Journal of Corporate Finance 29 (2014) 579–593

Contents lists available at SciVerse ScienceDirect

Journal of Corporate Finance

journal homepage: www.elsevier.com/locate/jcorpfin

The prevention of excess managerial risk taking☆

Edward D. Van Wesep *, Sean Wang [1]

University of North Carolina's Kenan-Flagler Business School, Campus Box 3490, McColl Building, Chapel Hill, NC 27599, United States

ARTICLE INFO

Article history:
Received 11 May 2012
Received in revised form 8 April 2013
Accepted 23 April 2013
Available online 14 May 2013

JEL codes:
J33
M55

Keywords:
Severance pay
Executive compensation
Empire building

ABSTRACT

Executives with poor prior performance may be inclined to take excessive risk in the hope of meeting performance targets, in which case a compensation contract featuring severance pay can be optimal. While prior work has shown that severance can induce managers to take positive NPV risks, we show that it can also keep them from taking negative NPV risks. We show that severance should be contingent on results: complete failure should nullify any payments. We also show that mandating a firm size that is larger than first-best, while costly, can help screen for good managers.

1. Introduction

Managers are typically viewed as overly risk averse (Guay, 1999; Smith and Stulz, 1985), and the prior literature on contract design has focused on increasing managers' willingness to take risk via equity grants, stock options, and severance pay.[2] These efforts appear, in practice, appear to be effective.[3] But there are many situations in which even a risk averse manager would take excessive risks, in particular when she finds herself close to missing earnings targets.[4] The CEO may prefer to take negative NPV risks to at least have some shot at meeting performance targets. We show that an optimal contract will often set high performance goals, higher even than first-best for any type of manager, and offer severance for executives who do not meet the bar. We also show that the features of these contracts appear to match several stylized facts about severance agreements that we see in practice.

☆ The authors would like to thank B Douglas Bernheim, Jon Levin, Andres Almazan, and Scott Schaefer for helpful comments. Edward Van Wesep would like to thank the Taube Foundation for financial support. All errors are ours.

* Corresponding author. Tel.: +1 919 360 9689.
E-mail addresses: vanwesep@unc.edu (E.D. Van Wesep), sean_wang@unc.edu (S. Wang).
[1] Tel.: +1 919 962 0284.

[2] See, for example, Core and Guay (1999), Core and Guay (2002), Lys et al. (2007), Yermack (2006), and Manso (2011).

[3] Some examples of riskier decisions, as reported in Coles et al. (2006) include relatively higher investment in R&D, lower investment in PP&E, a more excessive use of cash (lower cash balances), a stronger focus on fewer lines of business, and higher leverage.

[4] Indeed, prior literature studying the actions of money managers, such as Chevalier and Ellison (1997), show that mutual fund managers increase the riskiness of portfolios in the fourth quarter when their return is below comparable benchmarks, while Green (2009) shows that hedge fund managers use discretion in their abilities to mark-to-market Level 3 Assets in order to achieve their desired reporting targets. A large literature on CEO behavior has shown that they are willing to make value-destroying decisions regarding advertising, pricing, production levels, R&D and asset sales in order to meet financial reporting targets. For example, see Baber et al. (1991), Dechow and Sloan (1991), Bartov (1993), Bushee (1998), Graham et al. (2005), Roychowdhury (2006), and Wang and D'Souza (2008). These papers largely concern the tendency of managers to manipulate earnings to beat internally or externally defined benchmarks, like analyst forecases, prior earnings, or management guidance. While there may be no explicit incentive tied to these benchmarks, the implicit incentive appears sufficient to drive behavior. The benchmarks we derive in our paper may be formally written or informally understood, but serve an identical purpose in either case.

