > 1$  for  $1 < \beta < \beta^*$ .

We now need to check if  $IC_\beta$  is satisfied. This is so provided that

$$V_\beta^\beta \geq \beta rk_1^* \frac{r}{4c} - \frac{k_1^{*2}}{2} = V_1^1 + rk_1^* \frac{r}{4c}(\beta - 1)$$

Substituting (41) and  $V_1^1 = \frac{k_1^{*2}}{2}$ , it is equivalent to  $k_\beta^s \beta \geq k_1^*$ , which is satisfied since  $\beta > 1$  and  $k_\beta^s > k_\beta^* > k_1^*$ .

The principal's payoff is obtained by substituting (41):

$$V_s^* = \frac{k_1^{*2}}{2} + \pi k_\beta^s \frac{r^2\beta}{4c}(\beta - 1). \quad (42)$$

**Proof of Proposition 4:** The principal's payoff under separation can be rewritten as  $V_s^* = \frac{k_1^{*2}}{2} + \pi(1 - \frac{1}{\beta})\Delta k_\beta^{*2}$ , after substituting  $k_\beta^* = \frac{r^2\beta^2}{4c}$  into (42). Therefore, pooling dominates separation if and only if

$$\frac{\bar{k}^{*2}}{2} - \frac{k_1^{*2}}{2} \geq \pi(1 - \frac{1}{\beta})\Delta k_\beta^{*2}. \quad (43)$$

Substituting  $\bar{k}^* = \frac{r^2\bar{\theta}^2}{4c}$ ,  $k_\theta^* = \frac{r^2\theta^2}{4c}$ , and  $\Delta = \frac{\beta + \sqrt{\beta - 1}}{\beta^2}$  into (43), we have  $\frac{[1 + \pi(\beta - 1)]^4}{2} - \frac{1}{2} \geq \pi(1 - \frac{1}{\beta})\beta^2(\beta + \sqrt{\beta^2 - 1})$ . Let  $\delta \equiv \frac{[1 + \pi(\beta - 1)]^4}{2} - \frac{1}{2} - \pi(\beta - 1)[\beta^2 + \beta\sqrt{\beta^2 - 1}]$ . Pooling dominates separation if and only if  $\delta \geq 0$ .

We first fix  $\beta$  and find  $\pi^*(\beta)$ . First notice that  $\frac{\partial \delta}{\partial \pi} \Big|_{\pi=0} = (\beta - 1)[2 - \beta(\beta + \sqrt{\beta^2 - 1})]$ . Since  $[2 - \beta(\beta + \sqrt{\beta^2 - 1})] \geq 0$  iff  $\beta \leq \frac{2}{\sqrt{3}}$ ,  $\frac{\partial \delta}{\partial \pi} \Big|_{\pi=0} \geq 0$  if and only if  $\beta \leq \frac{2}{\sqrt{3}}$ . In this case, because  $\delta(\pi)$  is convex in  $\pi$ ,  $\frac{\partial \delta}{\partial \pi} > 0$  for all  $\pi > 0$ . Together with  $\delta(\pi = 0) = 0$ , we can conclude  $\delta > 0$  for all  $\pi > \pi^*(\beta) = 0$  on  $(1, \frac{2}{\sqrt{3}}]$ .

If  $\beta > \frac{2}{\sqrt{3}}$ , then  $\frac{\partial \delta}{\partial \pi} \Big|_{\pi=0} < 0$ . That is,  $\delta < 0$  when  $\pi$  is close to 0. We also know that  $\delta(\pi = 1) \geq 0$  because  $V_p^*(\pi = 1) = \frac{k_p^2}{2}$  achieves the full information payoff. We thus can conclude that there exists a  $\pi^*(\beta) > 0$  such that  $\delta = 0$  by continuity. By the convexity of  $\delta(\pi)$ , we can conclude that  $\delta(\pi) \leq 0$  for any  $0 \leq \pi \leq \pi^*(\beta)$ . Furthermore, because  $0 = \delta(\pi^*) - \delta(0) = \int_0^{\pi^*} \frac{\partial \delta}{\partial \pi}(\pi) d\pi < \int_0^{\pi^*} \frac{\partial \delta}{\partial \pi}(\pi^*) d\pi = \frac{\partial \delta}{\partial \pi}(\pi^*) \pi^*$ , we have

