A Theory of Socialistic Internal Capital Markets*

Antonio E. Bernardo†

Jiang Luo‡

James J.D. Wang§

January 21, 2005

Abstract

We develop a model of a two-division firm in which the "strong" division has, on average, higher quality investment projects than the "weak" division. We show that the firm optimally biases its project selection policy in favor of the weak division and this bias is stronger when there is a greater spread in average project quality. The cost of such a policy is that the firm sometimes funds an inferior project but the benefit is that it motivates the manager of the strong division to set (and meet) more aggressive cash flow targets.

*All errors are ours.

†UCLA Anderson School of Management, Los Angeles, CA 90095-1481, USA; e-mail: abernard@anderson.ucla.edu

‡Department of Finance, HKUST, Clear Water Bay, Hong Kong

§Department of Economics and Finance, City University of Hong Kong, Kowloon, Hong Kong

1 Introduction

A vast majority of corporate investment is financed with internal funds.[1] Since investment decisions made using internal funds are not subject to the same scrutiny from the external capital markets as those funded with new equity or debt issues it is important to examine how effectively internal capital markets allocate funds to their best use. Although this research question is difficult to tackle directly, the availability of data on investments by major business lines (segments) of U.S. public companies has allowed researchers to compare investment decisions in conglomerate firms to investment decisions in focused firms. This literature has typically shown that investment in one division of a conglomerate firm is sensitive to the cash flows of unrelated divisions and that conglomerate firms invest less in divisions from industries with good investment opportunities and more in divisions from industries with poor investment opportunities compared to their focused counterparts (see, e.g., Lamont, 1997; Scharfstein, 1998; Shin and Stulz, 1998; Rajan, Servaes, and Zingales, 2000; and Gertner, Powers, and Scharfstein, 2002).[2] Moreover, this socialistic behavior is more severe when there is more diversity in the quality of investment opportunities across divisions in the firm (Rajan, Servaes, and

[1]The use of internal funds as a percentage of total investment for non-financial corporations in the U.S. ranged from 72% to 108% annually between the years 1990 and 2000 (Brealey and Myers, 2003). The percentage exceeds two-thirds in Germany, Japan, and the U.K. (Corbett and Jenkinson, 1997).

[2]The typical empirical strategy is to use the median Tobin's q of (traded) stand-alone firms in an industry to measure the quality of investment opportunities in a (non-traded) division of a conglomerate firm. Chevalier (2000), Whited (2001), Bernardo and Chowdhry (2002), Campa and Kedia (2002), Gomes and Livdan (2004) argue against this empirical strategy because the decision to diversify (i.e., form a conglomerate) is endogenous and may be determined in part by the quality of the division's investment opportunities thus the median q of stand-alone firms may not be an accurate measure of the investment opportunities of a conglomerate division. A notable exception to the finding of socialistic cross-subsidization is Maksimovic and Phillips (2002) who use plant-level data from manufacturing firms and find that these firms re-allocate resources in favor of strong divisions when they experience a positive demand shock.

Zingales, 2000; Lamont and Polk, 2002; Billett and Mauer, 2003).[3]

These empirical observations are difficult to reconcile with agency models at the level of the CEO. For example, if the CEO has empire-building preferences such models would predict overinvestment in all divisions instead of a re-allocation of funds from strong to weak divisions. Two recent models go a level deeper into a firm's hierarchy and focus on the role of rent-seeking activity by division managers. Scharfstein and Stein (2000) develop a model in which managers divert their time away from productive effort to enhance their outside options and increase their bargaining power when negotiating total compensation. They argue that such behavior is more problematic for managers of weak divisions because the opportunity cost to them is relatively low. One way to mitigate this problem is to offer the manager a cash bribe to refrain from such behavior. However, they argue that if there is another layer of agency between the CEO and shareholders, the CEO may prefer to distort investment in favor of the weak division rather than increase cash payments to the manager because the latter comes from discretionary funds (which the CEO can control and potentially divert to himself) rather than investment funds (which are assumed to be under the control of shareholders). One problem with this explanation is that it seems implausible that CEOs would prefer to misallocate potentially hundreds of millions of investment dollars in order to maintain discretion over a relatively small cash payment to the division manager. Rajan, Servaes, and Zingales (2000) develop a model in which division managers have autonomy to choose between an efficient investment and a "defensive" investment which protects the surplus created from other managers. While the efficient investment maximizes firm value, a manager may prefer the "defensive" investment particularly when the surplus created is far greater than in other divisions. The firm can mitigate this inefficiency by tilting the capital budget in favor of the division with lower-quality investments.

A key feature of both these models is that the firm cannot write managerial incentive

[3]Khanna and Tice (2001) show that firms operating in multiple divisions of the same broad industry do not engage in inefficient cross-subsidization.

contracts that depend on the project cash flows. If even crude contracts could be written, the rent-seeking behavior in these models would likely be mitigated at a much lower cost than tilting the capital budget. In this paper, we develop a model that explains why firms invest too much in weak divisions and too little in strong divisions relative to first-best even when they can write managerial compensation contracts that depend on division cash flows. We consider a two-division firm in which the "strong" division has, on average, higher quality investment projects than the "weak" division. The firm selects one project to fund based on analyses (or "reports") provided by division managers with private information about the quality of their projects. These reports establish cash flow targets for each project. The presence of moral hazard and competition for capital gives the division managers opposing incentives to misrepresent their private information. On one hand, by setting low, easily attainable targets the division manager can shirk from privately costly effort and still produce the cash flows expected by the firm. On the other hand, by setting aggressive targets the division manager is more likely to get her project funded and receive the pecuniary benefits that come with it. To obtain truthful representations of project quality, we show that the firm optimally increases the likelihood of selecting the project and links the (winning) manager's pay more closely to performance when she reports a higher quality project. In this way, the firm discourages a manager from setting a low target by making funding less likely and discourages a manager from setting a high target by making her "buy" a larger share in the project.[4]

Our main result shows that the firm optimally biases the project choice in favor of the weak division; specifically, the strong division's project is funded only when its reported (and, in equilibrium, true) quality exceeds the weaker division's project quality by a positive "premium". The cost of such a policy is that the firm sometimes funds an

[4]Our paper also contributes to the literature examining the effectiveness of monitoring, compensation, and reputation to mitigate various agency and information problems in the capital budgeting process (e.g., Harris and Raviv, 1996, 1998; Holmstrom and Ricart i Costa, 1986; Bernardo, Cai, and Luo, 2001, 2004; Garcia, 2002; Berkovitch and Israel, 2004; Ozbas, 2004). The emphasis in this paper, however, is on the allocation of resources in relatively weak and strong divisions.

inferior project; however, the benefit is that it increases competition for internal capital and reduces the incentive for the manager of the strong division to shirk on effort by setting low, easily achievable targets. We also show that the firm increases the bias in favor of the weak division when the spread between the ex ante quality of investments in the two divisions is greater, consistent with the empirical evidence that socialistic cross-subsidization is more pronounced when there is greater spread in the industry Tobin's-q across divisions (Rajan, Servaes, and Zingales, 2000; Lamont and Polk, 2002). The firm also increases the bias in favor of the weak division when the firm is less effective at monitoring the managers' actions, consistent with the empirical evidence that cross-subsidization is more pronounced when firms have ineffective boards of directors (Palia, 1999) and when the CEO has poorly-aligned incentives (Scharfstein, 1998). We also show that when the two divisions have the same expected project quality but differ in the uncertainty about project quality, the firm biases project selection against the division with more uncertainty. Thus, we argue that idiosyncratic risk may be relevant for project evaluation not because it affects the appropriate discount rate - all our agents are risk-neutral - but instead because it affects equilibrium incentives for managerial effort. Finally, we show that division manager compensation depends on the attributes of the other divisions in the firm; for example, expected compensation and performance-pay is greater for division managers in low-growth businesses when the other divisions in the firm are in high-growth businesses.[5]

The remainder of the paper is organized as follows. Section 2 presents our model.

[5]Goel, Nanda, and Narayanan (2004) offers an explanation of socialistic internal capital markets due to CEO career concerns. In their model, the CEO wishes to maximize her perceived ability and thus finds it optimal to allocate more intangible (unobservable) resources to divisions with cash flows that are more informative about her ability. Anticipating this bias, the firm's owners also find it optimal to allocate more physical (observable) capital to these divisions. The model's cross-sectional predictions about socialistic cross-subsidization follow mainly from across-division differences in informativeness about CEO ability whereas our model's predictions follow mainly from across-division differences in the investment opportunity set.

Section 3 derives the first-best allocation and the optimal second-best mechanism in the case where the distribution of project qualities in the two divisions have the same mean but different variance. We derive implications for project selection bias, hurdle project qualities, and division manager compensation. Section 4 considers the case where the distribution of project qualities in the two divisions have the same variance but different mean. Section 5 concludes and gives direction for future research.

2 The model

We consider a model of a firm with two divisions, indexed by $i = 1, 2$. Each division has a single investment opportunity and is run by a manager with private information, denoted t_i, about its quality.[6] The two projects have the same initial cost and the firm is assumed to have capital sufficient to fund only one of them. If her project is funded, the successful division manager can increase her project's future expected cash flows by taking privately costly, unverifiable, and non-contractible actions, denoted e_i. These actions may include firing other top managers, shutting down an inefficient plant, etc. In keeping with the moral hazard literature, we refer to these actions as managerial "effort". We use a specific functional form for the cash flows of investment project i:

$$V_i \quad = \quad \delta t_i + \theta e_i + \epsilon_i,$$

where $\delta > 0$ measures the importance of unknown (to the firm) project quality, $\theta > 0$ measures the importance of the division manager's effort, and ϵ_i are independent noise terms with mean zero.[7] For simplicity, we assume no discounting. While each manager

[6] In our model, the scope of the firm is given exogenously. However, if top management, and in particular the CEO, have more precise prior information about the t_i than the external market then the firm's relative ability to "pick winners" may explain why these two divisions exist within one firm, see., e.g. Alchian (1969), Williamson (1975), and Stein (1997).

[7] The ϵ_i represent measurement error which makes it impossible for the board to infer an exact relation between t_i and e_i by observing the cash flows V_i. However, since everyone is risk neutral in our model, the mean zero noise term and its distribution have no effect on our results.

is assumed to know the precise quality, t_i, of her own project the firm only knows that t_i is drawn from a normal distribution with mean η_i and variance σ_i^2, i.e., $t_i \sim N(\eta_i, \sigma_i^2)$. The firm's prior belief is that division 2 has better investment opportunities, on average, than division 1, i.e., $\eta_2 > \eta_1$. We normalize $\eta_1 = 0$ and let $\eta_2 = \eta > 0$ for notational convenience. We will refer to division 2 as the strong division and division 1 as the weak division. For now we assume that the variance of project quality is equal in the two divisions and normalize $\sigma_1^2 = \sigma_2^2 = 1$. We consider the case of equal means but different variances in Section 4.