0929-1199/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jcorpfin.2013.04.008

In our model, a board chooses a potential CEO it believes to be a good fit with the firm and makes a contract offer. A contract describes what she is to be paid and whether she will be fired, both as a function of the profitability and size of the firm under her leadership. After accepting the contract and joining the firm, the CEO may find that her best efforts will be insufficient to reach the performance targets set by the board. Her decision, then, is whether to “do her best” and under-perform, knowing that she may be fired, or to take negative NPV risks to attempt to meet her performance targets.[5] Our results demonstrate that if the board wishes to prevent the manager from taking excessive risk, it is often the case that a contract featuring severance pay is the least expensive way to induce good behavior from mediocre managers.

The model requires two critical features. First, there must be multiple periods for the concept of “firing” to make sense. Second, there must be a trade-off between expected profit and risk, where an increase in the latter corresponds to a decrease in the former. The model incorporates these two features in the simplest possible way: there are two periods, and profit can take either a high value or a low value. The low value is always equal to zero, so the high value can be defined as the “targeted profit”. As this value increases, the expected profit is assumed to decrease.

While the simplicity is attractive in its own right, an additional benefit to the model is that we need not restrict the space of contracts that the firm may offer. It is often the case in contracting papers that attention is restricted to linear contracts, option contracts, etc. in order to make the problem tractable. By restricting the analysis to two potential outcomes, we are able to consider the full contracting space while still analytically deriving a simple closed form solution for the optimal contract.

We show that compensation packages including severance pay are most valuable when the additional expected profit from having a good manager is high and when ability of the board to screen candidates prior to the contracting stage is high. Importantly, the optimal severance contract pays CEOs who achieve mediocre profits, but not those who fail completely. Pay for truly poor performance actually provides incentives for risk taking, as poor performance typically results from a failed risky action.

The contract will also mandate empire building, in the sense that good managers are expected to choose a firm size above that which is first-best.[6] The intuition is that increases in firm size, beyond first-best, have two effects. First, the profitability of the firm decreases. Second, the probability of achieving targeted profitability for the lower quality CEO is decreasing in firm size, so a larger required size allows the firm to decrease severance pay while still ensuring that the lower quality CEO does not engage in risk taking. In the neighborhood of the first-best firm size, the former effect is second-order while the latter effect is first-order.

In fact, trivial extensions to the model would yield optimal contracts requiring high quality CEOs to take many actions that are excessive relative to the first-best. Any firm attribute that is complementary to CEO skill in determining firm profit could be inserted instead of, or in addition to, firm size without changing our results. For example, if it is more difficult for a lower quality CEO to manage a firm with more divisions, than the firm will require a high quality CEO to choose more than the first-best number of divisions. This model produces empire building and conglomerating as part of an optimal contract.

The existing literature on severance pay as a component of an optimal compensation scheme has focused on a number of different themes. Kahn (1985) demonstrates that when profits are noisy and firing of high quality executives takes place on the equilibrium path, severance payments can allow for efficient risk sharing. Similarly, Manso (2011) suggests that severance can induce optimal risk taking when failure can be a result of either shirking or (efficient) risk taking. Almazan and Suarez (2003) use severance as a method of solving a commitment problem by allowing boards to only fire executives when a replacement is significantly better.[7] Walkling and Long (1984) and Lambert and Larcker (1985) demonstrate that severance payments can align management and shareholder interests by compensating managers for loss of a job in the case that their firm is acquired. Schleifer and Vishny (1989) and Scharfstein and Stein (2000) argue that severance may result from non-arms-length bargaining where managers are entrenched and are effectively able “write their own paychecks”.

Laux (2008) examines a setting that varies CEO-board independence and concludes that because the board's propensity to terminate the CEO is positively correlated with independence, severance pay should be included in an optimal contract to induce self-revelation of low-type CEOs. He then shows that because these severance contracts have adverse incentives on a CEO's effort level, optimal contracts should compensate for this by increasing the level of stock options. Roman and Mueller (2010) reach similar conclusions regarding severance pay and incentive pay, without varying levels of board independence. The motivation of both of these papers is Levitt and Snyder (1997), who show in a more generalized setting that poorly-performing agents must be rewarded by the principal in order to receive early warning of the agent's performance.