$$\frac{\partial \delta}{\partial \pi} \Big|_{\pi=\pi^*} > 0. \quad (44)$$

That is,  $\frac{\partial \delta}{\partial \pi} > 0$  for all  $\pi > \pi^*$  by convexity. Thus  $\delta > 0$  for all  $\pi > \pi^*(\beta)$ . Now we look at  $\delta(\beta, \pi^*(\beta)) = 0$ . Some tedious algebra shows that,

$$\frac{\partial \delta}{\partial \beta} \Big|_{\pi=\pi^*} < 0. \quad (45)$$

Applying the Implicit Function Theorem to  $\delta(\beta, \pi^*(\beta))$ , we have

$$\frac{d\pi^*}{d\beta} = - \frac{\frac{\partial \delta}{\partial \beta} \Big|_{\pi=\pi^*}}{\frac{\partial \delta}{\partial \pi} \Big|_{\pi=\pi^*}} > 0, \quad (46)$$

the inequality follows because of (44) and (45).  $\pi^*(\beta)$  is thus strictly increasing on  $(\frac{2}{\sqrt{3}}, \beta^*)$ .

We denote  $\beta^{**}(\pi) : (0, 1) \rightarrow (\frac{2}{\sqrt{3}}, \beta^*)$  the inverse function of  $\pi^*(\beta)$ . We conclude the proof by showing that for fixed  $\pi$ ,  $\delta(\beta) < 0$  if and only if  $\beta^{**}(\pi) < \beta < \beta^*$ . Suppose  $\beta_1 > \beta^{**}$ . We have  $\delta(\beta_1, \pi^*(\beta_1)) = 0$ . Since  $\pi^*(\cdot)$  is strictly increasing,  $\pi^*(\beta_1) > \pi$ . We thus conclude that for  $(\pi, \beta_1)$ , separation is optimal. The case where  $\beta_1 < \beta^{**}$  is proved symmetrically.

**Proof of Proposition 5:** First, we show that  $\delta = 1$  is not optimal. Take derivative of  $V_\delta$  with respect to  $\delta$ , we have

$$\frac{dV_\delta}{d\delta} = \frac{\partial V_\delta}{\partial \alpha_\delta(1)} \frac{d\alpha_\delta(1)}{d\delta} + \frac{\partial V_\delta}{\partial \alpha_\delta(\beta)} \frac{d\alpha_\delta(\beta)}{d\delta} + \frac{\partial V_\delta}{\partial k_\delta(1)} \frac{dk_\delta(1)}{d\delta} + \frac{\partial V_\delta}{\partial k_\delta^\delta} \frac{dk_\delta^\delta(\beta)}{d\delta}.$$



Since  $\alpha_\delta(\theta)$  satisfies the FOC  $\left. \frac{\partial V_s}{\partial \alpha_\theta} \right|_{\alpha(\theta)=\frac{1}{2}} = \left. \frac{\partial V_\delta}{\partial \alpha_\theta^\delta} \right|_{\delta=1} = 0$ . Similarly,  $\frac{\partial V_\delta}{\partial k_1^\delta} = 0$  at  $\delta = 1$ , since  $k_1^{s*}$  satisfies FOC  $\left. \frac{\partial V_s}{\partial k_1^s} \right|_{k_1^s=k_1^{s*}} = \left. \frac{\partial V_\delta}{\partial k_1^\delta} \right|_{\delta=1} = 0$ . Finally,  $\left. \frac{\partial V_s}{\partial k_\beta^s} \right|_{k_\beta^s=k_\beta^{s*}} = \left. \frac{\partial V_\delta}{\partial k_\beta^\delta} \right|_{\delta=1} < 0$  because  $k_\beta^{s*} > k_\beta^*$ . Therefore, since  $\frac{dk_\delta(\beta)}{d\delta} > 0$ ,

$$\left. \frac{dV_\delta}{d\delta} \right|_{\delta=1} = \left. \frac{\partial V_\delta}{\partial k_\beta^\delta} \right|_{\delta=1} \frac{dk_\delta(\beta)}{d\delta} \Big|_{\delta=1} < 0.$$

