Available online at www.sciencedirect.com

ScienceDirect

J. Finan. Intermediation 17 (2008) 357–378

Journal of Financial Intermediation

www.elsevier.com/locate/jfi

The role of profit-based and stock-based components in incentive compensation

Yuri Khoroshilov [a], M.P. Narayanan [b,*]

[a] *University of Ottawa, Canada*

[b] *Ross School of Business, University of Michigan, Ann Arbor, MI 48109-1234, USA*

Received 28 June 2006

Available online 16 March 2008

Abstract

This paper argues that the presence of both profit-based and stock price-based components in compensation contracts provides senior managers the incentive to optimally allocate effort to both implementing previously devised strategies that provide current profits and to formulating new strategies that create shareholder value. If managers are concerned about their reputation and if outcomes of strategy implementation are more informative about their ability than outcomes of strategy formulation, compensation based only on profit will incent managers to boost their reputation by over-allocating effort to strategy implementation. To restore the balance the contract needs to contain some stock-based compensation.

JEL classification: D86; G34

Keywords: Executive compensation; Stock-based compensation; Profit-based compensation

1. Introduction

There has been considerable interest recently in various aspects of executive compensation. While there is general agreement among both academics and practitioners that executive compensation should contain a significant incentive component to align the executives' interests to that of shareholders, there is little consensus regarding the size or the structure of incentive compen-

* Corresponding author.
E-mail address: mpn@umich.edu (M.P. Narayanan).

1042-9573/$ – see front matter
doi:10.1016/j.jfi.2008.02.005

sation. In this paper, we focus on the structure of incentive compensation, namely, the trade-offs between profit-based incentive compensation and stock-based incentive compensation.

There has been significant intertemporal variation within industries, and cross-sectional variation across industries, in the mixture of profit-based and stock-based incentive compensation. While there was a considerable increase in stock-based compensation, particularly stock options, in the last decade of the twentieth century (Murphy, 2004), recently several companies have been revising their executive compensation packages by reducing stock options and increasing restricted stock grants and profit-based bonuses (Plitch, 2004; Boyle, 2004). In addition, there is wide variation in the proportions of profit-based compensation and stock compensation across different industries as reported in the 2006 CEO Compensation Survey conducted by the Wall Street Journal and Mercer Human Resource Consulting.

In order to understand the temporal and cross-sectional variations in the structure of incentive compensation, we need to clarify the roles of profit-based and stock-based compensation in aligning managerial interests to that of shareholders. Even though using a mixture of profit-based and stock-based compensation is customary, it is not clear why *both* forms are needed. Indeed, since the stock price should equal the present value of future dividends, a manager who plans to stay with the firm will naturally care about the firm's expected future profit. Thus, de facto, she will care about the current stock price and there is no need to explicitly include a stock-based component in the manager's incentive contract: a profit-based compensation contract alone can provide a manager with the right incentives. In our paper, we provide an explanation for the use of both profit-based and stock-based compensations even if the manager is expected to stay with the firm.

Consider a manager with ability unknown to everyone including herself. The firm's profit is directly related to the manager's effort and her ability. The manager needs to allocate effort to two types of activities: formulating strategy (i.e., discovering new projects) and implementing strategy (i.e., implementing discovered projects). Effort exerted toward discovery does not lead to immediate cash flows but increases the probability that an implementable project will be available in the future. Implementation of a previously discovered project results in immediate cash flows. One of the main premises of our model is that implementation outcomes provide more precise information about the manager's ability than outcomes of discovery. Indeed, implementation of existing projects results in immediate cash flows that are observable to everyone, while the market can only obtain imprecise information about the success of discovery efforts. A second premise of our model is that manager cares about her reputation, i.e., her ability as perceived by the market.

In the absence of reputation concerns shareholders may provide the manager with a simple profit-based incentive contract: such a contract provides her the incentive to maximize the present value of the firm's profits. Reputation concerns, however, distort the manager's incentives. In order to boost her perceived ability the manager is willing to exert more than the optimal effort toward the more informative activity of implementation. Shareholders, who would like to restore the balance in effort allocation, need to provide her the incentive to expend less effort toward implementation. Since implementation results in immediate cash flows, shareholders can reduce the manager's incentive to expend implementation effort by reducing the profit-based compensation. This reduction, however, will also reduce the manager's expected future payoff from discovery, making her less willing to expend discovery effort. To increase the manager's incentive to discover new projects without affecting her incentive to implement them, shareholders need to offer her a stock-based component as well. If compensation is based exclusively on stock performance, however, the manager's wage becomes independent of profits. Her incentive to implement is then driven entirely by her reputation concern, which may be insufficient to incent the manager to exert the optimal level of implementation effort.

Our model provides several empirical implications relating the structure of the optimal incentive contract to industry characteristics. First, the greater the demand for managerial labor the greater the stock-based compensation. In environments where demand for managerial labor is high, managers are more concerned about perceived ability. As a result, in such environments, shareholders need to provide more stock-based and less profit-based compensation. Second, if the market knows the manager's ability with greater precision, the manager cannot easily mislead the market about her ability and, therefore, shareholders may offer her less stock-based compensation. A result that follows is that we should expect more stock-based compensation in industries with fast-changing environments where managers' ability to generate cash flows may fluctuate more over time due to the need to learn new skills and adjust to new technologies. Third, firms in industries where it is relatively difficult to ascertain the success probability of strategies at the formulation stage should offer the manager more stock-based and less profit-based compensation. As the relative signal-to-noise ratio of discovery outcomes declines, the manager's incentive to exert more than optimal effort toward implementation increases and, to offset it, the profit-based compensation has to be decreased. However, lower profit-based compensation reduces the manager's expected payoff from discovery, which reduces her incentive to exert discovery effort. To restore this incentive, the stock-based compensation has to be increased. And finally, in knowledge-based industries both profit-based and stock-based compensations will be greater. In such industries the manager receives a higher proportion of the profit because her contribution to profits is greater. As the manager's share of the profit increases, however, the stock price, which is net of wage, becomes less sensitive to discovery effort, and to compensate, the relative weight of stock compensation needs to be greater.

There are other reasons why shareholders use both profit-based and stock-based compensation. Banker and Datar (1989) consider a one-period model in which they show that shareholders use a mixture of bonus-based and stock-based compensation in order to provide better risk-sharing for a risk-averse manager and that the relative weight of the bonus-based compensation is higher when accounting returns contain less noise than stock prices. Holmstrom and Tirole (1993) shows that both profit-based and stock-based compensation are required for a risk-averse manager when there is a rational speculator on the market who can independently investigate the potential value of the firm. Dechow and Sloan (1991) considers a hypothesis consistent with our model, i.e., a manager with profit-based compensation who plans to quit her job will be biased toward short-run projects. The paper finds that CEOs reduce the R&D expenditures in their terminal year in the office. The bias toward the short-run project, however, disappears if the CEO stays in the firm long enough to reap the benefits of the long-term investment. Bushman and Indjejikian (1993) seek the optimal incentive contract for a risk-averse manager with two-dimensional effort that affects the firm's current and future profits. In their one-period model, the manager's contract is set as a function of both accounting data and stock price, which are noisy indicators of firm's value (defined as the sum of current and future profits). In this setup, a mixture of bonus-based and stock-based compensation can be used to filter some of the noise in the firm's value and allow better risk-sharing between the risk-averse manager and risk-neutral shareholders. In addition, since the accounting data and stock price provides information about different portions of the firm's value (namely, accounting data provides no information about the future firm's profit), using a mixture can allow shareholders to influence the manager's effort allocations for producing current and future profits and make her allocate more effort into the riskier activity that is potentially more profitable. Our explanation does not use risk-sharing arguments. In this paper we investigate how managerial reputation concerns affect the choice between stock-based and bonus-based compensation and show that even in the absence of risk-sharing (e.g., when managers are risk-neutral or may diversify most of the unsystematic risk) the optimal compensation should include both profit-based and stock-based components. By focusing on the differences in the industry structure, we are able to provide some cross-industry implications.

One of the premises of our model, that reputation concerns may affect the manager's behavior, is not new and is well-studied in the literature.[1] Our model uses the multitasking idea of Holmstrom and Milgrom (1991), and similar to Goel et al. (2004), models reputation concern in a multiple-dimensional action environment.