While our paper has similarities with the aforementioned, our main intent is not to shed further light on the use of compensation to reveal negative information or the poor fit of CEOs with their respective firms. Rather, we contribute by demonstrating that severance pay can be used to avoid excessive managerial risk-taking and firm-value destruction, particularly when the executive has significant pressures to meet performance-benchmarks. The two most important empirical implications from our model that differ from others in the literature are that, (i) severance should not be paid for truly horrendous performance, and (ii) empire building is optimal for good, but not bad, managers. These aspects of the manager's contract should arise only if (iii) there is significant opportunity for her to take unobservable, negative NPV risks, and (iv) it is likely that the manager is high quality.

Section 2 outlines the simple model underlying the discussion. Section 3 derives the set of possible optimal contracts and shows how parameters affect which out of this set is best. Section 4 concludes. Proofs are provided in the Appendix A.

[5] Whether she will be fired for missing performance targets is endogenously specified in the initial contract.

[6] Empire building, or the managerial practice of choosing a sub-optimally large group of underlings (the entire firm, for a CEO), has been documented for some time. The typical explanation for empire building is that it confers some private benefit to the manager in the form of job safety, perks, additional income etc. In this paper, the manager will be *required* to empire build by the board.

[7] In particular, severance in their setting commits boards to only replace managers when replacement is most valuable. This ex-post protection induces ex-ante optimal investment by managers.

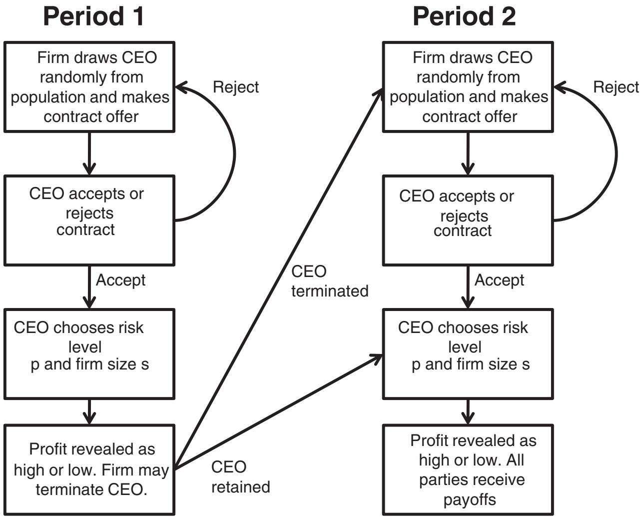

Fig. 1. Timeline.

2. The model

2.1. Players and timeline

There is a firm and an infinite set of risk-neutral agents (potential CEOs) that each live for two periods. Agents are either high type or low type, $\theta \in \{\theta_L,\theta_H\}$, and type is privately observed.[8] The following game structure is shown in Fig. 1. In period one, the board nominates an agent randomly from the CEO pool and makes a take-it-or-leave-it contract offer, details of which we discuss below. The board has some power to screen potential candidates before reaching the contracting stage, but is unable to do so perfectly. With exogenous probability λ, any given nominee is of type θ_H, and with probability $1 - \lambda$ the manager is of type θ_L.

The nominee may accept or decline the contract offer. If the offer is declined then the board nominates a new candidate with a new contract offer. This continues until a candidate accepts the offer of employment. The firm may also choose to hire nobody, in which case the game ends.

Once a nominee accepts a contract offer, she becomes CEO and chooses firm size $s \in [0,\infty)$ and risk $p \in (0,1]$. Nature then chooses firm profit as either high, with probability p, or low, with probability 1-p. Both firm size and profitability are observed by the board.