We assume both managers are risk neutral and their expected utility is given by:

$$U_i = E_{\epsilon_i} w_i - 0.5\gamma e_i^2,$$

where w_i is her compensation, $E_{\epsilon_i} w_i$ is the expected compensation (expectation over ϵ_i), and γ parameterizes the managers' effort cost. Each manager has an outside option yielding reservation utility $\bar{U} = 0$. We assume that a manager receives compensation w_i when her project gets funded and receives her reservation utility, $\bar{U} = 0$, when her project is not funded.[8]

In the capital budgeting process, the firm chooses which project to fund based on analyses or "reports", denoted $\hat{t}_i$, provided by the division managers. For example, these reports might include a detailed analysis of the project's discounted cash flows. As we shall see below, the presence of moral hazard and competition for capital gives the division managers opposing incentives to misrepresent their private information. On one hand, each manager will have an incentive to understate project quality so that they can shirk on effort and still produce the cash flows expected by the firm. On the other hand, each manager will have an incentive to overstate project quality to enhance the probability of selection.

The firm's problem is to choose a project selection and managerial compensation policy to mitigate these agency and information problems and maximize the expected

[8]Although this seems to restrict us to a narrow class of mechanisms, the derived mechanism can be shown to be globally optimal.

payoff to its risk-neutral shareholders. Specifically, the firm designs an optimal mechanism consisting of (i) a selection policy $\{p_i(\hat{t})\}$ which is the probability that project i gets funded as a function of both reports $\hat{t} \equiv \{\hat{t}_1, \hat{t}_2\}$, and (ii) a compensation schedule $\{w_i(\hat{t}, V_i)\}$ which is the compensation to manager i if her project is selected as a function of both reports and the project's cash flows. Importantly, we assume that the managers' private information is not directly observable or verifiable by the firm ex post, therefore, contracts cannot be written on t_i directly. Moreover, since the effort choice is unobservable and unverifiable by the firm, contracts cannot be written on e_i directly. We allow selection policy to be probabilistic in order to admit the largest set of feasible mechanisms; however, we shall show below that the equilibrium selection policy becomes deterministic once both managers announce their type, i.e., the firm chooses the candidate with the highest "score" (not necessarily the highest quality) with probability one and rejects the other. Therefore, the optimal selection policy is implemented deterministically.

The sequence of moves of the game is as follows:

date 1. The firm offers the mechanism $\{w_i(\hat{t}, V_i), p_i(\hat{t})\}$.

date 2. Each manager reports on the quality of her project, $\hat{t}_i$.

date 3. The firm selects project i according to the selection probabilities $\{p_i(\hat{t})\}$.

date 4. Successful manager i chooses the effort level e_i.

date 5. The cash flows, V_i, are realized and distributed to shareholders less the compensation $w_i(\hat{t}, V_i)$ paid to manager i.

We make the standard assumption in these types of models that the firm can commit to the mechanism offered at date 1. Absent a commitment device, it might be optimal for the board to choose a different project at date 3 from the one specified by the selection policy offered at date 1. However, if the managers knew this they would not report truthfully in the first place. Although we don't model this explicitly, the firm's commitment to the mechanism offered at date 1 may be optimal if it recognizes that it

will play this game repeatedly with the division managers in the future.[9]

3 Optimal mechanism

To provide a benchmark, we begin by demonstrating the following result for the socially efficient (first-best) solution assuming no moral hazard or information problems. This solution maximizes the expected total surplus (expectation over ϵ_i's).

Proposition 1. *The first-best project selection policy and managerial effort are given by:*

$$\begin{aligned} p_i^{fb}(t) &= \begin{cases} 1 & \text{if } t_i > \max(t_j, -0.5\frac{\theta^2}{\delta\gamma}), \\ 0 & \text{otherwise;} \end{cases} \\ e_i^{fb} &= \frac{\theta}{\gamma}. \end{aligned}$$

The proof is in the Appendix. Since each manager's effort has the same marginal productivity (θ) and cost (γ), the firm would request the same level of effort no matter which project is financed. Therefore, in the first-best solution, the ranking of each project is determined solely by the ranking of its quality, t_i. The firm should write a complete contract specifying the effort choice in Proposition 1 for the manager of the higher quality project and this manager should receive a fixed wage satisfying the division manager's participation constraint.

We now solve for the firm's optimal mechanism assuming the firm does not know either project quality t_i or effort e_i. By the Revelation Principle we can, without loss of generality, restrict our attention to direct revelation mechanisms in which the managers

[9]The assumption of commitment allows us to apply the Revelation Principle to find the optimal mechanism. This mechanism specifies investment rules and compensation that motivate the managers to reveal their information truthfully. Thus, our model says nothing about who should make investment decisions. Marino and Matsusaka (2005) provide an alternative model without commitment to study the optimal decentralization of investment decisions.

report their project qualities truthfully. Thus, the firm's mechanism design problem can be stated as:

$$\max_{w_i(t,V_i),p_i(t),e_i(t)} E\Pi \equiv \sum_{i=1,2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_i(t) E_{\epsilon_i} \left[V_i - w_i(t, V_i) \right] d\Phi(t_1) d\Phi(t_2 - \eta)$$

such that, $\forall i$

(IC1) $e_i(t) \in \arg\max_{\hat{e}_i} \; E_{\epsilon_i} w_i(t, V_i(t_i, \hat{e}_i, \epsilon_i)) - 0.5\gamma \hat{e}_i^2$,

(IC2) $t_i \in \arg\max_{\hat{t}_i} \; U_i(t_i, \hat{t}_i) \equiv \int_{-\infty}^{\infty} p_i(\hat{t}_i, t_j) \left[E_{\epsilon_i} w_i(\hat{t}_i, t_j, V_i(t_i, , \hat{e}_i, \epsilon_i)) - 0.5\gamma \hat{e}_i^2 \right] d\Phi(t_j - \eta_j)$,

(IR) $U_i(t_i, t_i) \geq 0$,

(NN) $e_i(t), p_i(t) \geq 0$, and $p_1(t) + p_2(t) \leq 1$,

where $\Phi(.)$ denotes the c.d.f. of the standard normal distribution, and $U_i(t_i, \hat{t}_i)$ is the expected utility of the manager who reports $\hat{t}_i$, has true type t_i, and assumes that the other manager is reporting her true type. The first incentive compatibility constraint (IC1) requires that the successful division manager chooses effort to optimize her expected utility. The second incentive compatibility constraint (IC2) requires that, for every project type, each manager finds it optimal to report truthfully given that the other manager also reports truthfully. The constraint (IR) is the standard interim individual rationality constraint requiring that each manager's expected equilibrium payoff is at least as large as her outside reservation utility, $\bar{U} = 0$. The non-negativity constraint (NN) requires that effort allocations are non-negative and the selection policy is well-defined.

Let $\phi(.)$ denote the p.d.f. of the standard normal distribution and $\mu(.) \equiv (1 - \Phi(.))/\phi(.)$ denote the inverse of its hazard rate.[10] Define the "score" function:

$$H(t_i, \eta_i) \equiv \delta t_i + 0.5\gamma \left[\max \left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta} \mu(t_i - \eta_i) \right) \right]^2.$$

[10]Below we shall use two important properties of the standard normal distribution: $\mu' < 0$ and $\mu'' > 0$.

The following proposition specifies the optimal mechanism and demonstrates that the firm chooses the project with the highest score, $H(t_i, \eta_i)$, which is not necessarily the project with the highest quality.

Proposition 2. *The optimal mechanism can be implemented in dominant strategies by the following project selection policy and linear compensation contract:*

$$\begin{aligned} p_i(t) &= \begin{cases} 1 & \text{if } H(t_i,\eta_i) > \max(0, H(t_j,\eta_j)) \ , \\ 0 & \text{otherwise;} \end{cases} \\ w_i(t, V_i) &= a_i(t) + b_i(t) V_i \end{aligned}$$

where

$$\begin{aligned} b_i(t) &= \max\left(0, 1 - \frac{\delta\gamma}{\theta^2}\mu(t_i - \eta_i)\right), \\ a_i(t) &= \int_{-\infty}^{t_i} \delta p_i(s, t_j) b_i(s, t_j) ds - b_i(t)(\delta t_i + 0.5\frac{\theta^2}{\gamma} b_i(t)). \end{aligned}$$

If project i is selected, manager i will provide effort

$$e_i(t) = \frac{\theta b_i(t_i)}{\gamma} = \max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_i - \eta_i)\right).$$

The proof is in the Appendix.

The optimal mechanism can be implemented by a project selection policy, $p_i(t)$, and a linear compensation contract consisting of salary, $a_i(t)$, and profit-sharing, $b_i(t)$. It can be shown that the selection policy and profit-sharing are monotone increasing and the salary is monotone decreasing in the manager's (truthful) report of her own project's quality, t_i. These properties of the optimal mechanism balance the agency and information problems faced by the firm by increasing the likelihood of selection when a project is reported to be of higher quality while at the same time linking the successful manager's pay more closely with performance. In this way, it is costly for the manager to understate quality because it reduces the chances her project will be selected and it is costly for the manager to overstate quality because it requires the successful manager to "buy" more shares (i.e., higher performance pay) with cash (i.e., lower salary).

The optimal mechanism is implementable in dominant strategies, thus it is robust to each manager's beliefs about the other manager's actions. Moreover, the optimal compensation contract is not affected by the noise term, ϵ_i. The reason is that the compensation contract trades off the benefit of providing effort incentives against the cost of eliciting truthful reporting. In our model, the additive noise term does not affect the firm's ability to infer t_i and e_i separately (e.g., even if the noise term $\epsilon_i \equiv 0$ the firm can only infer $\delta t_i + \theta e_i$ by observing V_i) or the cost of providing incentive because both managers are risk-neutral.[11]

3.1 Project selection bias

The optimal mechanism in Proposition 2 reveals that the firm selects the project with the highest "score", given by $H(t_i, \eta_i)$, not necessarily the project with the highest quality. The following result demonstrates that the optimal selection policy biases against the stronger division.

Corollary 1. (Existence of selection bias) *The firm chooses to finance project 2 only if its quality is higher than the quality of project 1, i.e., given* t_1, $p_2 = 1$ *only if* $t_2 > t_1 + \tau$ *where* $\tau \geq 0$ *and for high* t_1, $\tau > 0$.