The rest of the paper proceeds as follows. Section 2 describes the model. Section 3 provides the first-best effort allocations. Section 4 describes the equilibrium incentive contract and how it is affected by the manager's reputation concerns. Section 5 discusses the empirical implications and Section 6 concludes.

2. The model

2.1. Technology

Consider a firm run by a manager in a three-period model in which the periods are numbered by $t = 1, 2$, and 3. In order to isolate the reputation effect from the risk-sharing effect, and to obtain an analytic solution, we assume that the manager is risk-neutral. At the beginning of the first and second periods, the manager needs to exert some effort d_t, $t = 1, 2$, and allocate investors' capital towards discovery of new projects. This discovery effort succeeds with some probability, resulting in a project that is available for implementation in period $t + 1$. If a project is available for implementation in periods 2 and 3, the manager needs to exert effort i_t and allocate capital to implement the project.

Examples of discovery activities include formulating a new business strategy or setting a new direction for the company, identifying targets for acquisition, identifying the need to develop new products or modify existing products to meet evolving customer demand, and recognizing the need to improve the organizational capability for better operating efficiency and to better serve the customer. Implementation of the above discoveries involves activities such as communicating the discovery to subordinates and motivating them to implement it, extracting value from an acquisition by integrating the acquired company, shepherding the development of a new product to completion, and installing new systems that will improve the organizational capability. Discovery by itself does not produce profits; only implementation of previously discovered projects produces profits.

For simplicity, each project is a one-period project: that is, if a project is discovered in period $t - 1$, it can be implemented and produces profits only in period t. The profit from implementing the project in period t is stochastic and depends on the manager's ability, the amount of investment, and the managerial effort expended towards implementation. In particular, we assume that each unit of effort must be matched by one unit of investment in order to generate profits and vice versa. The firm's profit in period t is given by,

$$\pi_t = N_t f(i_t) + a + \varepsilon_t, \quad t = 2, 3, \tag{1}$$

[1] See, for example, Fama (1980), Gibbons and Murphy (1990), Holmstrom (1999), Hirshleifer and Thakor (1992), Milbourn et al. (2001), and Milbourn (2003).

where π_t is the firm's profit in period t, a is the manager's ability, and ε_t is a normally distributed stochastic noise term with zero mean and precision h_ε. The production function $f(\cdot)$ is increasing and strictly concave. The indicator function N_t equals 1 if the manager has a project to implement in period t (i.e., a project that was discovered in period $t-1$), and 0 otherwise. To simplify the notation, we write the production function in Eq. (1) as a function of i_t only, keeping in mind that the capital needed for implementation must be equal to i_t.

If the manager exerts effort d_t to discover a new project in period t and it is matched by d_t units of capital, a new project will be available for implementation in period $(t+1)$ with some probability p_t given by,

$$p_t = p(d_t), \quad t = 1, 2, \tag{2}$$

where $p(\cdot)$ is an increasing and strictly concave function. For example, the manager could expend effort to discover a new drug, but not always succeed. As in case of implementation, we write the probability function in Eq. (2) as a function of d_t only, keeping in mind that the capital needed for discovery must be equal to d_t.

One of our main premises is that discovery outcomes are less informative about managerial ability than implementation outcomes. For example, discovery may involve identifying a geographic growth strategy through roll-ups. While the identification of the strategy reveals something about the manager's ability, profits from the successful implementation of the roll-up strategy is a stronger signal of her ability. For analytical ease, we model the differential signal-to-noise ratios of discovery and implementation outcomes by assuming that the probability of discovery does not depend on the manager's ability at all.[2]

The manager's utility in period t, u_t, is given by

$$u_t = w_t - \lambda(i_t + d_t), \tag{3}$$

where w_t is her wage in period t. The above specification assumes that the cost of effort to the manager in period t is a linear function of the combined effort for discovery and implementation in that period and that the manager has no preference between the two activities.[3] The parameter $\lambda > 0$ represents the managerial labor to capital ratio: since one unit of capital needs to be matched by one unit of managerial effort for both discovery and implementation, the cost of managerial effort is λ units per unit of capital invested. This specification is equivalent to assuming that every dollar of capital is matched by λ dollars of effort.

2.2. *Information and sequence of events*

The manager's ability is not observable to anyone, including the manager. At the beginning of the first period, both the manager and the market possess identical information about the manager's ability a and it is common knowledge that a is normally distributed with mean m_1 and precision h_1. The effort levels the manager expends for discovery and implementation are not observable to anyone but the manager. To be consistent with the asymmetric information

[2] This simplification enables us to solve the model analytically; an alternative assumption that the probability of discovery depends on the manager's ability but provides the market with the noisier signal about the ability than the profits will not affect the intuition behind our main results.

[3] The assumption of a linear effort cost function is sufficient to satisfy the second-order conditions of the optimization program in our model because of the concavity of the productions functions. The intuition behind the model stays the same if we assume that the cost function is convex, but the linearity assumption significantly simplifies the analysis.

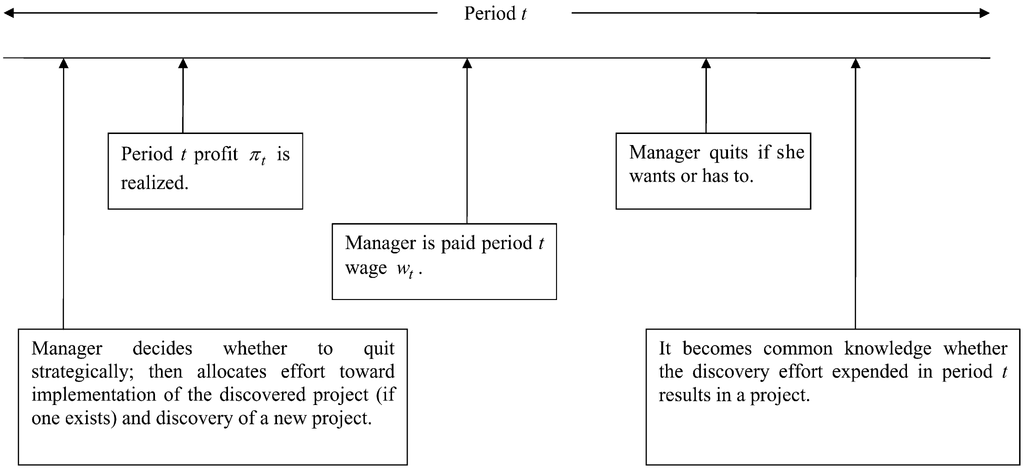

Fig. 1. The sequence of events during period 2. Note that there is a slight change in the sequence of events at $t = 1$ and $t = 3$. For $t = 1$: prior to expending effort, the manager is offered a compensation contract. In this period there is no previously discovered project, no profit, and no compensation paid to the manager. For $t = 3$: there is no future project to be discovered.

assumption about the manager's effort allocation, we need to assume that the capital amounts invested in discovery and implementation are not observable to anyone but the manager.

The sequence of events is as follows (see Fig. 1). At the beginning of the first period the manager is offered an incentive contract that specifies her wage for periods 2 and 3. At the beginning of each period t, the manager first decides if she wishes to quit (see discussion on manager's tenure that follows) based on her opportunities in the labor market and the wage contract offered by the firm. She then expends effort to discover new projects (in periods 1 and 2) and to implement any previously discovered projects (in periods 2 and 3). Then the firm's profit π_t from project implementation is realized (in periods 2 and 3) and is observable to everyone. After this the manager is paid (in periods 2 and 3). Then everyone observes if the effort expended toward discovery at the beginning of the periods 1 and 2 results in an implementable future project. At the end of final period, all residual cash flows are paid out to investors.

2.3. Manager's tenure

We assume that the firm is bound by the wage contract while the manager can quit at the end of any period. In the second and third periods the manager has outside job opportunities that will provide her an expected utility of $E_{I_\tau}(a)$ net of effort costs in each future period, where I_τ is the information available to the market at the beginning of period τ, $\tau = 2, 3$.