The board now may propose a change to the original contract terms, which the CEO must accept or reject. If rejected, the original contract remains in place and, if accepted, the new contract is put in place. This includes the termination rule agreed in the initial contract. That is, if the CEO was to be terminated, based upon the observed s and π, her rejection of the new contract offer will result in her termination. If the CEO was to be retained, her rejection of the new contract will leave her in power under the terms of the initial contract.

In period 2, the board may either have retained the CEO from period 1, or must find a new one. If the period 1 CEO was either terminated or voluntarily resigned, then a new CEO is chosen with the same process as in period 1. Once the firm has a CEO in period 2, whether the initial choice was retained or a new one was chosen, the game proceeds with the same parameters as period 1: the CEO chooses risk p and firm size s, nature chooses profitability, and payouts occur.

2.2. Payoffs

In each period, after the CEO has chosen firm size and risk, nature chooses profit to be either high or low. Because we are concerned with managerial risk taking, the CEO chooses the degree of variability in profit. For simplicity, we set low profit to be 0 regardless of the CEO's actions. High profit is denoted $\pi(s,p,\theta)$. A value of $p = 1$ implies that nature will select high profit with certainty. For risk taking ever to make sense, from the CEO's perspective, the high profit outcome must be increasing as the probability of failure, 1-p, increases. We therefore have:

Assumption 1. Profit following success is decreasing in the probability of success, p:

$$\frac{\partial \pi(s,p,\theta)}{\partial p} < 0 \text{ for all } s > 0,\ p \in (0,1) \text{ and } \theta \in \{\theta_L,\theta_H\}.$$

[8] In Section 6, we consider the case where type is unknown to all prior to hiring. The main implications of the model are strengthened.

Risk in this model is *ex ante* detrimental to firm value. This assumption allows for the starkest contrast between the drivers of severance optimality in this model versus previous work, in which severance could be used to protect a CEO and thus incentivize risk-taking. We therefore need expected profit to increase when the CEO chooses a higher probability of success.

Assumption 2. Expected profit is increasing in the probability of success, p:

$$\frac{\partial}{\partial p} p\pi(s, p, \theta) > 0 \text{ for all } s > 0, p \in (0, 1) \text{ and } \theta \in \{\theta_L, \theta_H\}.$$

For any given choice of risk and size, let expected profits be higher for high type CEOs (in some sense this is the definition of "higher quality").

Assumption 3. Profit is higher at firms managed by high types versus low types, ceterus paribus:

$$\pi(s, p, \theta_H) > \pi(s, p, \theta_L) \text{ for all } s > 0 \text{ and } p \in (0, 1).$$

We want the problem to remain bounded, and solvable using standard first-order methods. Therefore:

Assumption 4. π is twice continuously differentiable and concave in s, achieving a maximum at a finite value of s for all $p \in (0,1)$ and $\theta \in \{\theta_L,\theta_H\}$.

Let the size that maximizes expected profits, given a risk choice p, be denoted $s_i^*(p) = argmax_s \pi(s,p,\theta_i)$. We assume that managers' skill and firm size are complementary:

Assumption 5. $s_H^*(p) > s_L^*(p)$, $\forall p \in (0,1)$.

Finally, we must ensure that the bad manager can always attempt to imitate the good manager.

Assumption 6. For any $s > 0$ and $p \in (0,1)$, there exists $p' > 0$ such that the low quality manager can achieve s and $\pi(s, p, \theta_H)$ with probability p':

$$\text{For all } s > 0 \text{ and } p \in (0, 1], \ \exists \ p' > 0 \ \ s.t. \pi(s, p', \theta_L) = \pi(s, p, \theta_H).$$

The principal's payoff equals expected firm profit. Since CEOs are risk-neutral, an actual or potential CEO's payoff is expected lifetime earnings. There is no discounting, and the reservation wage for a type θ_i, $i \in \{L,H\}$, CEO is u_i in each period, with $u_H > u_L$.