The proof is in the Appendix. The intuition for the result is as follows. A division manager with a high-quality project has more scope to shirk on effort and still meet any fixed cash flow target. Thus, in order to get truthful information from a manager of a high-quality project, the firm provides effort incentives in the form of greater performance-based pay, b_i. We show in the proof that this has the effect of increasing the manager's expected compensation, net of private effort costs, i.e., managers with

[11] In contrast, noise terms do affect the cost of providing incentives, and thus the optimal contract, when managers are risk-averse. We include the noise term because it shows that our results hold more generally and it is natural to assume that cash flows are subject to measurement errors and other random shocks. A more technical proof of the robustness of the linear contract to additive noise is given in Laffont and Tirole (1993, pp. 72-73).

higher quality projects enjoy *information rents* in the form of higher expected compensation. However, to induce truthtelling, these information rents must be increasing in the manager's project quality, t_i, otherwise she would have an incentive to claim that the project is of lower quality. In other words, if the firm commits to giving large information rents to a manager with a project of a given quality it must also commit at least as much rents to all managers with better quality projects. This makes the marginal cost of providing effort incentives b_i very costly to the firm, particularly when there are many possible higher types. For a fixed project quality, there are more possible higher types in the strong division because its distribution of project types has a higher mean than in the weak division. Consequently, it is costlier for the firm to provide appropriate incentives to the manager of the strong division, *all else equal*, and the firm optimally biases project selection in favor of the weak division. In sum, each manager's report balances the benefit of under-reporting quality (shirking on effort) against the cost (decreasing the likelihood of selection). Since the manager of the strong division has greater scope for shirking on effort, the firm optimally biases the project selection against her to encourage more aggressive target-setting and thus more effort.

It is important to note that the above result states that the firm invests too much (little) in the weak (strong) division relative to first-best. We do not claim that multi-division firms invest too much (little) in weak (strong) divisions relative to single-segment firms; thus, we cannot adequately address the value consequences of integration and, in particular, whether socialistic internal capital markets can help to explain the "diversification discount" - the empirical observation that conglomerate firms are valued at discount to a matched portfolio of focused firms (Lang and Stulz, 1994; Berger and Ofek, 1995).[12] To address this issue we would need to specify more fully the costs and benefits

[12]The theory that inefficient internal capital markets explains the diversification discount has some proponents (e.g., Rajan, Servaes, and Zingales, 2000; Scharfstein and Stein, 2000) and opponents (e.g., Chevalier (2000), Whited (2001)). A recent paper by Villalonga (2004) argues that the diversification discount is an artifact of the COMPUSTAT segment data. She defines business units using the Business Information Tracking Series (BITS) census database and instead finds a diversification premium.

of forming conglomerates, describe the nature of capital constraints in single-segment versus multi-segment firms, and model explicitly why it is that the firm has chosen to bring together the two divisions. While this issue is extremely important, we simply take the scope of the firm as given and examine the consequences for the allocation of internal funds.

The following result shows when the project selection bias is weak or strong.

Corollary 2. (Comparative statics for selection bias) *The quality premium project 2 must offer, τ, increases in η (the mean difference between the two divisions) and θ (importance of effort); and decreases in δ (importance of unknown project quality) and γ (effort aversion).*

The proof is in the Appendix. The intuition for these results can be understood as follows. In Proposition 1 and Corollary 1, we established that for a given project quality the information rents enjoyed by the manager of the strong division is greater than for her counterpart in the weak division. In response, the firm tilts the capital budget in favor of the weak division to improve the strong division manager's effort incentives. When η is higher, the spread between the average quality in the two divisions is greater, and the existing differences in the managers' effort incentives become more exaggerated and this results in the firm optimally tilting the capital budget even more in favor of the weak division. Fixing η, a higher value of θ (or a lower δ or a lower γ) increases the marginal benefit of providing effort incentives (b_i) in the optimal contract. The effect of this is to increase managerial information rents but moreso for the strong division and this again results in the firm optimally tilting the capital budget even more in favor of the weak division.

The most interesting empirical implication here is that the bias in favor of the weak division is stronger when the difference between the average quality of projects in the two divisions is greater. This prediction is also made by Scharfstein and Stein (2000) and Rajan, Servaes, and Zingales (2000) and is consistent with the empirical evidence that socialist cross-subsidization is more pronounced when there is greater spread in the

industry Tobin's-q across divisions (Rajan, Servaes, and Zingales, 2000; Lamont and Polk, 2002).

Further empirical implications follow from reasonable interpretations of the other key parameters in our model: δ, θ and γ. The parameter δ represents the relevance of the division manager's private information for the project cash flows thus we expect δ to be greater, for example, when the firm is early in its lifecycle or in a high-growth industry. The parameter θ represents the importance of unverifiable managerial effort thus we expect θ to be greater, for example, when the division managers' job tasks require more firm-specific human capital. Finally, the parameter γ represents effort cost; however, an alternative and more empirically useful interpretation of $1/\gamma$ (more generally, the inverse of the second derivative of the cost function $1/C''(e_i)$) is the responsiveness of unverifiable actions to an increase in incentives (Milgrom and Roberts, 1992). In this interpretation, γ is lower when the division managers have more discretion which is more likely true if, for example, the firm's board of directors is ineffective. Similarly, while there is no conflict between the CEO and shareholders in our model, γ is lower when the CEO has poorly aligned incentives because he will not be motivated to provide monitoring effort and division managers will have more discretion. Thus, we predict that internal capital markets are more socialistic (i.e., biased more in favor of weak divisions) when (i) the firm is relatively mature; (ii) the division managers require more firm-specific human capital, (iii) the board is less effective at monitoring managers, and (iv) the CEO has poorly aligned incentives. The last two predictions are also made by Scharfstein and Stein (2000) and are consistent with the empirical evidence that cross-subsidization is more pronounced when firms have ineffective boards of directors (Palia, 1999) and when the CEO has low-powered incentives (Scharfstein, 1998).

3.2 Hurdle quality

In the first-best solution, the firm funds a project only when $t_i > \underline{t}^{fb} = -0.5\theta^2/(\delta\gamma)$. The hurdle quality level, $\underline{t}^{fb}$, takes into consideration the contribution of managerial effort

therefore it is negative. In the second-best solution, the firm funds a project only when $t_i > \underline{t}_i^{sb}$ where $H(\underline{t}_i^{sb}, \eta_i) = 0$. Comparing these investment policies yields the following result:

Corollary 3. (Hurdle quality) *The firm has a higher hurdle quality in the second-best than in the first-best, i.e.,* $\underline{t}_i^{sb} > \underline{t}^{fb}$*; and the firm has a higher hurdle quality for the strong division than for the weak division, i.e.,* $\underline{t}_2^{sb} \geq \underline{t}_1^{sb}$*, where the inequality holds strictly for some parameter values.*

The proof is in the Appendix. In the second-best mechanism, effort incentives are weaker than in the first-best (i.e., $b_i(t_i) \leq 1$) thus the equilibrium level of managerial effort is smaller. Since the total project cash flows depend additively on unobserved project quality and managerial effort, the hurdle project quality must be higher in the second-best to compensate for the lower effort. For the same reason, the firm optimally imposes a higher hurdle quality on the strong division than in the weak division. In other words, projects cannot simply be evaluated and ranked using the discounted cash flow (DCF) methodology; rather, the firm must introduce a "fudge-factor" when evaluating investment projects which depends on, among other things, the relative strength of its divisions.

3.3 Managerial compensation

The optimal mechanism can be implemented with a linear managerial compensation contract thus performance pay is measured by the slope of the contract, $b_i(t)$. Conditional on having her project selected, the manager of the weak division 1 receives expected performance pay, $\bar{b}_1$, and expected total compensation, $\bar{w}_1$, given by:

$$\begin{aligned} \bar{b}_1 &= E_{\{t_1,t_2\}}\Big[b_1(t)|H(t_1,0) > \max(0, H(t_2,\eta))\Big], \\ \bar{w}_1 &= E_{\{t_1,t_2\}}\Big[a_1(t) + b_1(t)E_{\epsilon_1}V_1|H(t_1,0) > \max(0, H(t_2,\eta))\Big]. \end{aligned}$$

We now show the following result.

Corollary 4. (Managerial compensation) *Conditional on having her project selected, the manager of the weak division receives higher-powered incentives and greater total compensation, on average, when the other division in the firm has better investment opportunities, i.e.,* $\bar{b}_1$ *and* $\bar{w}_1$ *increase in* η.

The proof is in the Appendix. The intuition for this follows from the following two features of the optimal mechanism. First, $b_i(t_i)$ is increasing in t_i because, as we argued earlier, the marginal cost of providing effort incentives is lower when there are fewer possible higher types. As t_i increases there are fewer higher types and the firm optimally increases b_i. Second, the stronger the other division the higher the hurdle quality, on average, the manager of the weak division must surpass in order to get funding. One empirical implication of this result is as follows. Consider a set of multi-division firms each of which have a division operating in a mature industry with few good growth opportunities. We predict that the performance-based pay and total compensation for these division managers will be higher in the subset of firms that also operate in other industries with many good growth opportunities.

3.4 N divisions

In our optimal mechanism, the firm selects the project with the highest "score" (given by $H(t_i, \eta_i)$) and compensates the winning manager with greater performance-pay, $b_i(t_i)$, when the project quality is higher. These results extend to the case of $N > 2$ divisions. To understand the effect of having N divisions we note two important features of our optimal mechanism. First, the investment distortion within a division vanishes as the true project quality approaches its upper bound. Second, for fixed quality t_i, performance-pay does not depend on the quality of the other project t_j. The reason for this is that the quality of the competing project only imposes a lower-bound on the threshold quality for selecting a project whereas the marginal cost of providing effort incentives depends only on the proportion of *higher* quality project types *within* a division. Thus, extending our model to N divisions does not change the form of the compensation contract. However,

increased competition does increase the expected quality of the winning project thus it impacts the average performance pay conditional on the project's selection and the average investment distortion. In sum, as we increase the number of divisions in the firm, the form of the optimal mechanism is unchanged but the expected agency and information costs go to zero.

4 Same mean quality but different variances

We now consider the case where the firm's prior belief is that the two projects have similar average quality, normalized to 0, but different variances, i.e., $t_1 \sim N(0, \sigma_1^2)$ and $t_2 \sim N(0, \sigma_2^2)$ where $\sigma_2 = \sigma > \sigma_1 = 1$. For what follows, we make the following assumption:

(A1) $\dfrac{\theta}{\gamma} - \dfrac{\delta\sigma}{\theta}\mu(0) < 0.$

Assumption (A1) requires that the quality of project 2 is significantly uncertain (high σ) so the two projects are sufficiently different. Our results still hold (for high project qualities) when this assumption is violated.