This implies that if we reduce  $\delta$ , we can strictly increase the value of  $V_\delta$ .

Second, we show that  $\delta > 1$  cannot be optimal. We first check that the optimal contract under  $\delta > 1$ ,  $(w_\theta^\delta, k_\theta^\delta)$  satisfies the type 1's IC under  $\delta = 1$ . The type 1's IC when  $\delta > 1$  is

$$\delta^4 k_\beta^{s*} r \left(1 - \frac{1}{2}\right) e_\beta^{s*} - \delta^4 \frac{k_\beta^{s*2}}{2} \leq \delta^4 k_1^{s*} r \left(1 - \frac{1}{2}\right) e_1^{s*} - \delta^4 \frac{k_1^{s*2}}{2}. \quad (47)$$

Substitute the contract  $(w_\theta^\delta, k_\theta^\delta)$  into the situation where  $\delta = 1$ , the type 1's IC becomes

$$\delta^3 k_\beta^{s*} r \left(1 - \frac{\delta}{2}\right) e_\beta^{s*} - \delta^4 \frac{k_\beta^{s*2}}{2} \leq \delta^3 k_1^{s*} r \left(1 - \frac{\delta}{2}\right) e_1^{s*} - \delta^4 \frac{k_1^{s*2}}{2}, \quad (48)$$

Comparing the two inequality, we find that deducting  $\delta^3 k_\beta^{s*} r (\delta - 1) e_\beta^{s*}$  from the LHS and deducting  $\delta^3 k_1^{s*} r (\delta - 1) e_1^{s*}$  on the RHS of (47) yields (48). But

$$\delta^3 k_\beta^{s*} r (\delta - 1) e_\beta^{s*} > \delta^3 k_1^{s*} r (\delta - 1) e_1^{s*}, \quad (49)$$

if and only if  $\delta > 1$ . Therefore, (48) is satisfied with strict equality. We can also check that  $(w_\theta^\delta, k_\theta^\delta)$  satisfies all the constraints in program (35) to (36). As a result,  $(w_\theta^\delta, k_\theta^\delta)$  is strictly suboptimal under  $\delta = 1$ , since at the optimum type 1's IC has to be binding. The firm value when under  $(w_\theta^\delta, k_\theta^\delta)$  is thus less than  $V_{\delta=1} = V_s^*$ .

Technically, reducing  $\delta$  from 1 relaxes the type 1's IC constraint, which is binding at  $\delta = 1$ . We can see this point by noticing that the inequality in (49) would flip for  $\delta < 1$ . Therefore, (48) would imply (47). Reducing  $\delta$  thus expands the opportunity set and as a result may increase the value of the objective function.

### Technical derivation 1: The optimal mechanism under non-commitment

Before proceeding to the optimal contract design, we need point out that the Revelation Principal in the mechanism design literature [Laffont and Green (1977), Myerson (1979), and Dasgupta, Hammand, and Maskin (1979)] fails to hold here as shown by proposition 4. This is because the agent's effort choice is a function of the principal's announcement. That

is, the agent's effort choice has to maximize the agent's payoff conditional on the principal's announcement. This incentive compatibility constraint breaks the equivalence between the direct mechanism, in which the principal tells the truth, and the indirect mechanism in which the principal does not.

It turns out that the optimal mechanism can be achieved by either pooling or separation. To show that, we first show that it is without loss of generality to consider only two contracts in our mechanism design.