A key premise of our model is that the manager values her reputation (i.e., perceived ability). One way to model this premise is to assume that the manager will need to quit the job with some probability $q < \lambda/(1 + \lambda)$ for exogenous reasons.[4] It may be that her position becomes

[4] The limit on q ensures that there is a role for profit-based component in the optimal contract. If the manager values her reputation too much (high q), she will be willing to exert implementation effort to boost her reputation without requiring any monetary compensation, or even be willing to pay to do it.

redundant after a takeover, or that she needs to move to another region for personal reasons.[5] We will refer to quitting for such reasons as *exogenous* quitting. Since the manager's expected utility from an outside job depends on her perceived ability, the possibility of exogenous quitting makes her value her reputation. In addition to exogenous quitting, the manager may quit in any period because she is not satisfied with the contract that she is offered. We will refer to quitting for this reason as *strategic* quitting.[6] The manager makes the strategic quitting decision before effort allocation since her effort decision may depend on the quitting decision. If the manager quits for either reason in period t, a new manager with the same perceived ability will be hired; however, the new project that the former manager discovered will be lost and the firm has no project to implement during the following period.[7] We assume that the reason for quitting is unobservable.

2.4. Incentive contract

When the manager is hired she is offered a long-term incentive contract $W \equiv \{w_2, w_3\}$, where the period t wage w_t depends on the firm's realized current and past profits π_j and cum-profit stock prices Π_j (stock prices just before profit realization) of current and past periods $j \leqslant t$:

$$w_t = w_t\big(\{\pi_j, \Pi_j\}\big|_{j=2}^{t}\big), \quad t = 2, 3. \tag{4}$$

The cum-profit stock prices in periods 2 and 3 are given by,

$$\Pi_t = E_{I_t}(\pi_2 - w_2 + \pi_3 - w_3 - d_1 - d_2 - i_2 - i_3). \tag{5}$$

Eq. (5) states that the period t stock price depends on the expected profit net of wages less the expected capital investments in implementation and discovery, and any retained profits, with the expectation based on the investors' information set.

Since the firm's profit depends on both discovery and implementation efforts, shareholders need to use two different instruments (stock price and profit) to provide the manager with appropriate incentives. Note that although the firm's profit in period t depends on both the discovery effort in period $t-1$ and the implementation effort in period t, the cum-profit stock price in period t depends on the *expected* profit at period t, which, in turn, is affected only by the discovery effort in period $t-1$ and is independent of the actual implementation effort in period t. As a result, shareholders may use these two instruments to provide different incentives for discovery

[5] Alternatively, we may assume that the manager values her reputation because of prestige and self-esteem and directly include the perceived ability into her utility function.

[6] One might wonder whether the possibility of strategic quitting alone would make her value her perceived ability. However, since the optimal contract will ensure that the manager will never want to quit strategically, she obtains no benefit from misleading other employers about her ability. Moreover, in the absence of exogenous quitting it can be shown that the optimal contract will pay her the entire share of her profit which depends only on true ability. As a result, her wage inside the firm will not depend on her perceived ability as well. Therefore, without the possibility of exogenous quitting, the manager does not value her perceived ability. All our results, except those pertaining to the deferred wage, will hold even if only exogenous quitting is allowed.

[7] If, alternatively, we assume that the project that the manager discovered during her final period of employment stays alive but drops in value after she quits, then the main intuition behind the model will not be affected. However, such an assumption will increase the shareholders' effective intertemporal discount factor, making it different from the manager's discount factor; or, alternatively, such an assumption will make the manager's time horizon shorter than the time horizon of the firm. As a result, even if the manager's ability is observable, the optimal incentive contract will include some portion of stock-based compensation.

and implementation. In addition, the ex-profit stock price (stock price just after the profit is realized) depends on the realized profit in period t, and, as such, does not provide any information about the effort level beyond the information provided by the current profit.

We assume that shareholders wish to maximize their expected terminal payouts while the manager wants to maximize the sum of her expected future utility.

3. First-best effort allocation

To find the first-best effort allocation toward implementation and discovery, consider a social planner who wishes to maximize the sum over time of the expected profit less the cost of effort. There are two possible states in each period: success (denoted by s) in which there is a project to implement, or failure (denoted by f) in which there is no project to implement.

At the beginning of each period t, given the state $k \in \{s, f\}$, the social planner chooses the effort levels $i_{k,t}$ and $d_{k,t}$ that maximize the firm's current profit net of the costs of effort and investment,

$$N_t f(i_{k,t}) + E_t(a) - (1+\lambda)(i_{k,t} + d_{k,t}),$$

plus the future expected social surplus, defined as expected profit less expected costs of effort and investment. $E_t(.)$ is the expectation operator based on the manager's (hence, the social planner's) information set at the beginning of period t before she chooses her effort. Denote $S^*_{s,t}$ and $S^*_{f,t}$ to be the sum of future expected social surplus (from period t to period 3) in states s and f, respectively, given first-best effort levels $i^*_{k,t}$, and $d^*_{k,t}$. Thus, the optimization problem for period t and state k can be written as

$$\max_{i_{k,t}, d_{k,t}} N_t f(i_{k,t}) + E_t(a) - (1+\lambda)(i_{k,t} + d_{k,t}) + \big\{(1-q)p_{k,t} S^*_{s,t+1} + \big(1-(1-q)p_{k,t}\big) S^*_{f,t+1}\big\}. \tag{6}$$

The last term in Eq. (6) is the expected social surplus. The manager will be with the firm in period $(t+1)$ and have a project to implement with probability $(1-q)p_{k,t}$, in which case the social surplus will be $S^*_{s,t+1}$. With probability $1-(1-q)p_{k,t}$ there will be no project to implement in period $(t+1)$, in which case the social surplus will be $S^*_{f,t+1}$. The following lemma characterizes the first-best solution. All proofs are provided in Appendix A.

Lemma 1. *The first-best implementation and discovery efforts satisfy the following set of equations*:

$$i_{f,t} = i^*_f = 0 \quad \textit{for } t = 2, 3; \tag{7a}$$

$i_{s,t} = i^*_s$, $t = 2, 3$, *and is given by,*

$$f'\big(i^*_s\big) = 1 + \lambda; \tag{7b}$$

and, $d_{k,t} = d^*$, $k \in \{s, f\}$, $t = 1, 2$, *and is given by,*

$$(1-q)p'(d^*)\big(f\big(i^*_s\big) - (1+\lambda)i^*_s\big) = 1 + \lambda. \tag{7c}$$

The following points are worth noting: (1) Since there is no project to implement in state f, the optimal implementation effort is $i^*_f = 0$ in any period t. (2) Eq. (7b) states that the marginal profit from implementation of a discovered project must equal the total marginal costs of implementation, which is the sum of one unit of marginal cost of capital and λ units of marginal cost of the manager's effort. The optimal implementation effort when there is a project to implement is the same in every period since it does not affect future social surplus. (3) Eq. (7c) states that the marginal benefit of discovery must equal the total marginal cost of discovery. The marginal benefit of discovery is equal to the marginal probability that the project will be available for implementation next period, $(1-q)p'(d^*)$, times the net profit of implementation, which in turn is equal to the expected profit $f(i_s^*)$ minus total costs of implementation $(1+\lambda)i_s^*$. The marginal cost of discovery is one unit of capital plus λ units of effort. The discovery effort and the probability of discovery are independent of time and state in our model because of the assumption that the marginal costs of discovery and implementation efforts do not depend on the level of implementation and discovery effort, respectively. As a result, the optimal choice of the discovery effort is not affected by the existence of a project to implement in the current period.

To simplify the notation, we will drop the subscript representing state when referring to the implementation effort since it is expended only if there is a project to implement.

4. Optimal incentive contract

When the manager's effort allocation is not observable, shareholders have to rely on an incentive contract to extract the optimal levels of discovery and implementation efforts. We will seek a contract, if one exists, that (i) induces the manager to expend first-best levels of discovery and implementation efforts (Incentive Compatibility (IC) constraint); (ii) at any period t provides the manager a sum of current and future expected utility (from period t to period 3) that is greater than or equal to the sum of her current and future expected utility from outside job opportunities (Individual Rationality (IR) constraint); and (iii) at $t=1$ provides the manager a sum of current and future expected utility (from period t to period 3) that is exactly equal to the sum of her current and future expected utility from an outside job opportunity. If such a contract exists, it will (a) satisfy IC and IR conditions, (b) deliver the highest social surplus, and (c) provide the manager an expected utility exactly equal to her reservation utility. Because both the manager and shareholders are risk neutral, such a contract will provide the shareholders the highest possible profit among all contracts that satisfy IC and IR constraints, i.e., it will be the optimal incentive contract.