It is straightforward to show that the first-best solution is same as in Proposition 1; in the absence of asymmetric information and moral hazard, the firm would always select the higher-quality project. Following the derivation in Proposition 2, we can solve for the optimal second-best mechanism. First, define the "score" function:

$$R(t_i, \sigma_i) \equiv \delta t_i + 0.5\gamma \left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta\sigma_i}{\theta}\mu(\frac{t_i}{\sigma_i})\right)\right]^2.$$

Proposition 3. *The optimal mechanism can be implemented in dominant strategies by the following selection policy and linear compensation scheme:*

$$\begin{aligned} p_i(t) &= \begin{cases} 1 & \text{if } R(t_i, \sigma_i) > \max(0, R(t_j, \sigma_j)) \ , \\ 0 & \text{otherwise;} \end{cases} \\ w_i(t, V_i) &= a_i(t) + b_i(t)V_i \end{aligned}$$

where

$$\begin{aligned} b_i(t) &= \max\left(0, 1 - \frac{\delta\gamma\sigma_i}{\theta^2}\mu(\frac{t_i}{\sigma_i})\right), \\ a_i(t) &= \int_{-\infty}^{t_i} \delta p_i(s,t_j) b_i(s,t_j) ds - b_i(t)(\delta t_i + 0.5\frac{\theta^2}{\gamma} b_i(t)). \end{aligned}$$

If project i is selected, manager i will provide effort

$$e_i(t) = \frac{\theta b_i(t)}{\gamma} = \max\left(0, \frac{\theta}{\gamma} - \frac{\delta\sigma_i}{\theta}\mu(\frac{t_i}{\sigma_i})\right).$$

The proof is omitted since it parallels that of Proposition 2. We now demonstrate the existence of a selection bias in favor of the division with lower variance in project quality.

Corollary 5. (Existence of selection bias) *The firm chooses to finance project 2 only if its quality is higher than the quality of project 1, i.e, given t_1, $p_2 = 1$ only if $t_2 > t_1 + \rho$ where $\rho \geq 0$ and for high t_1, $\rho > 0$.*

The proof is in the Appendix. The intuition is similar to the case of different means. In this case, for a fixed (positive) project quality, there are more possible higher types in division 2 because its distribution of project types has a higher variance. Consequently, it is costlier for the firm to provide appropriate incentives to the manager of division 2 and the firm optimally biases project selection in favor of division 1. Moreover, the comparative statics results for the relative strength of the selection bias are identical to the case of different means.

Corollary 6. (Comparative statics for selection bias) *Given that project 1 has quality t_1, the quality premium project 2 has to offer, ρ, increases in σ (the variance difference between the two divisions) and θ (importance of effort); and decreases in δ (importance of unknown project quality) and γ (effort aversion).*

The proof is in the Appendix.

Recall, in the first-best solution, the firm funds a project only when $t_i > \underline{t}^{fb} = -0.5\theta^2/(\delta\gamma)$. The hurdle quality level, $\underline{t}^{fb}$, takes into consideration the contribution of

managerial effort; therefore, it is negative. In the second-best solution, the firm funds a project only when $t_i > \underline{t}_i^{sb}$ where $R(\underline{t}_i^{sb}, \sigma_i) = 0$. Comparing these investment policies yields the following result:

Corollary 7. (Hurdle quality) *The firm has a higher hurdle quality in the second-best than in the first-best, $\underline{t}_i^{sb} > \underline{t}^{fb}$; and the firm has higher hurdle quality for the division with more uncertainty, $\underline{t}_2^{sb} \geq \underline{t}_1^{sb}$, where the inequality holds strictly for some parameter values.*

The proof is in the Appendix. The firm requires a higher hurdle quality for projects from the division with more uncertainty about its investment opportunities. Since this uncertainty is idiosyncratic this result implies that firms bias against making investments with greater idiosyncratic risk. In standard applications of the discounted cash flow (DCF) methodology without agency or information problems, idiosyncratic risk does not affect project evaluation because such risks vanish in a well-diversified portfolios. In the presence of agency and information problems, however, it is costlier for the firm to provide effort incentives when it is more uncertain about the project's true quality. The firm compensates for this by biasing its investment policy against projects with greater idiosyncratic risk. This is one potential explanation for the survey evidence that firms use a much higher hurdle rate of return than justified by standard asset pricing models (Poterba and Summers, 1995).

5 Conclusions

In this paper, we argue that when division managers have private information about project quality and can enhance cash flows with privately costly effort the firm will optimally commit to tilt the capital budget in favor of divisions with relatively poor investment opportunities. The cost of such a commitment is that the firm sometimes funds an inferior project but the benefit is that it increases competition for internal funds and reduces the incentive for the managers of the strong divisions to shirk on effort by

setting low, easily achievable cash flow targets for their projects. This socialistic behavior is more severe when there is greater diversity in the quality of investment opportunities across divisions. One important advantage of our model over competing theoretical models based on managerial rent-seeking behavior (Scharfstein and Stein, 2000; Rajan, Servaes, and Zingales, 2000) is that our results are robust to the plausible assumption that firms can write managerial compensation contracts that depend on division cash flows. While our model makes several similar predictions to these competing theories, we make numerous novel testable predictions about the severity of the bias in internal capital markets, the effect of idiosyncratic risk on project evaluation, and the form of managerial compensation.

In the present model, we can show that the firm invests too much (little) in the weak (strong) division relative to first-best but we cannot compare investment in multi-division firms to single-division firms. For example, our multi-division firm is capital constrained so we would have to specify the constraint faced by similar focused firms if we want to compare investment policies. This is critical because one plausible benefit of integration is that it slackens financial constraints (e.g., Billett and Mauer, 2003). Similarly, we ignore the possibility that the CEO may be better informed about the firm's investment opportunity set than the external capital market (e.g., Stein, 1997) or that the CEO may privately benefit from allocating capital towards "pet" projects.

Appendix

Proof of Proposition 1. The first-best solution maximizes the expected total surplus (expectation over ϵ_i):

$$\max_{p_i(t),e_i(t)} E\Pi \equiv \sum_{i=1,2} p_i(t)\left[E_{\epsilon_i}V_i - 0.5\gamma e_i(t)^2\right] = \sum_{i=1,2} p_i(t)\left[\delta t_i + \theta e_i(t) - 0.5\gamma e_i(t)^2\right]$$

such that $p_i(t) \geq 0$ and $p_1(t) + p_2(t) \leq 1$.

The f.o.c. for $e_i(t)$ is: $\theta - \gamma e_i(t) = 0$, thus $e_i^{fb}(t) = \theta/\gamma$. The s.o.c. holds obviously. Substituting the expression for e_i^{fb} into the surplus yields:

$$E\Pi = \sum_{i=1,2} p_i(t)(\delta t_i + 0.5\theta^2/\gamma).$$

Therefore, $p_i^{fb}(t) = 1$ if $t_i > \max(t_j, -0.5\dfrac{\theta^2}{\delta\gamma})$; otherwise, $p_i^{fb}(t) = 0$. Q.E.D.

Proof of Proposition 2. Denote the firm's optimization problem as P. Our proof strategy for finding the optimal Bayesian-Nash mechanism follows the two-step approach of Laffont and Tirole (1986). In step 1, we consider a program (denoted R) that relaxes some constraints in P. With relaxed constraints, the firm should get at least as much expected payoff in R as in P. In step 2, we consider a program (denoted L) that uses compensation contracts linear in cash flows. Since L restricts to a narrower class of mechanisms, the firm can get at most the same expected payoff in L as in P. Finally, we demonstrate that the firm's expected payoffs from R and L are the same. Consequently, the optimal solution for L must be the optimal solution for P.

Step 1: Consider a relaxed program, denoted R, in which the firm can verify $\delta t_i + \theta e_i$. In this hypothetical case, assuming that the other manager will report her true type, the manager of the funded project who reports $\hat{t}_i$ must choose $\hat{e}_i$ such that $\delta t_i + \theta\hat{e}_i = \delta\hat{t}_i + \theta e_i(\hat{t}_i, t_j)$. Otherwise, the firm would be sure that the manager had either lied about her true type or didn't exert the required effort, and could punish her (arbitrarily) severely. Consequently, if the manager of a true type t_i reports $\hat{t}_i$, she must choose

$\hat{e}_i = e_i(\hat{t}_i, t_j) + \delta(\hat{t}_i - t_i)/\theta$, resulting in the cash flow $V_i(\hat{t}_i, t_j, \epsilon_i) = \delta\hat{t}_i + \theta e_i(\hat{t}_i, t_j) + \epsilon_i$. Compared with P, the (IC1) constraint for effort is completely relaxed in R.

In program R, if the manager of a true type t_i reports $\hat{t}_i$, her expected payoff is:

$$U_i(t_i, \hat{t}_i) = \int_{-\infty}^{\infty} p_i(\hat{t}_i, t_j) \left[E_{\epsilon_i} w_i(\hat{t}_i, t_j, V_i(\hat{t}_i, , t_j, \epsilon_i)) - 0.5\gamma(e_i(\hat{t}_i, t_j) + \frac{\delta(\hat{t}_i - t_i)}{\theta})^2 \right] d\Phi(t_j - \eta_j).$$

By the Envelope Theorem,

$$\begin{aligned} \frac{dU_i(t_i, t_i)}{dt_i} &= \frac{\partial U_i(t_i, \hat{t}_i)}{\partial t_i}\Big|_{\hat{t}_i = t_i} + \frac{\partial U_i(t_i, \hat{t}_i)}{\partial \hat{t}_i}\Big|_{\hat{t}_i = t_i} \\ &= \frac{\partial U_i(t_i, \hat{t}_i)}{\partial t_i}\Big|_{\hat{t}_i = t_i} = \int_{-\infty}^{\infty} \frac{\delta\gamma p_i(t_i, t_j) e_i(t_i, t_j)}{\theta} d\Phi(t_j - \eta_j). \end{aligned}$$

Integrating yields:

$$U_i(t_i) = U_i(-\infty) + \int_{-\infty}^{t_i} \int_{-\infty}^{\infty} \frac{\delta\gamma p_i(s, t_j) e_i(s, t_j)}{\theta} d\Phi(t_j - \eta_j) ds.$$