Denote a contract  $c_i \equiv (w_i(q), k_i)$ . Since revelation principal does not hold any more, choosing a contract from the menu  $\{c_i\}_{i \in I}$  may not fully reveal his type.<sup>20</sup> Let  $\varphi(\theta)$  be the set of contracts that  $\theta$  chooses with positive probability (and for a slight abuse of notation, it also includes the probability of each type choosing  $c_i$ ,  $\pi_\theta(c_i)$  for  $c_i \in \varphi(\theta)$ ). Formally, the principal's problem is

$$\begin{aligned} & \max_{c_i, i \in I} \sum_{\theta \in \{1, \beta\}} \Pr(\theta) \sum_{c_i \in \varphi(\theta)} \pi_\theta(c_i) V_\theta(c_i) \\ \text{s.t.} \quad & e_i \in \arg \max_e \{ {}^0 w_i + ({}^{k_i r} w_i - {}^0 w_i) E[\theta | c_i] - h(e, k_i) \}, \end{aligned} \quad (50)$$

$${}^0 w_i + ({}^{k_i r} w_i - {}^0 w_i) E[\theta | c_i] - h(e, k_i) \geq 0, \quad (51)$$

$$V_\theta(c_i) \geq V_\theta(c_j), \text{ for any } c_i \in \varphi(\theta) \quad (52)$$

$$V_\theta(c_i) \geq 0 \quad (53)$$

$$\sum_i \pi_\theta(c_i) = 1, \quad (54)$$

$${}^q w_i \geq 0 \quad (55)$$

Let  $a_i \equiv (rk_i - w_i)e_i$  where  $w_i \equiv {}^{rk_i} w_i - {}^0 w_i$ .  $\alpha_i = \frac{w_i}{k_i r}$ . By the standard revealed preference argument, we have the following result.

**Lemma 3** (*Monotonicity*) *In any mechanism, we have*

*i) if type  $\beta$  picks contract  $c_i$  with positive probability, and type 1 picks contract  $c_j$  with positive probability, then  $a_i \geq a_j$ ;*

*ii) if type  $\beta$  is indifferent between  $c_i$  and  $c_j$  with  $a_i > a_j$ , then type 1 strictly prefers  $c_j$  to  $c_i$ ; If type 1 is indifferent between  $c_i$  and  $c_j$  with  $a_i > a_j$ , then type  $\beta$  strictly prefers  $c_i$  to  $c_j$ .*

---

<sup>20</sup> $I$  need not be finite or even countable..

Proof: type 1 and  $\beta$ 's ICs imply

$$a_i\beta - \frac{1}{2}k_i^2 - w_{i,0} \geq a_j\beta - \frac{1}{2}k_j^2 - w_{j,0} \quad (56)$$

and

$$a_j - \frac{1}{2}k_j^2 - w_{j,0} \geq a_i - \frac{1}{2}k_i^2 - w_{i,0} \quad (57)$$

Adding up both inequalities yields

$$(a_i - a_j)(\beta - 1) \geq 0, \quad (58)$$

which implies i). For part ii) we first consider type  $\beta$ ' IC (56). Indifference implies  $\beta(a_i - a_j) = \frac{1}{2}k_i^2 - \frac{1}{2}k_j^2 + w_i - w_j$ . That is  $(a_i - a_j) < \frac{1}{2}k_i^2 - \frac{1}{2}k_j^2 + w_{i,0} - w_{j,0}$ , which implies that type 1 strictly prefers  $c_j$ . The case where type 1 is indifferent follows by similar argument. *q.e.d.*

**Lemma 4** *For any mechanism, there is a mechanism with at most two contracts that is no worse than the original one for either types.*

Proof: There are two cases. First, in the original mechanism there are two contracts  $c_\theta \in \{1, \beta\}$  such that  $c_\theta$  is chosen by type  $\theta$  only. In this cases, a separation mechanism with only two contracts,  $c_\theta \in \{1, \beta\}$  and type  $\theta$  only chooses  $c_\theta$  with probability 1 will be no worse than the original mechanism. The reason is that for each type,  $V_\theta(c_\theta)$  is the best value it can achieve in the original mechanism by its IC and in the new mechanism it achieves it with probability 1. Notice that  $c_\theta$  satisfies all the constraints from (50) to (55).