We consider incentive contracts of the form

$$w_t = \alpha_t \pi_t + \beta_t \Pi_t + c_t, \quad t = 2, 3. \tag{8}$$

Note that if there exists an incentive contract of the form in Eq. (8) that satisfies all three conditions specified above, then it is an optimal incentive contract among all possible contracts specified by Eq. (4). Substituting the definition of stock price in Eq. (5) into Eq. (8), and solving the resulting expression for $E_{I_t}(w_t)$ yields,

$$E_{I_t}(w_t) = \frac{\alpha_t + \beta_t}{1+\beta_t} E_{I_t}(\pi_t) + \frac{\beta_t}{1+\beta_t} E_{I_t}\big(\pi_j - w_j - (i_2 + i_3 + d_1 + d_2)\big) + \frac{c_t}{1+\beta_t}, \tag{9}$$

where $j=2$ for $t=3$ and $j=3$ for $t=2$. Substituting Eq. (9) into the expression for Π_t in Eq. (5) and then using the result in Eq. (8) yields the following expression for w_t:

$$w_t = \alpha_t \pi_t + \frac{\beta_t}{1+\beta_t}\big[(1-\alpha_t)E_{I_t}(\pi_t) + E_{I_t}\big(\pi_j - w_j - (i_2 + i_3 + d_1 + d_2)\big)\big] + \frac{c_t}{1+\beta_t}, \tag{10}$$

where $j = 2$ for $t = 3$ and $j = 3$ for $t = 2$.

The first term in Eq. (10) represents the profit-based wage component. It indicates that the manager receives α_t portion of the firm's current profit. The second term represents the stock-based component and consists of two parts. This term is better understood by noting that the period 2 stock price depends on the both current and period 3 expected profits, whereas the period 3 stock price depends on both the current expected profit and the period 2 realized profit since the residual cash flows are retained in the firm. The first part of the second term in Eq. (10) represents the stock-based portion of the wage derived from the current period expected profit. Since the manager receives α_t portion of the firm's current profit, only $(1-\alpha_t)$ portion of the profit is included into the stock price. The second part of the stock-based component represents the effect of the past or future profit on the current stock price, and, consequently, on the manager's current stock-based compensation. The second term in Eq. (10) is scaled by $1/(1+\beta_t)$ to account for the fact that the stock price is net of wages, decreasing the sensitivity of the stock-based compensation to the expected profit. The last term in Eq. (10) is a fixed cash component. Since one unit of cash compensation reduces the stock price by the same amount and, consequently, reduces the managers's stock-based compensation by β_t, the cash compensation is also scaled by $1/(1+\beta_t)$.

The manager's optimization problems in periods 1, 2, and 3, respectively, are given by the following set of equations:

$$\max_{d_1} E_1 \left\{ \begin{array}{c} -\lambda d_1 + 2qm_2 + (1-q)\big(w_2 - p(d_1)\lambda i^* - d_2\big) \\ + q(1-q)m_3 + (1-q)^2\big(w_3 - p(d_2)\lambda i^*\big) \end{array} \right\}, \tag{11a}$$

$$\max_{i_2, d_2} E_2\big\{w_2 - \lambda i_2 - \lambda d_2 + qm_3 + (1-q)\big[w_3 - p(d_2)\lambda i^*\big]\big\}, \tag{11b}$$

$$\max_{i_3} E_3\{w_3 - \lambda i_3\}, \tag{11c}$$

where m_t is the market's perception of her ability at the beginning of period t. In period 3 the manager trades off her wage against her implementation effort (Eq. (11c)). In period 2 her utility function (Eq. (11b)) includes, in addition to the trade off between her wage and her efforts in period 2, her expected utility in period 3. If she quits in period 2 her utility in period 3 is m_3 and if she stays she receives w_3 but will exert implementation effort i^* with probability $p(d_2)$. Similarly, in period 1 she has to consider her expected utility in periods 2 and 3. In Eq. (11a), the second and fourth terms represent the manager's utility if she quits in periods 1 and 2 respectively, and the third and last terms represent her utility if she stays with firm in the same periods.

Since the manager's share of the factors of production is $\lambda/(1+\lambda)$, her marginal benefit from discovery and implementation efforts should be set to $\lambda/(1+\lambda)$ times the marginal profit produced by the respective efforts for her to receive full credit for her efforts.

As the firm's expected profit at the beginning of a period depends only on the availability of the project and the expected implementation effort level and not on the actual implementation effort level, the stock-based compensation does not affect the manager's choice of implementation effort. Therefore, shareholders can extract the optimal implementation effort in period t by choosing an appropriate profit-based component α_t. They can then select the stock-based component β_t such that manager receives full credit for her discovery effort.

Implementation effort increases the expected current period profit, creating two sources of benefit for the manager: increased expected compensation from current profit and increased perceived ability (if she cares about her reputation).[8] In the terminal period the manager does not care about her reputation. Therefore, her marginal benefit from implementation effort comes only from the profit-based compensation in that period and, in state s of period 3, is equal to $\alpha_3 f'(i_3)$. Thus, it must be the case that $\alpha_3 = \lambda/(1+\lambda)$. In period 2, however, the manager's marginal benefit of implementation may also come from her increased perceived ability. Therefore, it is possible that $\alpha_2 < \lambda/(1+\lambda)$.

Discovery effort increases the probability that a project will be available for implementation next period, creating two sources of benefit for the manager: increased compensation from both higher expected profit and higher stock price next period.[9] Since setting $\alpha_3 = \lambda/(1+\lambda)$ provides the manager full credit for the terminal period implementation effort, the increase in the expected profit-based compensation in the terminal period is sufficient to induce her to exert the optimal discovery effort in period 2. Therefore, $\beta_3 = 0$. If, however, $\alpha_2 < \lambda/(1+\lambda)$, additional incentives may be needed to extract the optimal discovery effort in period 1 and, therefore, β_2 is likely to be positive. It can be seen from Eq. (10) that the manager gets α_2 share of the firm's realized profit and $(1-\alpha_2)\beta_2/(1+\beta_2)$ share of the firm's current expected profit. Therefore, in order to give the manager full credit for her discovery effort, the incentive contract must satisfy the condition $\alpha_2 + (1-\alpha_2)\beta_2/(1+\beta_2) = \lambda/(1+\lambda)$. The following lemma summarizes the above discussion.

Lemma 2. *The optimal contract must satisfy the following conditions*: $\beta_3 = 0$ *and,*

$$\alpha_t + (1-\alpha_t)\frac{\beta_t}{1+\beta_t} = \frac{\lambda}{1+\lambda}, \quad t = 2, 3. \tag{12}$$

Given the result in Lemma 2, $\alpha_3 = \lambda/(1+\lambda)$ and, therefore, we only need to derive α_2 and β_2 that satisfy Eq. (12). In order to highlight the impact of the manager's reputation concern on the optimal wage contract, we will derive α_2 and β_2 for the case of known ability first.

4.1. Case 1: Known ability

In this case the manager does not care about her reputation in any period as she cannot influence the market's perception of her ability. Therefore, her marginal benefit from implementation effort in period 2 comes only from her profit-based compensation in that period. Therefore, it must be the case that $\alpha_2 = \lambda/(1+\lambda)$ to give her full benefit. The following theorem formalizes the above discussion.

Theorem 1. *If the manager's ability is known, a linear contract of the form in Eq.* (8) *is optimal if and only if the parameters satisfy*

$$\left.\begin{aligned} \alpha_t &= \lambda/(1+\lambda) \\ \beta_t &= 0 \end{aligned}\right\}, \quad t = 2, 3, \tag{13}$$

[8] Period 2 implementation effort i_2 can potentially have an impact on period 3 wage w_3 through the stock based component since period 2 cash flows are retained in the firm till the terminal period. However, as we show shortly, $\beta_3 = 0$, and, therefore, this effect is absent.

[9] Period 1 discovery effort d_1 can potentially have an impact on period 3 wage w_3 through the stock based component since period 2 cash flows are retained in the firm till the terminal period. Since $\beta_3 = 0$ (as shown shortly), this effect is absent in this case and therefore we need not consider the second component of the stock-based compensation in Eq. (10).

and c_t, $t = 2, 3$, structured to satisfy the IR constraint. If the manager's ability is high enough, then $c_t \geqslant 0$, $t = 2, 3$.

As one can expect in a model with risk-neutral agents, no cash constraints, and unobservable efforts, the form of the optimal incentive contract is along the line of selling part of the firm to the manager. Indeed, since the production costs are divided between the manager and the shareholders with the manager bearing $\lambda/(1+\lambda)$ portion of the total costs, the incentive contract that provides her with $\lambda/(1+\lambda)$ portion of the profit each period is an optimal incentive contract. As we will show later, a simple sale of a portion of the firm to the manager is not an optimal incentive contract when the manager is concerned about her reputation.