Imposing $U_i(-\infty) = 0$ from the (IR) condition and taking expectation yields:

$$\begin{aligned} EU_i &= \int_{-\infty}^{\infty} \int_{-\infty}^{t_i} \int_{-\infty}^{\infty} \frac{\delta\gamma p_i(s, t_j) e_i(s, t_j)}{\theta} d\Phi(t_j - \eta_j) ds d\Phi(t_i - \eta_i) \\ &= -\int_{-\infty}^{\infty} \int_{-\infty}^{t_i} \int_{-\infty}^{\infty} \frac{\delta\gamma p_i(s, t_j) e_i(s, t_j)}{\theta} d\Phi(t_j - \eta_j) ds d(1 - \Phi(t_i - \eta_i)) \\ &= -\left[(1 - \Phi(t_i - \eta_i)) \int_{-\infty}^{t_i} \int_{-\infty}^{\infty} \frac{\delta\gamma p_i(s, t_j) e_i(s, t_j)}{\theta} d\Phi(t_j - \eta_j) ds \right]_{-\infty}^{\infty} \\ &\quad + \int_{-\infty}^{\infty} (1 - \Phi(t_i - \eta_i)) \int_{-\infty}^{\infty} \frac{\delta\gamma p_i(t_i, t_j) e_i(t_i, t_j)}{\theta} d\Phi(t_j - \eta_j) dt_i \\ &= \int_{-\infty}^{\infty} \frac{1 - \Phi(t_i - \eta_i)}{\phi(t_i - \eta_i)} \int_{-\infty}^{\infty} \frac{\delta\gamma p_i(t) e_i(t)}{\theta} d\Phi(t_j - \eta_j) d\Phi(t_i - \eta_i) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\delta\gamma p_i(t) e_i(t)}{\theta} \mu(t_i - \eta_i) d\Phi(t_j - \eta_j) d\Phi(t_i - \eta_i). \end{aligned}$$

Substituting the expression $E(p_i w_i) = EU_i + 0.5\gamma E(p_i e_i^2)$ into the firm's expected payoff yields:

$$\begin{aligned} E\Pi^R &= \sum_{i=1,2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_{\epsilon_i}\Big[p_i(t) V_i - 0.5\gamma p_i(t) e_i(t)^2 - U_i \Big] d\Phi(t_1) d\Phi(t_2 - \eta) \qquad (1) \\ &= \sum_{i=1,2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_i(t) \left[\delta t_i + \theta e_i(t) - 0.5\gamma e_i(t)^2 - \frac{\delta\gamma e_i(t)}{\theta} \mu(t_i - \eta_i) \right] d\Phi(t_1) d\Phi(t_2 - \eta). \end{aligned}$$

Since program R provides the firm more flexibility to discipline its managers, it should yield the firm at least the same expected payoff as P does, i.e., $E\Pi^R \geq E\Pi^P$.

Step 2: Now consider a program, denoted L, in which the firm is constrained to use linear compensation contracts: $w_i(t, V_i) = a_i(t) + b_i(t)V_i$. Under this contract and assuming the other manager reports her true type, the manager of type t_i can get the following expected payoff by reporting $\hat{t}_i$ and choosing effort $\hat{e}_i$:

$$U_i(t_i, \hat{t}_i) = \int_{-\infty}^{\infty} p_i(\hat{t}_i, t_j) \left[a_i(\hat{t}_i, t_j) + b_i(\hat{t}_i, t_j)(\delta t_i + \theta \hat{e}_i) - 0.5\gamma \hat{e}_i^2 \right] d\Phi(t_j - \eta_j). \tag{2}$$

Ignoring the (NN) constraint for the moment, the f.o.c. for $\hat{e}_i$ is: $\hat{e}_i = \theta b_i(\hat{t}_i, t_j)/\gamma$. The s.o.c. holds obviously.

By the Envelope Theorem,

$$\begin{aligned} \frac{dU_i(t_i, t_i)}{dt_i} &= \frac{\partial U_i(t_i, \hat{t}_i)}{\partial t_i}\Big|_{\hat{t}_i = t_i} + \frac{\partial U_i(t_i, \hat{t}_i)}{\partial \hat{t}_i}\Big|_{\hat{t}_i = t_i} \\ &= \frac{\partial U_i(t_i, \hat{t}_i)}{\partial t_i}\Big|_{\hat{t}_i = t_i} = \int_{-\infty}^{\infty} \delta p_i(t_i, t_j) b_i(t_i, t_j) d\Phi(t_j - \eta_j). \end{aligned}$$

Integrating yields:

$$U_i(t_i) = U_i(-\infty) + \int_{-\infty}^{t_i} \int_{-\infty}^{\infty} \delta p_i(s, t_j) b_i(s, t_j) d\Phi(t_j - \eta_j) ds. \tag{3}$$

Imposing $U_i(-\infty) = 0$ from the (IR) condition and taking expectation yields:

$$\begin{aligned} EU_i &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta p_i(t) b_i(t) \mu(t_i - \eta_i) d\Phi(t_j - \eta_j) d\Phi(t_i - \eta_i) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\delta \gamma p_i(t) e_i(t)}{\theta} \mu(t_i - \eta_i) d\Phi(t_j - \eta_j) d\Phi(t_i - \eta_i), \end{aligned}$$

where the first equality obtains by integrating by parts, and the last equality obtains by substituting in the f.o.c. for e_i.

Substituting the expression $E(p_i w_i) = EU_i + 0.5\gamma E(p_i e_i^2)$ into the firm's expected payoff yields:

$$\begin{aligned} E\Pi^L &= \sum_{i=1,2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_{\epsilon_i} \left[p_i(t) V_i - 0.5\gamma p_i(t) e_i(t)^2 - U_i \right] d\Phi(t_1) d\Phi(t_2 - \eta) \qquad (4) \\ &= \sum_{i=1,2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_i(t) \left[\delta t_i + \theta e_i(t) - 0.5\gamma e_i(t)^2 - \frac{\delta \gamma e_i(t)}{\theta} \mu(t_i - \eta_i) \right] d\Phi(t_1) d\Phi(t_2 - \eta). \end{aligned}$$

Since program L allows only a subset of the mechanisms available in P, it must be that $E\Pi^L \leq E\Pi^P$. However, comparing Eqs. (1) and (4) shows that $E\Pi^L = E\Pi^R \geq E\Pi^P$. Therefore, $E\Pi^L = E\Pi^P$, so the optimal linear contract can not be further improved upon.

Optimal Bayesian-Nash Mechanism: Now we solve the optimal mechanism with linear contracts. The point-wise f.o.c. for e_i is

$$e_i(t) = \max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_i - \eta_i)\right).$$

The s.o.c. holds obviously.

Substituting the expression for e_i into the expected payoff yields:

$$E\Pi = \sum_{i=1,2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_i(t)\left[\delta t_i + 0.5\gamma\left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_i - \eta_i)\right)\right]^2\right] d\Phi(t_1)d\Phi(t_2 - \eta).$$

Define a function:

$$H(y, x) \equiv \delta y + 0.5\gamma\left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(y - x)\right)\right]^2.$$

Therefore, the firm chooses:

$$p_i(t) = \begin{cases} 1 & \text{if } H(t_i, \eta_i) > \max(0, H(t_j, \eta_j)) \ , \\ 0 & \text{otherwise.} \end{cases}$$

To derive other parts of the optimal mechanism, it is immediate that

$$b_i(t) = \frac{\gamma e_i(t)}{\theta} = \max\left(0, 1 - \frac{\delta\gamma}{\theta^2}\mu(t_i - \eta_i)\right).$$

It follows from Eqs. (2) and (3) that

$$\begin{aligned} U_i(t_i) &= \int_{-\infty}^{t_i} E_{t_j}\Big[\delta p_i(s, t_j) b_i(s, t_j)\Big] ds \\ &= E_{t_j}\left[p_i(t)\left(a_i(t) + b_i(t)(\delta t_i + \theta e_i(t)) - 0.5\gamma e_i(t)^2\right)\right], \end{aligned}$$

thus

$$E_{t_j}\left[p_i(t)a_i(t)\right] = \int_{-\infty}^{t_i} E_{t_j}\Big[\delta p_i(s, t_j) b_i(s, t_j)\Big] ds - E_{t_j}\left[p_i(t)\left(b_i(t)(\delta t_i + \theta e_i(t)) - 0.5\gamma e_i(t)^2\right)\right].$$

The Bayesian-Nash implementation of the firm's mechanism design problem only requires that the salary satisfy the above expectation expression.

We immediately obtain some monotonic properties of the mechanism. Note $\mu' < 0$ then $\partial H(t_i, \eta_i)/\partial t_i > 0$; therefore, $p_i(t)$ is non-decreasing in t_i. Also from $\mu' < 0$, $b_i(t)$ and $e_i(t)$ are non-decreasing in t_i. By standard arguments (Mirrlees, 1971), the (IC2) constraint is satisfied.

So far we have studied the Bayesian-Nash implementation of the firm's mechanism design problem. Since the optimal mechanism, in particular $p_i(t)$ and $b_i(t)$, is monotonic in both t_1 and t_2, by the results of Mookherjee and Reichelstein (1992), it can also be implemented in dominant strategies. We now consider a specific form of salary and show that the resulting mechanism can be implemented in dominant strategies.

Implementation in dominant strategies: Our proof strategy is as follows. In step 1, we consider a specific form of $a_i(t_i, t_j)$. Note that in the Bayesian-Nash mechanism solved above we only need to specify the expected salary (integrating over the t_j) assuming the other manager reports the truth, i.e., $\hat{t}_j = t_j$. In step 2, we express the manager's utility as a function of her true t_i, her reported $\hat{t}_i$, and the other manager's reported $\hat{t}_j$. Unlike the Bayesian-Nash mechanism we allow the other manager to misreport her t_j. In step 3, we show that, given every realization of $\hat{t}_j$ which may or may not be the true t_j, reporting the truth $\hat{t}_i = t_i$ maximizes the manager's utility and therefore is her dominant strategy.

Consider a specific form of a_i:

$$\begin{aligned} a_i(t) &= \int_{-\infty}^{t_i} \delta p_i(s, t_j) b_i(s, t_j) ds - \left(b_i(t)(\delta t_i + \theta e_i(t)) - 0.5\gamma e_i(t)^2 \right) \\ &= \int_{-\infty}^{t_i} \delta p_i(s, t_j) b_i(s, t_j) ds - b_i(t)(\delta t_i + 0.5\frac{\theta^2}{\gamma} b_i(t)), \end{aligned}$$

where the last equality obtains by substituting into the expression of the effort level $e_i = \theta b_i/\gamma$. Plugging the compensation into the payoff of the manager who observes t_i and reports $\hat{t}_i$, assuming the other manager observes t_j and reports $\hat{t}_j$, yields:

$$U(t_i, \hat{t}_i | \hat{t}_j) = p_i(\hat{t}_i, \hat{t}_j) \left[\int_{-\infty}^{\hat{t}_i} \delta p_i(s, \hat{t}_j) b_i(s, \hat{t}_j) ds + \delta b_i(\hat{t}_i, \hat{t}_j)(t_i - \hat{t}_i) \right].$$

Note this payoff is independent of the true t_j. We just need to show that the manager i finds it dominant strategy to report $\hat{t}_i = t_i$.