2.3. Allowable contracts

A contract given to a new CEO in period 1 is given by $C_1 = \{w_1(s_1, \pi_1), T(s_1, \pi_1), w_2(s_1, \pi_1, s_2, \pi_2)\}$. The first contract element is the wage paid at the end of period 1, which may depend on firm size s_1 and profitability π_1. The second contract element is the termination decision $T \in [0,1]$, which is the probability that the CEO is terminated at the end of period 1, and may depend on firm size and profitability. The third contract element is the wage paid at the end of period 2, which may depend on observed firm size and profitability in both period 1 and 2. The wage payments and termination decision may be unrelated to each other. That is, it is allowable for the firm to make a wage payment in period 2 even while terminating the worker.

If the period 1 CEO is fired or quits, a contract must be given to a new CEO in period 2. This contract is given by $C_2 = \{w_3(s_2,\pi_2)\}$. It simply specifies a wage, which may depend upon firm size and profitability in period 2.

The functions w_1, w_2, w_3, and T may be increasing, decreasing, continuous or discontinuous in their arguments. Our sole restriction is limited liability: $w_i \geq 0$ for $i = 1, 2, 3$.

The limited liability restriction has two components. First, there is only so much that the CEO can be punished for bad performance. We normalize the maximum punishment to a wage of zero, but other choices would yield a qualitatively identical model. Second, there is limited liability period-by-period. This implies that there cannot be negative wages in period 1, compensated with positive wages in period 2. This restriction is important theoretically – if the CEO could post an arbitrarily large bond to the firm, then the contracting problem would be degenerate. It is also reasonable empirically – it is unclear how a potential CEO would come to have half a lifetime's worth of wages with which to post the bond. Period-by-period limited liability also means that the firm cannot "claw back" prior pay. This aspect of the limited liability assumption will not bind, as the optimal contract will feature pay that is deferred as long as possible.

3. Optimal compensation schemes

Now that the game is fully specified, we can solve for the optimal contract design and the choices that follow. For the rest of the paper, let $s_i = s_i^*(1)$ be the socially optimal firm size for a manager of type i taking the socially optimal level of risk, which is none. Define ρ to be the probability of success for a low type attempting the socially optimal size and profit of the high type: $\pi(s_H, \rho, \theta_L) = \pi(s_H, 1, \theta_H)$. Socially optimal decision-making henceforth refers to type i choosing $p = 1$ and $s = s_i$. Let the profit from socially optimal decision-making be $\pi_i = \pi(s_i, 1, \theta_i)$.

If the board could perfectly observe the ability of potential hires, the first-best contract would induce socially optimal decision-making from the hired CEO in both periods, and would only be offered to the type of agent that is more efficient for the firm. Firm profit is revenue from the efficient choices of size and risk, π_i, net of reservation pay u_i, so profit equals $\pi_i - u_i$. The firm would choose the type of CEO for whom $\pi_i - u_i$ is greater, and then only pay the CEO when outcomes are consistent with socially optimal decisions.

When the firm cannot observe agents' types, the problem is more complicated. Without limited liability, the firm could do as well as if it could observe agent type. If it were efficient to hire low type managers, it could simply offer a wage u_L, paid only when outcomes were consistent with socially optimal decisions. This contract would attract only low-type managers and would induce profit-maximizing decisions. If it were efficient to hire high type managers, the firm could offer very intense incentives, delivering u_H for outcomes consistent with socially optimal decision-making by the high type of agent, and delivering some payoff $-X$ for any other observed outcome. A low type agent could only achieve profit associated with $s = s_H$, $p = 1$ and $\theta = \theta_H$ with probability $\rho < 1$. Therefore, if $-X$ were sufficiently negative, the contract would not deliver expected utility u_L to a low type, so low types would not accept such a contract offer. Thus, the optimal contract could screen for the desired type of CEO while costing the firm no more than if it could observe CEO quality. We therefore have:

Lemma 1. *Without limited liability, the optimal contract would achieve first-best outcomes, identical to the case where agent type is observable.*

The model in this paper would not be particularly interesting if limited liability were not an issue, but in practice it is a constraint. A first step to finding the optimal contract is to find the best contract that achieves a particular goal. A goal might be, for example, "only hiring low types", or "trying to hire high types, but being content with a low type if you get stuck with one". We define contracts that are optimal for achieving a particular goal as "constrained optimal." To find the "globally optimal" contract, we compare firm value across potential goals, assuming optimal contracting toward that goal.