In the out-of-equilibrium move in which the manager decides to quit strategically in periods 1 and 2, she will not exert any effort toward discovery. Saving the cost of discovery effort will lead to a higher utility for her in the current period. To prevent such strategic quitting, the wage compensation should have a deferred wage component, i.e., in equilibrium, the manager's expected utility in period 1 should be less than that in period 3. As the proof of Theorem 1 shows, an appropriately specified cash component that results in a deferred wage structure prevents strategic quitting (i.e., satisfies the IR constraint).

Since the manager contributes only $\lambda/(1+\lambda)$ share of the project's input, in equilibrium she receives $\lambda/(1+\lambda)$ share of the firm's profit, and hence $\lambda/(1+\lambda)$ of the value of her ability, through the incentive part of her contract (stock or profit-based compensation). This means that she must be compensated for the remaining $1/(1+\lambda)$ part of her ability through the fixed wage component c_t. As shown in the proof of Theorem 1, if the manager's ability is high enough, the fixed components of the manager's compensation, c_t, will be positive to compensate her fully. Positive values for c_t will accommodate a manager with zero wealth by ensuring that he does not have to pay to accept the job.

4.2. Case 2: Unknown ability

We now consider the case in which the manager's ability is not observable. In equilibrium, the market updates its beliefs about the manager's ability based on the observed profit π_t. To simplify the notation, we define

$$y_t = \pi_t - f(i^*), \quad t = 2, 3. \tag{14}$$

The information provided by y_t is a sufficient statistic for π_t. Based on this information, the market updates its beliefs about the manager's ability according to the following rule:

$$a|I_{t+1} \sim N(m_{t+1}, 1/h_{t+1})$$

where m_{t+1} and h_{t+1} are given by,

$$m_{t+1} = \frac{h_t m_t + h_\varepsilon y_t}{h_t + h_\varepsilon}, \quad h_{t+1} = h_t + h_\varepsilon.$$

Since there is no output in period 1, $m_2 = m_1$ and $h_2 = h_1$. The perceived ability at the beginning of period 3 is given by,

$$m_3 = \frac{h_2 m_2 + h_\varepsilon y_2}{h_2 + h_\varepsilon}. \tag{15}$$

By deviating from the first-best equilibrium effort levels in period 2 the manager can influence the information that the market uses to form its beliefs about her expected ability. Indeed, as Eqs. (14) and (15) imply, i_2 affects not only the current profit w_2, but also m_3, the manager's perceived ability next period. Since the implementation outcome provides more information about the manager's ability than the discovery outcome, the manager may want to exert more than the first-best effort toward implementation of the existing project, if one exists. By doing so the manager will increase the firm's current profit in state s, which, given the equilibrium beliefs, will lead the market to believe that the manager's ability is higher. As a result, a reputation-concerned manager benefits from implementation effort in two ways: from an increase in current wage and from improved reputation. Hence, a wage contract that incents a reputation-concerned manager to extract the first-best implementation effort should be less sensitive to the firm's profit, i.e., should be designed such that $\alpha_2 < \lambda/(1+\lambda)$. This adjustment by itself, however, is not sufficient. A reduction in α_2 also reduces the manager's expected benefit from period 1 discovery which, in turn, reduces her incentive to exert discovery effort d_1. Therefore, shareholders need to include a stock-based component in the compensation contract (i.e., set $\beta_2 > 0$), to induce first-best discovery effort in the presence of the reputation concerns.

Similar to the known-ability case, shareholders need to structure the cash component of the manager's compensation in such a way that manager will never want to quit strategically. The following theorem summarizes the discussion above.

Theorem 2. *If the manager's ability is not known, a linear contract of the form in Eq. (8) is optimal if and only if the parameters satisfy*

$$\left.\begin{aligned} \alpha_2 &= \frac{\lambda}{1+\lambda} - q\frac{h_\varepsilon}{h_2+h_\varepsilon}, \\ \beta_2 &= q(1+\lambda)\frac{h_\varepsilon}{h_2+h_\varepsilon}, \\ \alpha_3 &= \frac{\lambda}{1+\lambda}, \\ \beta_3 &= 0, \end{aligned}\right\} \tag{16}$$

and c_t, $t = 2, 3$ structured to satisfy the IR constraint. If the manager's perceived ability at $t = 1$ is high enough, then $c_t \geqslant 0$, $t = 2, 3$.

Similar to the known-ability case, the cash component is structured to include some deferred compensation. However, now a larger portion of the wage needs to be deferred compared to the known-ability case. This is because, in addition to saving discovery effort, now the manager who plans to quit strategically can also potentially receive additional expected utility from outside job opportunities by boosting her perceived ability through increased implementation effort. For the same reasons as in the case of known ability, the fixed components of the manager's compensation, c_t, will be positive if her perceived ability is high enough. As before, positive values of c_t will accommodate a manager with zero wealth by ensuring that he does not have to pay to accept the job; positive values of α_t (note that $q < \lambda/(1+\lambda)$) guarantee that the manager's wage will not decrease with the firm's profit.

Theorem 2 shows that if the manager's ability is not known the sale of a fraction of the firm to the manager will not be an optimal incentive contract. Selling a part of the firm to the manager is equivalent to a contract with $\alpha_2 > 0$ and $\beta_2 = 0$, but the theorem shows that β_2 is strictly positive. The intuition behind this result is as follows. Since the manager's reservation utility is based on the market's perception of her ability, even if the manager owns all the firm (i.e., supplies all the effort and provides all the capital), she has the incentive to exert more than the socially optimal level of implementation effort in order to boost the value of her outside opportunities.

As Eq. (16) shows, if the reputation concern is not too high ($q < \lambda/(1+\lambda)$), a wage contract based only on stock-based compensation (i.e., $\alpha_2 = 0$) is not optimal. Intuitively, if compensation is exclusively stock-based, the manager's wage becomes independent of realized profits. Her incentive to implement is then driven entirely by reputation concerns, which is not sufficient to extract the optimal implementation effort if q is low enough. It also follows from Eq. (16) that, in the absence of exogenous quitting (i.e., $q = 0$) there is no need for a stock-based component.

5. Empirical implications

Our model provides several empirical implications that relate firm, industry and managerial characteristics to the design of the manager's optimal incentive contract. The trade-off between stock-based and profit-based components is driven by the extent to which the manager values her reputation and the ease with which she can attempt to boost it. The extent to which the manager values her reputation in turn is influenced by the parameter q, the probability of exogenous quitting. The greater this probability the more the manager cares about her reputation and, thus, the more effort she puts into implementation; to offset this incentive shareholders will have to provide more stock-based compensation. The following result is obtained by taking the appropriate partial derivatives of Eq. (16).[10]

Result 1. The greater the manager's concern about her reputation the greater (lower) the stock-based (profit-based) component in the optimal incentive contract, i.e.,

$$\frac{\partial \alpha_2}{\partial q} < 0 \quad \text{and} \quad \frac{\partial \beta_2}{\partial q} > 0.$$

Reputation concerns are likely to be greater when the demand for managers is high which in turn is more likely in growth industries and in industries with emerging technologies. Evidence in Core and Guay (1999, 2001) shows that industry growth is positively related to increased stock-based compensation. Ittner et al. (2003) provide evidence that new economy firms (firms in computer, software, internet, or telecommunications fields) offer more stock-based compensation relative to profit-based compensation. In addition, managers with transferable skills are more likely to be concerned about their reputation. Therefore, such managers should be offered contracts with a greater proportion of stock-based component. Such contracts should be more prevalent in economic booms when the demand for managers will be high.

The ease with which the manager can attempt to boost her reputation is related to two parameters in our model: the precision of the manager's perceived ability at the time of hiring (h_1) and the precision of the firm's profit (h_ε). If the manager's ability is known to the market with greater precision, less of the unanticipated increase in profit will be attributed to her ability. As a result, it becomes harder for the manager to manipulate market perceptions, lowering her incentive to boost her perceived ability by allocating more effort into implementation. Therefore, a manager whose ability is known with greater precision should be offered a contract with higher

[10] All comparative statics in the results that follow are for period 2. Beginning and ending periods are ignored.

profit-based compensation and lower stock-based compensation. Since the precision of the manager's ability in any period is directly related to the precision of her ability at the time of hiring, stock-based compensation is inversely related to the precision of the manager's perceived ability at the time of hiring.