Given the manager i's belief about the other manager's reported $\hat{t}_j$, consider two cases:

- Suppose $p_i(t_i, \hat{t}_j) = 1$. The manager must report $\hat{t}_i$ such that $p_i(\hat{t}_i, \hat{t}_j) = 1$; otherwise, $p_i(\hat{t}_i, \hat{t}_j) = 0$ leads to $U(t_i, \hat{t}_i|\hat{t}_j) = 0$ which is obviously inferior to telling the truth. Thus,

$$U(t_i, \hat{t}_i|\hat{t}_j) = \int_{-\infty}^{\hat{t}_i} \delta p_i(s, \hat{t}_j) b_i(s, \hat{t}_j) ds + \delta b_i(\hat{t}_i, \hat{t}_j)(t_i - \hat{t}_i).$$

  Taking derivatives yields

$$\begin{aligned} \frac{\partial U(t_i, \hat{t}_i|\hat{t}_j)}{\partial \hat{t}_i} &= \delta p_i(\hat{t}_i, \hat{t}_j) b_i(\hat{t}_i, \hat{t}_j) - \delta b_i(\hat{t}_i, \hat{t}_j) + \delta(t_i - \hat{t}_i) \frac{\partial b_i(\hat{t}_i, \hat{t}_j)}{\partial \hat{t}_i} \\ &= \delta(t_i - \hat{t}_i) \frac{\partial b_i(\hat{t}_i, \hat{t}_j)}{\partial \hat{t}_i}, \end{aligned}$$

  which, from that $b_i(t)$ is non-decreasing in t_i, is non-negative (non-positive) for $\hat{t}_i < t_i$ ($\hat{t}_i > t_i$). The manager will report the true quality, i.e., $\hat{t}_i = t_i$.

- Suppose $p_i(t_i, \hat{t}_j) = 0$. If the manager reports $\hat{t}_i$ such that $p_i(\hat{t}_i, \hat{t}_j) = 1$, then

$$\begin{aligned} U(t_i, \hat{t}_i|\hat{t}_j) &= \int_{t_i}^{\hat{t}_i} \delta p_i(s, \hat{t}_j) b_i(s, \hat{t}_j) ds + \delta b_i(\hat{t}_i, \hat{t}_j)(t_i - \hat{t}_i) \\ &\leq \int_{t_i}^{\hat{t}_i} \delta b_i(\hat{t}_i, \hat{t}_j) ds + \delta b_i(\hat{t}_i, \hat{t}_j)(t_i - \hat{t}_i) \\ &= 0. \end{aligned}$$

  Here we have used the facts: $p_i(t) = 0, 1$ and $p_i(t)$ is non-decreasing in t_i; $b_i(t)$ is non-decreasing in t_i. If the manager reports $\hat{t}_i$ (particularly $\hat{t}_i = t_i$) such that $p_i(\hat{t}_i, \hat{t}_j) = 0$, she gets $U(t_i, \hat{t}_i|\hat{t}_j) = 0$. Thus she will report the true quality, i.e., $\hat{t}_i = t_i$.

To summarize, the manager i will report the true t_i for every belief of the other manager's reported $\hat{t}_j$. Therefore, the mechanism is implementable in dominant strategies.

Finally, to examine the monotonicity of $a_i(t)$, note that when $p_i(t) = 1$, i.e., $H(t_i, \eta_i) > \max(0, H(t_j, \eta_j))$, we have

$$\begin{aligned}
\frac{\partial a_i(t)}{\partial t_i} &= \delta p_i(t) b_i(t) - \delta b_i(t) - \frac{\partial b_i(t)}{\partial t_i}(\delta t_i + 0.5\frac{\theta^2}{\gamma} b_i(t)) - b_i(t) 0.5 \frac{\theta^2}{\gamma}\frac{\partial b_i(t)}{\partial t_i} \\
&= -\frac{\partial b_i(t)}{\partial t_i}(\delta t_i + 0.5\frac{\theta^2}{\gamma} b_i(t)) - b_i(t) 0.5 \frac{\theta^2}{\gamma}\frac{\partial b_i(t)}{\partial t_i} \\
&\leq -\frac{\partial b_i(t)}{\partial t_i}(\delta t_i + 0.5\frac{\theta^2}{\gamma} b_i(t)) \\
&\leq 0
\end{aligned}$$

since $b_i(t)$ is non-decreasing in t_i and

$$\delta t_i + 0.5\frac{\theta^2}{\gamma} b_i(t) = \delta t_i + 0.5\gamma\frac{\theta}{\gamma} e_i(t) \geq \delta t_i + 0.5\gamma e_i(t)^2 = H(t_i, \eta_i) \geq 0.$$

Therefore, $a_i(t)$ is non-increasing in t_i.

Q.E.D.

Proof of Corollary 1. From Proposition 2, if the firm chooses to finance project 2, it must be $H(t_2, \eta) > H(t_1, 0)$.

Define t_1^* such that $H(t_1^*, \eta) = H(t_1, 0)$, i.e.,

$$\delta t_1^* + 0.5\gamma \left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1^* - \eta)\right)\right]^2 = \delta t_1 + 0.5\gamma \left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right)\right]^2.$$

From $\mu' < 0$, $H(t_1^*, \eta)$ increases from $-\infty$ to ∞ as t_1^* increases from $-\infty$ to ∞; therefore, given t_1, t_1^* is uniquely determined. Since $H(y, \eta)$ increases in y, then from $H(t_2, \eta) > H(t_1, 0) = H(t_1^*, \eta)$ it must be $t_2 > t_1^*$; we just need to show $t_1^* \geq t_1$ with strict inequality for high t_1. As $\eta \to 0$, $t_1^* \to t_1$ thus we just need to show $dt_1^*/d\eta \geq 0$ with strict inequality for high t_1.

Note that

$$\begin{aligned}
\frac{\partial H(t_1^*, \eta)}{\partial t_1^*} &= \delta - \frac{\delta\gamma}{\theta}\mu'(t_1^* - \eta)\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1^* - \eta)\right) > 0 \\
\frac{\partial H(t_1^*, \eta)}{\partial \eta} &= \frac{\delta\gamma}{\theta}\mu'(t_1^* - \eta)\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1^* - \eta)\right) \leq 0,
\end{aligned}$$

where we use the fact $\mu' < 0$, and the last inequality holds strictly for high t_1 that implies $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1^* - \eta) > 0$. Therefore, $dt_1^*/d\eta = -[\partial H(t_1^*, \eta)/\partial\eta]/[\partial H(t_1^*, \eta)/\partial t_1^*] \geq 0$ with strict inequality for high t_1. Q.E.D.

Proof of Corollary 2. The monotonic property with respect to η follows immediately from the proof of Corollary 1. From the proof of Corollary 1, substituting $t_1^* = t_1 + \tau$ into $H(t_1^*, \eta) = H(t_1, 0)$ yields

$$\delta\tau + 0.5\gamma\left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1 + \tau - \eta)\right)\right]^2 = 0.5\gamma\left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right)\right]^2. \quad (5)$$

If $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1) \leq 0$, it is obvious that $\tau = 0$. It holds trivially that $\partial\tau/\partial\theta \geq 0$, $\partial\tau/\partial\delta \leq 0$, and $\partial\tau/\partial\gamma \leq 0$.

If $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1) > 0$ but $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1 + \tau - \eta) \leq 0$, Eq. (5) becomes

$$\delta\tau = 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right]^2 = 0.5\frac{\theta^2}{\gamma} - \delta\mu(t_1) + 0.5\frac{\gamma\delta^2}{\theta^2}\mu(t_1)^2.$$

It is clear that

$$\begin{aligned}
\frac{\partial\tau}{\partial\theta} &= \frac{\gamma}{\delta\theta}\left[(\frac{\theta}{\gamma})^2 - (\frac{\delta}{\theta}\mu(t_1))^2\right] > 0, \\
\frac{\partial\tau}{\partial\delta} &= -\frac{\gamma}{2\delta^2}\left[(\frac{\theta}{\gamma})^2 - (\frac{\delta}{\theta}\mu(t_1))^2\right] < 0, \\
\frac{\partial\tau}{\partial\gamma} &= -\frac{1}{2\delta}\left[(\frac{\theta}{\gamma})^2 - (\frac{\delta}{\theta}\mu(t_1))^2\right] < 0.
\end{aligned}$$

If $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1) > 0$ and $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1 + \tau - \eta) > 0$, Eq. (5) becomes

$$\delta\tau + 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1 + \tau - \eta)\right]^2 = 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right]^2.$$

Since $\tau > 0$ from the proof of Corollary 1, it must be

$$\begin{aligned}
\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1 + \tau - \eta) &< \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1), \\
\mu(t_1 + \tau - \eta) &> \mu(t_1).
\end{aligned}$$

Define

$$\begin{aligned} G &= \delta\tau + 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1+\tau-\eta)\right]^2 - 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right]^2 \\ &= \delta\tau - \delta\left[\mu(t_1+\tau-\eta) - \mu(t_1)\right] + 0.5\gamma\frac{\delta^2}{\theta^2}\left[\mu(t_1+\tau-\eta)^2 - \mu(t_1)^2\right]. \end{aligned}$$

Note

$$\begin{aligned} \frac{\partial G}{\partial \tau} &= \delta - \frac{\delta\gamma}{\theta}\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1+\tau-\eta)\right]\mu'(t_1+\tau-\eta) > 0, \\ \frac{\partial G}{\partial \theta} &= -\frac{\gamma\delta^2}{\theta^3}\left[\mu(t_1+\tau-\eta)^2 - \mu(t_1)^2\right] < 0, \\ \frac{\partial G}{\partial \delta} &= \frac{1}{\delta}\left[\delta\tau - \delta\left[\mu(t_1+\tau-\eta) - \mu(t_1)\right] + \gamma\frac{\delta^2}{\theta^2}\left[\mu(t_1+\tau-\eta)^2 - \mu(t_1)^2\right]\right] \\ &= \frac{\gamma\delta}{2\theta^2}\left[\mu(t_1+\tau-\eta)^2 - \mu(t_1)^2\right] > 0, \\ \frac{\partial G}{\partial \gamma} &= \frac{\delta^2}{2\theta^2}\left[\mu(t_1+\tau-\eta)^2 - \mu(t_1)^2\right] > 0. \end{aligned}$$

Therefore, $\partial\tau/\partial\theta = -(\partial G/\partial\theta)/(\partial G/\partial\tau) > 0$, $\partial\tau/\partial\delta = -(\partial G/\partial\delta)/(\partial G/\partial\tau) < 0$, and $\partial\tau/\partial\gamma = -(\partial G/\partial\gamma)/(\partial G/\partial\tau) < 0$. Q.E.D.

Proof of Corollary 3. Note $H(\underline{t}_i^{sb}, \eta_i) = 0$ but

$$H(\underline{t}^{fb}, \eta_i) = \delta\underline{t}^{fb} + 0.5\gamma\left[\max(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(\underline{t}^{fb} - \eta_i))\right]^2 < \delta\underline{t}^{fb} + 0.5\theta^2/\gamma = 0.$$

From $\partial H(y,x)/\partial y > 0$ since $\mu' < 0$, it must be $\underline{t}_i^{sb} > \underline{t}^{fb}$.