Second, one type always pools with the other. We first consider the case where this pooling type is  $\beta$ . More specifically, type 1 chooses any contract in  $\varphi(\beta)$  with positive probability. And the probability that a type 1 chooses a contract from  $\varphi(\beta)$  is  $\pi_1(\varphi(\beta)) \leq 1$ . By lemma 3, if there are two contracts  $c_i$  and  $c_j$  in  $\varphi(\beta)$ , then  $a_i = a_j = a$ . We now construct the new mechanism. We first construct a contract  $c'_p$  for both types to pool. Pick a contract,  $c_p$ , in  $\varphi(\beta)$  such that  $E[\theta|c_k] \leq E[\theta|\varphi(\beta)]$ , where  $E[\theta|\varphi(\beta)] \equiv E[E[\theta|c_i]|c_i \in \varphi(\beta)]$ , is the conditional expectation of  $\theta$  on a contract  $c_i \in \varphi(\beta)$  is chosen. We then modify  $c_p$  to  $c'_p$ . We let type  $\beta$  choose  $c'_p$  with probability 1 and type 1 choose it with probability  $\pi_1(\varphi(\beta))$ . Thus,  $E[\theta|c'_p] = E[\theta|\varphi(\beta)]$ . We will modify  $k'_p$  so that  $a'_p = k'_p r(1 - \alpha_p) \frac{\alpha_p E[\theta|c'_p]}{c} =$

$k_p r(1 - \alpha_p) \frac{\alpha_p E[\theta|c_1]}{c} = a$ . Since  $E[\theta|c'_p] \geq E[\theta|c_p]$ ,  $k'_p \leq k_p$ . We then increase  $w_{p,0}$  to  $w'_{p,0}$  so that  $\frac{1}{2}k_p^2 + w_{p,0} = \frac{1}{2}k_p'^2 + w'_{p,0}$ . Thus  $c'_p = \{k'_p, \alpha_p, w'_{p,0}\}$ . We can check that both type's payoff from choosing  $c'_p$  are the same as that of choosing a contract in  $\varphi(\beta)$  in the original mechanism. Now we look at type 1. There are two cases. If  $\pi_1(\varphi(\beta)) = 1$ , then we are done because the original mechanism is the same as this one contract,  $c'_p$  with both type pooling on it. If  $\pi_1(\varphi(\beta)) < 1$ , then exists a contract  $c_1$  so that only type 1 chooses with positive probability in the original mechanism. We now keep  $c_1$  in the new mechanism and let type 1 chooses it with probability  $1 - \pi_1(\varphi(\beta))$ . The new mechanism has tow contracts,  $c'_p$  and  $c_1$ . We can see that type 1 is indifferent between the two and type  $\beta$  strictly prefers  $c'_p$ . Each type gets the same payoff as in the original mechanism.

The case where the pooling type is 1 is proved similarly. *q.e.d.*

**Proposition 8** *The principal's optimal expected payoff is achieved if he chooses either pooling or separation.*

Proof: We only need to consider  $\beta < \beta^*$ . For  $\beta \geq \beta^*$ , separation achieves the full information payoff. By lemma 4, we only need to consider two contracts as in the separation case. However, each type of principal may select the contracts randomly, i.e., revealing information partially. Now the mechanism is  $(w_{\hat{\theta},q}, k_{\hat{\theta}}, \sigma_{\theta})$ , As before,  $\hat{\theta}$  is the principal's claim and  $\theta$  is her true type. The first two terms are as before, the last term,  $\sigma_{\theta}$  is the probability that a type  $\theta$  principal chooses  $(w_{\theta,q}, k_{\theta})$ . Formally the contract design problem is