The greater the precision of the firm's profit the greater its informativeness as a signal of the manager's ability, which, in turn, makes it easier for her to mislead the market and increases her incentive to allocate more effort toward implementation. To mitigate this incentive to misallocate effort, the firm should decrease (increase) the profit-based (stock-based) compensation. The following result, which can be obtained by taking the appropriate partial derivatives of Eq. (16), summarizes the discussion above.

Result 2. Stock-based compensation is inversely related (and profit-based compensation is directly related) to the precision of the manager's perceived ability at the time of hiring. Stock-based compensation is directly related (and profit-based compensation is inversely related) to the precision of the employee's output. That is,

$$\frac{\partial \alpha_2}{\partial h_1} > 0, \qquad \frac{\partial \beta_2}{\partial h_1} < 0, \qquad \frac{\partial \alpha_2}{\partial h_\varepsilon} < 0 \quad \text{and} \quad \frac{\partial \beta_2}{\partial h_\varepsilon} > 0.$$

Consistent with the result that stock-based compensation is inversely related to the precision of the manager's ability, Ittner et al. (2003) find that newly hired managers are paid a higher stock-based component than managers at the same level already with the firm. The fact that lower level employees are paid less stock-based compensation (Core and Guay, 2001) is consistent with the result that stock-based compensation is inversely related to the noise of the employee output.

The parameter λ in our model represents the cost of effort per unit of capital invested in the project either for discovery or implementation. It can, therefore, be interpreted as the intensity of managerial labor to capital. We expect λ to be greater in industries that are less capital intensive and more knowledge-based. The following result shows that both profit-based and stock-based compensations should be greater in knowledge-based industries.

Result 3. Both stock-based and profit-based compensations are greater in knowledge-based industries, i.e.,

$$\frac{\partial \alpha_2}{\partial \lambda} > 0 \quad \text{and} \quad \frac{\partial \beta_2}{\partial \lambda} > 0.$$

The intuition for the above result is as follows. A higher λ implies that the cost of effort is greater relative to the cost of investment requiring that the manager receive a greater proportion of the profit, i.e., α_2 increases. As the manager's proportion of the profit increases, however, the stock price (which is net of wage) becomes less sensitive to discovery effort. Assuming reputation concerns stay the same (i.e., q does not change), β_2 also needs to increase to maintain the optimal level of discovery effort. The empirical evidence is consistent with this result. Ittner et al. (2003) finds that new economy firms with high R&D-to-sales ratio than old economy firms (durable and non-durable manufacturers) have significantly higher proportion of CEO pay linked to the stock price. Anderson et al. (2000) finds that firms in information technology industry offer larger stock-based compensation than traditional firms. Talmor and Wallace (1998) finds a similar result for CEO compensation in the computer industry.

We can easily relax the assumption that the manager's ability is constant by assuming, as in Holmstrom (1999), that ability follows a martingale process:

$$a_{t+1} = a_t + \rho_t,$$

where ρ_t is distributed normally with zero mean and precision h_ρ. It can be shown that the expression for profit and stock-based components will not change and will be determined by Eq. (16).[11] However, higher intertemporal volatility of the manager's ability will reduce the precision of the manager's ability in any period. In particular, the precision of the managers ability in period 2 will be given by $h_2 = h_1 h_\rho/(h_1 + h_\rho)$, i.e., it will be negatively related to the intertemporal volatility of the manager's ability. Combining the above analysis with Result 2, we obtain the following result.

Result 4. Profit based compensation is directly related (and stock-based compensation is inversely related) to the precision with which the manager's ability varies over time, i.e.,

$$\frac{\partial \alpha_2}{\partial h_\rho} > 0, \qquad \frac{\partial \beta_2}{\partial h_\rho} < 0.$$

This result may provide an explanation for the significant increase in the proportion of stock-based compensation since 1980s (Murphy, 1999). The fast growth of the technology-based industries such as computer hardware, communications, software, and biotechnology is likely to require managers to learn new skills and makes their accumulated knowledge, and hence ability, less useful. In the framework of our model, this development corresponds to greater intertemporal volatility of managerial ability and, thus, increased stock-based compensation.

In our model, the discovery outcome is assumed to be independent of managerial ability. In a more general model, if the outcomes of both implementation and discovery provide informative signals about managerial ability, the stock-based component of the compensation will be negatively related to the ratio of profit volatility to discovery volatility. Thus, ceteris paribus, we should expect to observe more stock-based compensation in industries in which it is more difficult to ascertain the probability of success relative to the volatility of profits once a discovery has been made. Examples of industries that are likely to meet these requirements are pharmaceuticals and biotechnology, in which the outcome of research and development is very uncertain but the cash flows from successful products are more certain because of patent protection.

6. Conclusion

This paper offers an explanation for the use of both profit-based and stock-based incentive compensation even though the stock price is related to expected future profits. We consider a three-period model in which the manager is offered a long-term compensation contract. The manager needs to allocate her effort to two activities: strategy formulation and strategy implementation. The firm's profit depends on investor capital, managerial ability, and implementation effort. Managerial ability and the effort allocated toward strategy formulation might result in discovery of projects for implementation.

The two premises of the model are that the outcome of implementation is more informative about managerial ability than the outcome of strategy formulation, and that the manager cares

[11] We can supply the proof of this assertion to interested readers.

about her reputation. We show that when the market can perfectly observe the manager's ability, the manager's wage in each period depends only on the firm's profit in that period and not on the firm's stock price. However, when the manager's ability is not observable, she will attempt to boost the market perception of her ability by shifting her effort toward the more informative implementation of existing projects. To restore the first-best outcome, shareholders have to offer some stock-based compensation for a reputation-concerned manager.

Besides offering an explanation for both profit-based and stock-based compensation, the paper also provides several results that explain the variation in compensation structures across industries and over time. The paper suggests that compensation contracts in growth industries that create value through innovation, such as communication, biotechnology, hardware, and software, are more likely to have a greater stock-based component than mature industries. There are several reasons for this. Managers in such industries are likely to have greater reputation concerns since there is greater demand for managers in these growth industries. Managers in these industries may be able to manipulate their reputation more easily because of the low volatility of profits relative to volatility of discovery outcome in these industries. In addition, managers in these fast-changing industries may have to adapt frequently and learn new skills resulting in greater volatility in intertemporal changes in their ability, which in turn makes it easier for them to manipulate their reputation. Finally, these industries are more knowledge-based than capital intensive, resulting in increased profit-based compensation which, in turn, lowers the sensitivity of the stock price to discovery effort. To compensate, the stock-based compensation needs to be higher. By the same token, the emergence of such industries as hardware, software, and biotechnology in which there is constant change requiring adaptability and learning new skills, might have increased the need for more stock-based compensation as observed in the last twenty years or so.

Acknowledgments

We thank Sugato Bhattacharyya, Gopalan Radhakrishnan, Uday Rajan, and the participants of the finance seminar at the University of Michigan for their comments. We also thank two anonymous referees and an editor of the journal for extensive comments that have benefited the paper.