To compare the hurdle rates for the two divisions, consider the following cases:

- If $H(0,0) = 0$, then $\underline{t}_1^{sb} = 0$. It follows from $\partial H(0,x)/\partial x \leq 0$ that $H(0,\eta) \leq 0$; in particular $H(0,\eta) = 0$. Thus $\underline{t}_2^{sb} = 0$, and therefore $\underline{t}_1^{sb} = \underline{t}_2^{sb}$.

- If $H(0,0) > 0$, then from $\partial H(y,0)/\partial y > 0$ it must be $\underline{t}_1^{sb} < 0$ which implies $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(\underline{t}_1^{sb}) > 0$. It follows immediately that

$$\begin{aligned} H(\underline{t}_1^{sb}, \eta) &= \delta\underline{t}_1^{sb} + 0.5\gamma\left[\max(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(\underline{t}_1^{sb} - \eta))\right]^2 \\ &< \delta\underline{t}_1^{sb} + 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(\underline{t}_1^{sb})\right]^2 = H(\underline{t}_1^{sb}, 0) = 0 = H(\underline{t}_2^{sb}, \eta) \end{aligned}$$

where the inequality obtains from $\mu' < 0$. From $\partial H(y,\eta)/\partial y > 0$, it must be $\underline{t}_2^{sb} > \underline{t}_1^{sb}$.

To summarize, $\underline{t}_2^{sb} \geq \underline{t}_1^{sb}$ where the inequality holds strictly for certain parameter values. Q.E.D.

Proof of Corollary 4. Write $H(t_2,\eta) = H(t_2-\eta,0) + \delta\eta$. Then it can be expressed

$$\begin{aligned}
\bar{b}_1 &= E_{\{t_1,t_2\}}\Big[\max(0, 1-\frac{\delta\gamma}{\theta^2}\mu(t_1))|H(t_1,0) > \max(0,H(t_2,\eta))\Big] \\
&= E_{\{t_1,t_2-\eta\}}\Big[\max(0, 1-\frac{\delta\gamma}{\theta^2}\mu(t_1))|H(t_1,0) > \max(0,H(t_2-\eta,0)+\delta\eta)\Big] \\
&= E_{\{t_1,x\}}\Big[\max(0, 1-\frac{\delta\gamma}{\theta^2}\mu(t_1))|H(t_1,0) > \max(0,H(x,0)+\delta\eta)\Big] \\
&= E_x\Big[E_{t_1}\Big[\max(0, 1-\frac{\delta\gamma}{\theta^2}\mu(t_1))|H(t_1,0) > \max(0,H(x,0)+\delta\eta)\Big]\Big]
\end{aligned}$$

where $x \equiv t_2 - \eta \sim N(0,1)$ and $x \perp t_1$. Note that $\max(0, 1-\frac{\delta\gamma}{\theta^2}\mu(t_1))$ is non-decreasing in t_1 and is independent of η; and $H(t_1,0)$ increases in t_1 and is independent of η. As η increases, given x, $E_{t_1}\Big[\max(0, 1-\frac{\delta\gamma}{\theta^2}\mu(t_1))|H(t_1,0) > \max(0,H(x,0)+\delta\eta)\Big]$ must increase since it puts more weight on high t_1. Therefore, $\bar{b}_1$ increases in η.

Now consider the total compensation. Write

$$\begin{aligned}
\bar{w}_i &= E_{\{t_1,t_2\}}\Big[a_i(t) + b_i(t)E_{\epsilon_i}V_i|H(t_i,\eta_i) > \max(0,H(t_j,\eta_j))\Big] \\
&= \frac{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))}\Big[a_i(t)+b_i(t)E_{\epsilon_i}V_i\Big]d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)}{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))}d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)} \\
&= \frac{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))}\Big[0.5\gamma e_i(t)^2 + \int_{-\infty}^{t_i}\delta p_i(s,t_j)b_i(s,t_j)ds\Big]d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)}{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))}d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)} \\
&= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[p_i(t)0.5\gamma e_i(t)^2 + p_i(t)\int_{-\infty}^{t_i}\delta p_i(s,t_j)b_i(s,t_j)ds\Big]d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)}{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))}d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[p_i(t)0.5\gamma e_i(t)^2 + p_i(t)\delta p_i(t)b_i(t)\mu(t_i-\eta_i)\Big]d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)}{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))} d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)} \\
&= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_i(t)\Big[0.5\gamma e_i(t)^2 + \frac{\delta\gamma e_i(t)}{\theta}\mu(t_i-\eta_i)\Big]d\Phi(t_i-\eta_i)}{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))} d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)} \\
&= \frac{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))} 0.5\gamma\max\Big[0,(\frac{\theta}{\gamma})^2-(\frac{\delta}{\theta}\mu(t_i-\eta_i))^2\Big]d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)}{\int_{H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))} d\Phi(t_i-\eta_i)d\Phi(t_j-\eta_j)} \\
&= E_{\{t_1,t_2\}}\left[0.5\gamma\max\Big[0,(\frac{\theta}{\gamma})^2-(\frac{\delta}{\theta}\mu(t_i-\eta_i))^2\Big]|H(t_i,\eta_i)>\max(0,H(t_j,\eta_j))\right].
\end{aligned}
$$

Here we have used the facts: $p_i(t)$ has piece-wise $\partial p_i(t)/\partial t_i = 0$; $p_i(t)^2 = p_i(t)$; the fifth equality obtains by integrating by parts. Thus,

$$
\bar{w}_1 = E_{\{t_1,t_2\}}\left[0.5\gamma\max\Big[0,(\frac{\theta}{\gamma})^2-(\frac{\delta}{\theta}\mu(t_1))^2\Big]|H(t_1,0)>\max(0,H(t_2,\eta))\right].
$$

Note that $0.5\gamma\max\Big[0,(\frac{\theta}{\gamma})^2-(\frac{\delta}{\theta}\mu(t_1))^2\Big]$ is non-decreasing in t_1 and is independent of η. Then similar to the above reasoning for performance pay, we obtain that $\bar{w}_1$ increases in η.

Q.E.D.

Proof of Corollary 5. From Proposition 3, if the firm chooses to fund project 2, it must be $R(t_2,\sigma) > R(t_1,1)$.

Define t_1^{**} such that $R(t_1^{**},\sigma) = R(t_1,1)$, i.e.,

$$
\delta t_1^{**} + 0.5\gamma\left[\max\left(0,\frac{\theta}{\gamma}-\frac{\delta\sigma}{\theta}\mu(\frac{t_1^{**}}{\sigma})\right)\right]^2 = \delta t_1 + 0.5\gamma\left[\max\left(0,\frac{\theta}{\gamma}-\frac{\delta}{\theta}\mu(t_1)\right)\right]^2.
$$

From $\mu' < 0$, $R(t_1^{**},\eta)$ increases from $-\infty$ to ∞ as t_1^{**} increases from $-\infty$ to ∞; therefore, given t_1, t_1^{**} is uniquely determined. Since $R(y,\sigma)$ increases in y, then from $R(t_2,\sigma) > R(t_1,1) = R(t_1^{**},\sigma)$ it must be $t_2 > t_1^{**}$; we just need to show $t_1^{**} \geq t_1$ with strict inequality for high t_1. Note as $\sigma \to 1$, $t_1^{**} \to t_1$; we just need to show $dt_1^{**}/d\sigma \geq 0$ with strict inequality for high t_1.

Before we proceed, we want to show that if $\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1^{**}}{\sigma}) > 0$, then $t_1^{**} > 0$. Suppose $t_1^{**} \leq 0$. It follows from $\mu' < 0$ that

$$\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(0) \geq \frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t^{**}}{\sigma}) > 0.$$

This contradicts Assumption (A1). Therefore, it must be $t_1^{**} > 0$.

Note

$$\begin{aligned}
\frac{\partial R(t_1^{**},\sigma)}{\partial t_1^{**}} &= \delta - \frac{\delta\gamma}{\theta}\mu'(\frac{t_1^{**}}{\sigma})\max\left(0, \frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1^{**}}{\sigma})\right) > 0 \\
\frac{\partial R(t_1^{**},\sigma)}{\partial \sigma} &= -\frac{\delta\gamma}{\theta}(\mu(\frac{t_1^{**}}{\sigma}) - \frac{t_1^{**}}{\sigma}\mu'(\frac{t_1^{**}}{\sigma}))\max\left(0, \frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1^{**}}{\sigma})\right) \leq 0,
\end{aligned}$$

where the last inequality follows from $\mu' < 0$ and $t^{**} > 0$ when $\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1^{**}}{\sigma}) > 0$, and holds strictly for high t_1 that implies $\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1^{**}}{\sigma}) > 0$. Therefore, $dt_1^{**}/d\sigma = -[\partial R(t_1^{**},\sigma)/\partial\sigma]/[\partial R(t_1^{**},\sigma)/\partial t_1^{**}] \geq 0$ with strict inequality for high t_1. Q.E.D.

Proof of Corollary 6. The monotonic property with respect to σ follows immediately from the proof of Corollary 5. We just need to the properties for other parameters.

From the proof of Corollary 5, substituting $t_1^{**} = t_1 + \rho$ into $R(t_1^{**},\sigma) = R(t_1, 1)$ yields

$$\delta\rho + 0.5\gamma\left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1+\rho}{\sigma})\right)\right]^2 = 0.5\gamma\left[\max\left(0, \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right)\right]^2. \quad (6)$$

If $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1) \leq 0$, it is obvious that $\rho = 0$. It holds trivially that $\partial\rho/\partial\theta \geq 0$, $\partial\rho/\partial\delta \leq 0$, and $\partial\rho/\partial\gamma \leq 0$.