$$\max_{w_{\theta}, k_{\theta}, \sigma_{\theta}} \sum_{\theta=1,\beta} p_{\theta} V_{\theta}(\theta)$$

$$s.t. \quad e_{\hat{\theta}} \in \arg \max_e \{w_{\theta,0} + [w_{\hat{\theta},rk_{\hat{\theta}}} - w_{\hat{\theta}}^0] E[\theta|\hat{\theta}] - k_{\hat{\theta}} f(e)\}, \quad (59)$$

$$w_{\hat{\theta}}^0 + [w_{\hat{\theta},rk_{\hat{\theta}}} - w_{\hat{\theta}}^0] E[\theta|\hat{\theta}] - k_{\hat{\theta}} f(e) \geq 0, \quad (60)$$

$$V_{\theta}^{\theta} \geq V_{\theta}^{\hat{\theta}}, \text{ for any } \hat{\theta} \neq \theta \quad (61)$$

$$(1 - \sigma_{\theta}) [V_{\theta}^{\theta} - V_{\theta}^{\hat{\theta}}] = 0, \text{ for } \theta = 1, \beta \quad (62)$$

Notice that this problem is similar to optimization problem in the separation case with two differences. First, the mechanism can be random; second, relatedly we have a new constraint, (62), which is the complementary slackness constraint. It says that if type  $\theta$  is

strictly better off choosing a contract, she will choose it with probability 1. On the other hand, if she is indifferent between the two contracts, she can randomize. (62) is a natural requirement for mixed strategies.

Notice that this program includes the full separation and full pooling equilibria studied before. This is because for the separation case,  $\sigma_\theta = 1$ . In the pooling case,  $({}^q w_\theta, k_\theta) = ({}^q w, k)$  for all  $\theta$ . In general, we need to consider three cases: I. Only type  $\beta$ 's IC of (61) is binding; II. only type 1's IC of (61) is binding; and III. both types ICs are binding.

Case I is impossible for the following reason. In this case, we can solve the problem with only type  $\beta$ 's IC (thus ignore type 1's IC and complementary slackness conditions). However, we know that the full information allocation in subsection (3.1) satisfies all the constraints. So the optimal solution should be the full information allocation (it should be clear that the program cannot get a better solution than the full information case). However, we know that is not possible since the full information allocation violates type 1's IC by assumption  $\beta < \beta^*$ .

Case III cannot be better than the pooling equilibrium studied in subsection 3.4. In case III, both types of principal are indifferent between the two contracts. Therefore the objective function becomes

$$\max_{w_{\theta,q}, k_\theta, \sigma_\theta} \sum_{\theta=1,\beta} p_\theta V_\theta^1$$

that is, the objective function only depends on contract for type 1. Without loss of generality, we assume  $E[\theta|1] \leq E[\theta]$ . That is, if the type 1 contract is chosen, the agent's posterior expectation of  $\theta$  is lowered (this assumption is without loss of generality because we can always define type 1 contract such that this assumption is satisfied). Now we claim that the solution can never be strictly better than the pooling equilibrium studied in subsection 3.4. Otherwise, let the contract  $(w_{1,q}, k_1)$ , be the contract in the pooling equilibrium. The agent would exert a higher effort in the pooling case than here when  $(w_{1,q}, k_1)$  is chosen since  $E[\theta|1] \leq E[\theta]$ . The contract satisfies the agent's IR in the pooling case too because (60) is satisfied and  $E[\theta|1] \leq E[\theta]$ . Therefore, under  $(w_{1,q}, k_1)$ , the principal has a better payoff than under  $(\bar{w}^*, \bar{k}^*)$  because the agent's effort is higher, which is a contradiction to the optimality of the contract  $(\bar{w}^*, \bar{k}^*)$ .