Appendix A

Proof of Lemma 1. First, note that if there is no project to implement (state f), the optimal implementation effort i_f^* is zero for all t. At $t=3$ the optimization problem in state s for the social planner can be written as,

$$\max_{i_{s,3}} f(i_{s,3}) + E_3(a) - (1+\lambda)i_{s,3},$$

and the optimal implementation effort $i_{s,3}^*$ is given by the first-order condition,

$$f'\big(i_{s,3}^*\big) = 1+\lambda.$$

At $t=2$ the optimization problem in state $k=\{s,f\}$ for the social planner can be written as,

$$\max_{i_{k,2},d_{k,2}} N_2 f(i_{k,2}) + E_2(a) - (1+\lambda)(i_{k,2}+d_{k,2}) \\ + \big\{(1-q)p(d_{k,2})\big(f\big(i_{s,3}^*\big) + E_2(a) - (1+\lambda)i_{s,3}^*\big) + \big(1-(1-q)p(d_{k,2})\big)E_2(a)\big\}.$$

The optimal implementation effort $i^*_{s,2}$ is given by the first-order condition,

$$f'\big(i^*_{s,2}\big) = 1 + \lambda.$$

The optimal discovery effort $d^*_{k,2}$ in state $k \in \{s, f\}$ is given by the first-order condition,

$$(1-q)p'\big(d^*_{k,2}\big)\big(f\big(i^*_{s,3}\big) - (1+\lambda)i^*_{s,3}\big) = 1 + \lambda.$$

At $t = 1$ the optimization problem for the social planner can be written as,

$$\begin{aligned}\max_{d_1} & -(1+\lambda)d_1 + \big\{(1-q)p(d_{k,1})\big[f\big(i^*_{s,2}\big) + E_1(a) - (1+\lambda)i^*_{s,2}\big] \\ & + \big(1-(1-q)p(d_{k,1})\big)E_1(a)\big\} \\ & + \big\{(1-q)p\big(d^*_2\big)\big[f\big(i^*_{s,3}\big) + E_1(a) - (1+\lambda)i^*_{s,3}\big] + \big(1-(1-q)p\big(d^*_2\big)\big)E_1(a)\big\}.\end{aligned}$$

The optimal discovery effort d^*_1 is given by the first-order condition,

$$(1-q)p'\big(d^*_1\big)\big(f\big(i^*_{s,2}\big) - (1+\lambda)i^*_{s,2}\big) = 1 + \lambda. \qquad \square$$

Proof of Lemma 2. We prove a slightly stronger result, namely, that the solution to the manager's optimization problem in Eqs. (11a)–(11c) will result in effort allocation levels $d_1 = d^*$, $d_2 = d^*$, and $i_3 = i^*$ if and only if $\beta_3 = 0$ and Eq. (12) is satisfied.

At $t = 3$, the manager's optimization problem is given by Eq. (11c). Using the definition of the wage contract in Eq. (8) and noting that Π_3 is unaffected by the manager's choice of i_3, the first-order condition in state s is $\alpha_3 f'(i_3) = \lambda$. Therefore, Eq. (7b) implies that $i_3 = i^*$ if and only if $\alpha_3 = \lambda/(1+\lambda)$.

At $t = 2$, the manager's optimization problem is given by Eq. (11b), where $m_3 = a$ if the manager's ability is observable. From Eq. (10) it follows that,

$$E_2 w_3 = \alpha_3\big[p(d_2)f(i^*) + E_2(a)\big] + (1-\alpha_3)\frac{\beta_3}{1+\beta_3}\big[p(d_2)f(i^*) + E_2(m_3)\big] + K_3, \quad \text{(A.1)}$$

where K_3 is given by,

$$K_3 \equiv \frac{\beta_3}{1+\beta_3}E_2\big\{\pi_2 - w_2 - E_{I_3}(d_1 + d_2 + i_2 + i_3)\big\} + \frac{c_3}{1+\beta_3}. \quad \text{(A.2)}$$

The marginal benefit from the manager's choice of d_2 arises from its effect on w_3, not w_2. Also, while K_3 depends on investors' expectation of d_2, it is independent of the manager's choice of d_2. In addition, d_2 does not affect neither $E_2(a)$ nor $E_2(m_3)$. Therefore, using Eqs. (11b) and (A.1), the first-order condition with respect to d_2 is given by,

$$(1-q)p'(d_2)\left[\left(\alpha_3 + (1-\alpha_3)\frac{\beta_3}{1+\beta_3}\right)f(i^*) - \lambda i^*\right] = \lambda. \quad \text{(A.3)}$$

Since $\alpha_3 = \lambda/(1+\lambda)$, it can be seen from Eqs. (A.3) and (7c) that $d_2 = d^*$ if and only if $\beta_3 = 0$. It follows that Eq. (12) is satisfied for $t = 3$ if and only if $i_3 = i^*$ and $d_2 = d^*$.

At $t = 1$ the manager's optimization problem is given by Eq. (11a). Similar to Eq. (A.1) we can write,

$$E_1 w_2 = \alpha_2\big(p(d_1)f(i^*) + E_1(a)\big) + (1-\alpha_2)\frac{\beta_2}{1+\beta_2}\big[p(d_1)f(i^*) + E_1(m_2)\big] + K_2, \quad \text{(A.4)}$$

where K_2 is given by,

$$K_2 \equiv \frac{\beta_2}{1+\beta_2} E_1\{E_{I_2}(\pi_3 - w_3 - d_1 - d_2 - i_2 - i_3)\} + \frac{c_2}{1+\beta_2}. \tag{A.5}$$

Although in general the manager's choice of d_1 can have an effect on w_3 through Π_3, this effect is absent here because $\beta_3 = 0$. Using the same logic as before, K_2 is independent of the manager's choice of d_1. In addition, d_1 does not affect neither $E_1(a)$ nor $E_1(m_2)$. Therefore, using Eqs. (A.4) and (A.5), the manager's first-order condition with respect to d_1 is given by,

$$(1-q)p'(d_1)\left[\left(\alpha_2 + (1-\alpha_2)\frac{\beta_2}{1+\beta_2}\right) f(i^*) - \lambda i^*\right] = \lambda. \tag{A.6}$$

From Eqs. (A.6) and (7c) it can be seen that $d_1 = d^*$ if and only if Eq. (12) is satisfied. □

Proof of Theorem 1. First, we check that the IC constraint is satisfied if and only if α_t and β_t are given by Eq. (13). As the proof of Lemma 2 shows, $d_1 = d^*$, $d_2 = d^*$, and $i_3 = i^*$ if and only if $\beta_3 = 0$ and Eq. (12) is satisfied. It follows from Eqs. (11b), (A.1), and (A.2) that the first-order condition with respect to i_2 in state s is,

$$\alpha_2 f'(i_2) = \lambda.$$

Combining this condition with Eq. (7b) implies that $i_2 = i^*$ if and only if $\alpha_2 = \lambda/(1+\lambda)$. Using Eq. (12) it follows that $\beta_2 = 0$.

To satisfy the IR constraint, that is, to guarantee that the manager accepts the contract and does not quit strategically in period 2, shareholders need to design the cash compensation to have a deferred component. If the manager's ability is known, her choice of implementation effort is independent of the quitting decision and is determined by the tradeoff between the marginal cost of effort and the marginal profit-based compensation. Hence the maximum utility she can gain from quitting strategically is λd^*, the cost of equilibrium discovery effort. Therefore, to prevent strategic quitting, an amount λd^* of the manager's compensation in each period should be deferred to the next period.[12] Since the deferred compensation will be paid only if the manager stays with the firm, which happens with probability $1-q$, the manager should be paid $\lambda d^*/(1-q)$ for the discovery effort of d^* exerted in the previous period.

The compensation in period 2 should provide the manager with an expected utility in period 2 equal to the sum of (1) the value of her expected ability a; (2) the expected costs of her effort $\lambda(p(d^*)i^* + d^*)$; and (3) the wage deferred from period 1, $\lambda d^*/(1-q)$, less the wage deferred to period 3, λd^*. That is, we must have,

$$E_{I_1} w_2 = a + \lambda\big(p(d^*)i^* + d^*\big) + \frac{\lambda d^*}{1-q} - \lambda d^* = a + \lambda p(d^*)i^* + \frac{\lambda d^*}{1-q}.$$

Using Eq. (10), the manager's expected wage in period 2 is,

$$E_{I_1} w_2 = \frac{\lambda}{1+\lambda}\pi_2 + c_2 = \frac{\lambda}{1+\lambda} p(d^*) f(i^*) + \frac{\lambda}{1+\lambda} a + c_2.$$

We can obtain exactly the same two expressions as above for period 3 as well (the lower expected cost of discovery effort is offset by the absence of deferred wage). Therefore, equating the above

[12] Strictly speaking, this deferral amount is the lower bound to prevent strategic quitting. If we defer too much, however, the cash component in period 2 will become negative.

two equations we find that c_t is given by,

$$c_t = -\left[p(d^*)\left(\frac{\lambda}{1+\lambda}f(i^*) - \lambda i^*\right) - \lambda\frac{d^*}{1-q}\right] + \frac{1}{1+\lambda}a, \quad t = 2, 3. \tag{A.7}$$