If $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1) > 0$ but $\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1+\rho}{\sigma}) \leq 0$, Eq. (6) becomes

$$\delta\rho = 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right]^2 = 0.5\frac{\theta^2}{\gamma} - \delta\mu(t_1) + 0.5\frac{\gamma\delta^2}{\theta^2}\mu(t_1)^2.$$

It is clear that

$$\frac{\partial\rho}{\partial\theta} = \frac{\gamma}{\delta\theta}\left[(\frac{\theta}{\gamma})^2 - (\frac{\delta}{\theta}\mu(t_1))^2\right] > 0,$$

$$\begin{aligned}
\frac{\partial \rho}{\partial \delta} &= -\frac{\gamma}{2\delta^2}\left[(\frac{\theta}{\gamma})^2 - (\frac{\delta}{\theta}\mu(t_1))^2\right] < 0, \\
\frac{\partial \rho}{\partial \gamma} &= -\frac{1}{2\delta}\left[(\frac{\theta}{\gamma})^2 - (\frac{\delta}{\theta}\mu(t_1))^2\right] < 0.
\end{aligned}$$

If $\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1) > 0$ and $\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1+\rho}{\sigma}) > 0$, Eq. (6) becomes

$$\delta\rho + 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1+\rho}{\sigma})\right]^2 = 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right]^2.$$

Since $\rho > 0$ from the proof of Corollary 5, it must be

$$\begin{aligned}
\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1+\rho}{\sigma}) &< \frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1), \\
\sigma\mu(\frac{t_1+\rho}{\sigma}) &> \mu(t_1).
\end{aligned}$$

Define

$$\begin{aligned}
K &= \delta\rho + 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1+\rho}{\sigma})\right]^2 - 0.5\gamma\left[\frac{\theta}{\gamma} - \frac{\delta}{\theta}\mu(t_1)\right]^2 \\
&= \delta\rho - \delta\left[\sigma\mu(\frac{t_1+\rho}{\sigma}) - \mu(t_1)\right] + 0.5\gamma\frac{\delta^2}{\theta^2}\left[\sigma^2\mu(\frac{t_1+\rho}{\sigma})^2 - \mu(t_1)^2\right].
\end{aligned}$$

Note

$$\begin{aligned}
\frac{\partial K}{\partial \rho} &= \delta - \frac{\delta\gamma}{\theta}\left[\frac{\theta}{\gamma} - \frac{\delta\sigma}{\theta}\mu(\frac{t_1+\rho}{\sigma})\right]\mu'(\frac{t_1+\rho}{\sigma}) > 0, \\
\frac{\partial K}{\partial \theta} &= -\frac{\gamma\delta^2}{\theta^3}\left[\sigma^2\mu(\frac{t_1+\rho}{\sigma})^2 - \mu(t_1)^2\right] < 0, \\
\frac{\partial K}{\partial \delta} &= \frac{1}{\delta}\left[\delta\rho - \delta\left[\sigma\mu(\frac{t_1+\rho}{\sigma}) - \mu(t_1)\right] + \gamma\frac{\delta^2}{\theta^2}\left[\sigma^2\mu(\frac{t_1+\rho}{\sigma})^2 - \mu(t_1)^2\right]\right] \\
&= \frac{\gamma\delta}{2\theta^2}\left[\sigma^2\mu(\frac{t_1+\rho}{\sigma})^2 - \mu(t_1)^2\right] > 0, \\
\frac{\partial K}{\partial \gamma} &= \frac{\delta^2}{2\theta^2}\left[\sigma^2\mu(\frac{t_1+\rho}{\sigma})^2 - \mu(t_1)^2\right] > 0.
\end{aligned}$$

Therefore, $\partial\rho/\partial\theta = -(\partial K/\partial\theta)/(\partial K/\partial\rho) > 0$, $\partial\rho/\partial\delta = -(\partial K/\partial\delta)/(\partial K/\partial\rho) < 0$, and $\partial\rho/\partial\gamma = -(\partial K/\partial\gamma)/(\partial K/\partial\rho) < 0$. Q.E.D.

Proof of Corollary 7. Note $R(\underline{t}_i^{sb}, \sigma_i) = 0$ but

$$R(\underline{t}^{fb}, \sigma_i) = \delta \underline{t}^{fb} + 0.5\gamma \left[\max(0, \frac{\theta}{\gamma} - \frac{\delta \sigma_i}{\theta} \mu(\frac{\underline{t}^{fb}}{\sigma_i})) \right]^2 < \delta \underline{t}^{fb} + 0.5\theta^2/\gamma = 0.$$

From $\partial R(y, x)/\partial y > 0$ since $\mu' < 0$, it must be $\underline{t}_i^{sb} > \underline{t}^{fb}$.

Now we compare the hurdle rates for the two division. Under Assumption (A1), $R(0, \sigma) = 0$; therefore, $\underline{t}_2^{sb} = 0$.

Consider the following cases for $\underline{t}_1^{sb}$:

- If $R(0, 1) = 0$, then $\underline{t}_1^{sb} = 0$; therefore, $\underline{t}_1^{sb} = \underline{t}_2^{sb}$.
- If $R(0, 1) > 0$, then from $\partial R(y, 1)/\partial y > 0$ it must be $\underline{t}_1^{sb} < 0$; therefore, $\underline{t}_1^{sb} < \underline{t}_2^{sb}$.

To summarize, $\underline{t}_2^{sb} \geq \underline{t}_1^{sb}$ where the inequality holds strictly for certain parameter values.

Q.E.D.

References

[1] Alchian, A.A., 1969, "Corporate management and property rights," in H. Manne, ed., Economic Policy and the Regulation of Corporate Securities. American Enterprise Institute, Washington D.C., 337-360.

[2] Berger, P., and E. Ofek, 1995, "Diversification's effect on firm value," *Journal of Financial Economics*, 37, 39-65.

[3] Berkovitch, E., and R. Israel, 2004, "Why the NPV criterion does not maximize NPV," *Review of Financial Studies*, 17, 239-255.

[4] Bernardo, A.E., H. Cai, and J. Luo, 2001, "Capital budgeting and compensation with asymmetric information and moral hazard," *Journal of Financial Economics*, 61, 311-344.

[5] Bernardo, A.E., H. Cai, and J. Luo, 2004, "Capital budgeting in multi-division firms: Information, agency, and incentives, *Review of Financial Studies* 17, 739-767.

[6] Bernardo, A.E., and B. Chowdhry, 2002, "Resources, real options, and corporate strategy," *Journal of Financial Economics*, 63, 211-234.

[7] Billett, M.T., and D.C. Mauer, 2003, "Cross-subsidies, external financing constraints, and the contribution of the internal capital market to firm value," *Review of Financial Studies*, 16, 1167-1201.

[8] Brealey, R.A., and S.C. Myers, 2003. Principles of Corporate Finance, 7th edition. McGraw-Hill Irwin, New York.

[9] Campa, J.M., and S. Kedia, 2002, "Explaining the diversification discount," *Journal of Finance*, 57, 1731-1762.

[10] Chevalier, J., 2000, "What do we know about cross-subsidization? Evidence from the investment policies of merging firms," working paper, Yale University.

[11] Corbett, J., and T. Jenkinson, 1997, "How is investment financed? A study of Germany, Japan, the United Kingdom, and the United States," *The Manchester School* 65 (supplement), 69-93.

[12] Garcia, D., 2002, "Optimal contracts with privately informed agents and active principals," working paper, Dartmouth College.

[13] Gertner, R., E. Powers, and D.S. Scharfstein, 2002, "Learning about internal capital markets from corporate spin-offs," *Journal of Finance*, 57, 2479-2506.

[14] Goel, A.M., V. Nanda, and M.P. Narayanan, 2004, "Career concerns and resource allocation in conglomerates," *Review of Financial Studies* , 17, 99-128.

[15] Gomes, J., and D. Livdan, 2004, "Optimal diversification: Reconciling theory and evidence," *Journal of Finance*, 59, 507-536.

[16] Harris, M., and A. Raviv, 1996, "The capital budgeting process: Incentives and information," *Journal of Finance*, 51, 1139-1174.

[17] Harris, M., and A. Raviv, 1998, "Capital budgeting and delegation," *Journal of Financial Economics*, 50, 259-289.

[18] Holmstrom, B., and J. Ricart i Costa, 1986, "Managerial incentives and capital management," *Quarterly Journal of Economics*, 101, 835-860.

[19] Khanna, N., and S. Tice, 2001, "The bright side of internal capital markets," *Journal of Finance*, 56, 1489-1526.

[20] Laffont, J.-J., and J. Tirole, 1986, "Using cost observation to regulate firms," *Journal of Political Economy*, 94, 614-641.

[21] Laffont, J.-J., and J. Tirole, 1993, *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge.

[22] Lamont, O., 1997, "Cash flow and investment: Evidence from internal capital markets," *Journal of Finance*, 52, 83-109.

[23] Lamont, O., and C. Polk, 2002, "Does diversification destroy value? Evidence from industry shocks," *Journal of Financial Economics*, 63, 51-77.

[24] Lang, L.H.P., and R. Stulz, 1994, "Tobin's q, corporate diversification, and firm performance," *Journal of Political Economy*, 102, 1248-1280.

[25] Maksimovic, V., and G.M. Phillips, 2002, "Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence," *Journal of Finance*, 57, 721-767.

[26] Marino, A.M., and J.G. Matsusaka, 2005, "Decision processes, agency problems, and information: An economic analysis of capital budgeting procedures," *Review of Financial Studies* 18, 301-325.

[27] Milgrom, P., and J. Roberts, 1992, *Economics, Organizations and Management*, Prentice Hall, Englewood Cliffs.

[28] Mirrlees, J.A., 1971, "An exploration in the theory of optimum income taxation," *Review of Economic Studies*, 38, 135-208

[29] Mookherjee, D., and S. Reichelstein, 1992, "Dominant strategy implementation of Bayesian incentive compatible allocation rules," *Journal of Economic Theory*, 56, 378-99.

[30] Ozbas, O., 2004, "Integration, organizational process, and allocation of resources," *Journal of Financial Economics*, forthcoming.

[31] Palia, D., 1999, "Corporate governance and the diversification discount: Evidence from panel data," working paper, Rutgers University.

[32] Poterba, J., and L. Summers, 1995, "A CEO survey of U.S. companies' time horizons and hurdle rates," *Sloan Management Review*, 37, 43-53.

[33] Rajan, R., H. Servaes, and L. Zingales, 2000, "The cost of diversity: The diversification discount and inefficient investment," *Journal of Finance*, 55, 35-80.

[34] Scharfstein, D.S., 1998, "The dark side of internal capital markets II: Evidence from diversified conglomerates," NBER Working Paper 6352, Cambridge, MA.

[35] Scharfstein, D.S., and J.C. Stein, 2000, "The dark side of internal capital markets: Divisional rent-seeking and inefficient investment," *Journal of Finance*, 55, 2537-2564.

[36] Shin, H., and R. Stulz, 1998, "Are internal capital markets efficient?", *Quarterly Journal of Economics*, 113, 531-552.

[37] Stein, J.C., 1997, "Internal capital markets and the competition for corporate resources," *Journal of Finance*, 52, 111-133.

[38] Villalonga, 2004, "Diversification discount or premium: New evidence from the Business Information Tracking Series ," *Journal of Finance*, 59, 479-506.

[39] Whited, T., 2001, "Is it inefficient investment that causes the diversification discount?," *Journal of Finance*, 56, 1667-1692.

[40] Williamson, O.E., 1975. Markets and Hierarchies: Analysis and Antitrust Implications. Collier Macmillan Publishers, Inc. : New York.