Therefore, we only need to consider case II as the potential optimal mechanism. In this case, the type 1 principal using a mixed strategy choosing the two contracts, and the type

$\beta$  always chooses  $(w_{\beta,q}, k_\beta)$ . Because when the type 1 chooses  $(w_{1,q}, k_1)$  she is fully revealed, the allocation in this case is the same as in the full information case as argued in the proof of proposition 2. Formally, the program is

$$\max_{w_\beta, k_\beta, \sigma_1} (1 - \pi)V_1^1 + \pi V_\beta^\beta \quad (63)$$

$$\text{s.t.} \quad e_\beta \in \arg \max_e \{w_{\beta,0} + [w_{\beta,k_\beta r} - w_\beta]E[\theta|\hat{\theta} = \beta]e - k_\beta f(e)\}, \quad (64)$$

$$V_1^1 \geq V_1^\beta, \quad (65)$$

$$w_{\beta,q} \geq 0, \quad (66)$$

$$k_\beta \geq 0, \quad (67)$$

$$0 \leq \sigma_1 \leq 1. \quad (68)$$

By Bayes rule,  $E[\theta|\hat{\theta} = \beta] = 1 + \frac{\pi}{\pi+(1-\pi)\sigma_1}(\beta - 1)$ . Let  $E[\theta|\hat{\theta} = \beta] = v\beta$ ,  $v \in [(\beta)^{-1}, 1]$ . Because  $v$  and  $\sigma_1$  are one to one, we change the control variable  $\sigma_1$  to  $v$ .

Similar to the proof of Proposition 2, we form the Lagrangian of the problem

$$\begin{aligned} L = & (1 - \pi)\frac{k_1^{*2}}{2} + \pi \left\{ rk_\beta(1 - \alpha_\beta)\beta\frac{\alpha_\beta r v \beta}{c} - g(k_\beta) - w_\beta^0 \right\} \\ & - \lambda \left\{ rk_\beta(1 - \alpha_\beta)\frac{\alpha_\beta r v \beta}{c} - g(k_\beta) - w_\beta^0 - \frac{k_1^{*2}}{2} \right\} + \mu w_\beta^0 + vk(\beta) \\ & - \xi(v - 1). \end{aligned}$$

By Kuhn-Tucker Theorem, we get the following FOCs with respect to  $w_\beta^0$ ,  $\alpha_\beta$ ,  $k_\beta$ , and  $v$ , respectively

$$\lambda - \pi + \mu = 0, \quad (69)$$

$$\alpha_\beta = \frac{1}{2}, \quad (70)$$

$$(\pi\beta - \lambda)r(1 - \alpha_\beta)\frac{\alpha_\beta r v}{c} - (\pi - \lambda)k_\beta + v = 0, \quad (71)$$

$$(\pi\beta - \lambda)rk(\beta)(1 - \alpha_\beta)\frac{\alpha_\beta r \beta}{c} - \xi = 0. \quad (72)$$

Similar argument as in the proof of Proposition 2, we have  $\pi\beta - \lambda > 0$ . The last equation thus implies that  $\xi > 0$ , which in turn implies that  $v = 1$ . That is, the optimal solution is actually the separation case as studied before. As a result, it is without loss of generality to focus on only pooling and separating equilibria.