It can be seen from Eq. (A.7) that c_t is positive if a is high enough. □

Proof of Theorem 2. First, we check that the IC constraint is satisfied if and only if α_t and β_t are given by Eq. (16). As the proof of Lemma 2 shows, $d_1 = d^*$, $d_2 = d^*$, and $i_3 = i^*$ if and only if $\beta_3 = 0$ and Eq. (12) is satisfied. Since $\beta_3 = 0$ and i_2 does not affect future realized profits, i_2 has no effect on $E_2 w_3$. Therefore, using Eq. (11b), the first-order condition of the manager's period 2 optimization problem with respect to i_2 in state s is given by,

$$\alpha_2 f'(i^*) + q\frac{\partial m_3}{\partial \pi_2}\frac{\partial \pi_2}{\partial i_2} = \lambda.$$

In equilibrium,

$$\frac{\partial \pi_2}{\partial i_2} = f'(i^*) = 1 + \lambda. \tag{A.8}$$

It follows from Eq. (15) that

$$\frac{\partial m_3}{\partial \pi_2} = \frac{h_\varepsilon}{h_2 + h_\varepsilon}. \tag{A.9}$$

From Eqs. (A.8) and (A.9) it follows that $i_2 = i^*$ if and only if

$$\alpha_2 = \frac{\lambda}{1+\lambda} - q\frac{h_\varepsilon}{h_2 + h_\varepsilon}. \tag{A.10}$$

From Eqs. (12) and (A.10) it follows that

$$\beta_2 = q(1+\lambda)\frac{h_\varepsilon}{h_2 + h_\varepsilon}.$$

The cash components c_t need to be chosen to satisfy the IR constraints. Since the manager's ability is not known, deciding to quit strategically has two potential benefits in period 2 in addition to saving the cost of discovery effort: By expending more than the equilibrium implementation effort the manager can try to boost her perceived ability and period 2 wage. Her additional benefit from expending effort i_2 instead of i^* is given by,

$$B \equiv \max_{i_2} E_2\big[\alpha_2\big(f(i_2) - f(i^*)\big) + (m_3 - m_1) - \lambda(i_2 - i^*)\big].$$

The first term is the increase in period 2 wage from profit component and the second term is the increase in reservation wage from increased perceived ability (note that $m_2 = m_1$). The final term is the cost of additional implementation effort. Hence, to prevent strategic quitting shareholders should defer $(\lambda d^* + B)$ of the manager's compensation each period.

The compensation in period 2 should provide the manager with the expected utility in period 2 which is equal to the sum of (1) her value of her expected ability m_1; (2) the expected costs of her effort $\lambda(p(d^*)i^* + d^*)$; and (3) the wage deferred from period 1, which is equal to $\lambda d^*/(1-q)$, less the wage deferred to period 3, $\lambda d^* + B$. That is, we must have,

$$E_{I_1} w_2 = m_1 + \lambda\big(p(d^*)i^* + d^*\big) + \frac{\lambda d^*}{1-q} - \lambda d^* - B = m_1 + \lambda p(d^*)i^* + \frac{\lambda d^*}{1-q} - B. \tag{A.11}$$

From Eq. (10) the manager's expected wage in period 2 is given by,

$$\begin{aligned} E_{I_1} w_2 &= \left(\alpha_2 + \frac{(1-\alpha_2)\beta_2}{1+\beta_2}\right) E_{I_1}(\pi_2) + A + \frac{c_2}{1+\beta_2} \\ &= \frac{\lambda}{1+\lambda} p(d^*) f(i^*) + \frac{\lambda}{1+\lambda} m_1 + A + \frac{c_2}{1+\beta_2}, \end{aligned} \tag{A.12}$$

where,

$$\begin{aligned} A &\equiv \frac{\beta_2}{1+\beta_2} E_{I_1}\big(\pi_3 - w_3 - (i_2 + i_3 + d_1 + d_2)\big) \\ &= \frac{\beta_2}{1+\beta_2}\left(\frac{1}{1+\lambda}\big(p(d^*) f(i^*) + m_1\big) - 2\big(p(d^*) i^* + d^*\big) - c_3\right). \end{aligned}$$

From Eqs. (A.11) and (A.12) it follows that,

$$c_2 = -\left\{p(d^*)\left(\frac{\lambda}{1+\lambda} f(i^*) - \lambda i^*\right) - \lambda \frac{d^*}{1-q} + B + A\right\}(1+\beta_2) + \frac{1}{1+\lambda} m_1 (1+\beta_2). \tag{A.13}$$

By applying a similar line of reasoning to period 3, noting that $\beta_3 = 0$, the wage deferred from period 2 is $(\lambda d^* + B)/(1-q)$, and there is no wage deferred to future periods, we obtain,

$$c_3 = -\left(p(d^*)\left(\frac{\lambda}{1+\lambda} f(i^*) - \lambda i^*\right) - \lambda \frac{d^*}{1-q} - \frac{B}{1-q}\right) + \frac{1}{1+\lambda} m_1. \tag{A.14}$$

Both A and B are independent of m_1. Using Eq. (A.14), A can be rewritten as

$$A = \frac{\beta_2}{1+\beta_2}\left(p(d^*) f(i^*) - \lambda i^* - \lambda \frac{d^*}{1-q} - \frac{B}{1-q} - 2\big(p(d^*) i^* + d^*\big)\right).$$

B is independent on m_1 because, using Eq. (15) and noting that $m_2 = m_1$ and $h_2 = h_1$, it follows that

$$m_3 - m_1 = \frac{h_\varepsilon}{h_2 + h_\varepsilon}\big(N_2\big(f(i_2) - f(i^*)\big) + \varepsilon_2\big).$$

Therefore, similar to the case of known ability, c_t is positive for managers with sufficiently high perceived ability m_1. □

References

Anderson, M., Banker, R., Ravindran, S., 2000. Executive compensation in the information technology industry. Manage. Sci. 46, 530–547.

Banker, R., Datar, S., 1989. Sensitivity, precision, and linear aggregation of accounting signals. J. Acc. Res. 27 (1), 21–39.

Boyle, M., 2004. When will they stop? Despite last year's loud cries for pay reform, FORTUNE 500 CEOs made more money than ever in 2003. Here's why change is taking so long. Fortune, May 3.

Bushman, R., Indjejikian, R., 1993. Accounting income, stock price, and managerial compensation. J. Acc. Econ. 16, 3–23.

Core, J., Guay, W., 1999. The use of equity grants to manage optimal equity incentive levels. J. Acc. Econ. 28, 151–184.

Core, J., Guay, W., 2001. Stock option plans for non-executive employees. J. Finan. Econ. 61, 253–287.

Dechow, P., Sloan, R., 1991. Executive incentives and the horizon problem. J. Acc. Econ. 14 (1), 51–89.

Goel, A., Nanda, V., Narayanan, M.P., 2004. Career concerns and resource allocation in conglomerates. Rev. Finan. Stud. 17 (1), 99–128.

Fama, E., 1980. Agency problems and the theory of the firm. J. Polit. Economy 88 (2), 288–307.
Gibbons, Murphy, 1990. Optimal incentive contracts in the presence of career concerns: Theory and evidence. J. Polit. Economy 100 (3), 468–505.
Hirshleifer, D., Thakor, A., 1992. Managerial conservatism, project choice, and debt. Rev. Finan. Stud. 5, 437–470.
Holmstrom, B., 1999. Managerial incentive problems: A dynamic perspective. Rev. Econ. Stud. 66, 169–182.
Holmstrom, B., Milgrom, P., 1991. Multi-task principal-agent problems: Incentive contracts, asset ownership, and job design. J. Law Econ. Org. 7, 24–52. Special Issue.
Holmstrom, B., Tirole, J., 1993. Market liquidity and performance monitoring. J. Polit. Economy 101, 678–709.
Ittner, C., Lambert, R., Larcker, D., 2003. The structure and performance consequences of equity grants to employees of new economy firms. J. Acc. Econ. 34, 89–127.
Milbourn, T., 2003. CEO reputation and stock-based compensation. J. Finan. Econ. 68, 233–262.
Milbourn, T., Shockley, R., Thakor, A., 2001. Managerial career concerns and investments in information. RAND J. Econ. 32, 334–351.
Murphy, K., 1999. Executive compensation. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3. North Holland.
Murphy, K., 2004. Explaining executive compensation: Managerial power versus the perceived cost of stock options. University of Chicago Law Rev. 69, 847–869.
Plitch, P., 2004. Worth of old CEO options ‘skyrocket’ as cos. scale back. Dow Jones Newswires, June 24.
Talmor, E., Wallace, J., 1998. Computer industry executives: An analysis of the new Barons’ compensation. Inform. Syst. Res. 9, 398–414.