## References

- [1] Ambrose, M. L. and C. T. Kulik (1999) “Old Friends, New Faces: Motivation Research in the 1990s” *Journal of Management* 25, pp. 231-92.
- [2] Beaudry, P. (1994) “Why An Informed Principal May Leave Rents To An Agent” *International Economic Review* 34, pp. 821-832.
- [3] Benabou R. and J. Tirole (2003) “Intrinsic and Extrinsic Motivation” *Review of Economic Studies*, 70, pp 489-520.
- [4] Berkovitch E. and R. Israel (2004) “Why the NPV Criterion Does Not Maximize NPV,” *Review of Financial Studies* 17, pp. 239-55.
- [5] Bernardo, A., H. Cai, and J. Luo, (2001) “Capital Budgeting and Compensation with Asymmetric Information and Moral Hazard,” *Journal of Financial Economics* 61, 311-344.
- [6] Bernardo, A., H. Cai, and J. Luo, (2004), “Capital Budgeting in Multi-Division Firms: Information, Agency, and Incentives,” *Review of Financial Studies* 17, 739-767.
- [7] Bernardo, A., J. Luo, and J. Wang, (2006), “A Theory of Socialistic Internal Capital Markets,” *Journal of Financial Economics* 80, 485-509.
- [8] Bester, H. and R. Strausz, (2001), “Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case,” *Econometrica*, vol. 69(4), pages 1077-98.
- [9] Burkart, Mike, D. Gromb, and F. Panunzi, (1997) “Large Shareholders, Monitoring, and the Value of the Firm,” *The Quarterly Journal of Economics*, MIT Press, vol. 112(3), pages 693-728, August. Burkart, Gromb and Panunzi (1997)
- [10] Dasgupta, P., P. Hammond, and E. Maskin, (1979), “The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility,” *Review of Economic Studies*, 46, pp. 185-216.
- [11] Green J. and J-J. Laffont, (1977) “On the revelation of preferences for public goods,” *Journal of Public Economics*, Elsevier, vol. 8(1), pages 79-93, August.

- [12] Graham, J. and C. Harvey (2001), “The theory and practice of corporate finance: evidence from the field,” *Journal of Financial Economics*, Elsevier, vol. 60(2-3), pages 187-243, May.
- [13] Grinstein, Y. and E. Tolkowsky (2004), “The role of the board of directors in the capital budgeting process – Evidence from S&P 500 Firms”. Working paper.
- [14] Gompers P. and J. Lerner (1999), “The Venture Capital Cycle” MIT Press.
- [15] Harris, M. and A. Raviv (1996), “The Capital Budgeting Process: Incentives and Information,” *The Journal of Finance*, 51, pp. 1139-1174.
- [16] Harris, M. and A. Raviv (1998), “Capital Budgeting and Delegation,” *Journal of Financial Economics*, 50, pp. 1139.
- [17] Hart, O. and J. Moore (1995), “Debt and Seniority: An Analysis in the Role of Hard Claims in Constraining Management,” *American Economic Review*, 85, pp. 567-85.
- [18] Hermalin, B. (1998), “Toward an Economic Theory of Leadership: Leading by Example,” *American Economic Review* 88, pp. 1188-1206.
- [19] Holmstrom B. (1982), “Moral Hazard in Teams,” *The Bell Journal of Economics*, Vol. 13, No. 2 pp. 324-340
- [20] Jensen, M., (1986) “The Agency Costs of Free Cash Flow: Corporate Finance and Takeovers,” *American Economic Review*, 76.
- [21] Komai M., Stegeman M. and B. Hermalin. (2007) “Leadership and Information.” *American Economic Review* 97 (3), 944-947.
- [22] Marino, A.M., and J.G., Matsusaka (2005) Decision processes, agency problems, and information: an economic analysis of budget procedures. *Review of Financial Studies*, 18 pp. 301-325.
- [23] Meier, I. and V. Tarhan (2007) “Corporate investment decision practices and the hurdle rate premium puzzle”, working paper.



- [24] Myers, S. (1977) Determinants of corporate borrowing, *Journal of Financial Economics* 5 pp. 147–175.
- [25] Myerson, R. (1979) “Incentive compatibility and the bargaining problem”, *Econometrica* 47, 61-73.
- [26] Poterba, J. and L. Summers (1995) “A CEO Survey of U.S. Companies’ Time Horizons and Hurdle Rates,” *Sloan Management Review*, MIT, Fall.
- [27] Strausz, R. (2009) “Entrepreneurial Financing, Advice, and Agency Costs,” *Journal of Economics and Management Strategy* 18-3 pp. 845-870.