This process is backward induction. The firm chooses a hiring goal and, given that goal, chooses the optimal contract for achieving it. To find the optimal contract, we solve backwards. Find the optimal contract for each goal, and then choose the optimal goal, given payoffs resulting from those contracts. The best contract for the goal yielding the highest payoff is the optimal contract. We identify constrained optimal contracts below.

3.1. A constrained optimal contract that attracts only low types

Lemma 2. *A constrained optimal contract that attracts only low types includes a flat wage in both periods and no termination:*

1. $w_1(s_1 = s_L, \pi_1 = \pi_L) = u_L$ *while* $w_1 = 0$ *for all other* $\{s_1,\pi_1\}$ *combinations*
2. $w_2(s_1 = s_L, \pi_1 = \pi_L, s_2 = s_L, \pi_2 = \pi_L) = u_L$ *while* $w_2 = 0$ *for all other* $\{s_1,\pi_1,s_2,\pi_2\}$ *combinations*
3. $T(s_1 = s_L, \pi_1 = \pi_L) = 0$ *(retain), while* $T = 1$ *(terminate) for all other* $\{s,\pi\}$ *combinations.*

The contract only offers sufficient compensation to attract low types, and pays them as little as possible while inducing socially optimal decision-making. This contract will always be globally optimal when $\pi_L - u_L \geq \pi_H - u_H$ and the value to the firm of this contract is $2(\pi_L - u_L)$. When $\pi_L - u_L < \pi_H - u_H$, the board may prefer to hire only high types, if possible.

3.2. A constrained optimal contract that attracts only high types

Lemma 3. *A constrained optimal contract that only attracts high types features deferred compensation with the option of termination:*

1. $w_1 = 0$ *for all* $\{s_1, \pi_1\}$ *combinations*
2. $w_2(s_1 = s_H, \pi_1 = \pi_H, s_2 = s_H, \pi_2 = \pi_H) = 2u_H$ *while* $w_2 = 0$ *for all other* $\{s_1, \pi_1, s_2, \pi_2\}$ *combinations*
3. $T(s_1 = s_H, \pi_1 = \pi_H) = 0$ *(retain) while* $T = 1$ *(terminate) for all other* $\{s_1, \pi_1\}$ *combinations.*

This contract offers the minimum necessary pay to attract high types, and does so in a way to be as unappealing as possible to low types. All pay is deferred until the end of the contract, which means that a low type must successfully imitate a high type twice in order to be paid. The contract gives an incentive for a high type to take socially optimal actions by only paying her when outcomes are consistent with those actions.

It is not always that the case that attracting only high types is feasible. A low type receiving the contract offer above makes the following calculation. If she accepts the offer, she will attempt to imitate the high type and will choose $s_1 = s_H$ and $p = \rho$. If she fails, which happens with probability $(1 - \rho)$, she is fired and receives reservation pay of u_L in period 2. If she succeeds, which happens with probability ρ, then she makes the same choices in period 2. If she succeeds, then she receives payoff $2u_H$. If she fails, she receives nothing. Her payoff from accepting the contract is therefore $(1 - \rho)u_L + \rho^2(2u_H)$. If she rejects the initial contract offer, she receives $2u_L$. She therefore only rejects the contract offer when

$$(1-\rho)u_L + \rho^2(2u_H) \leq 2u_L,$$

which yields the following necessary and sufficient condition for the existence of contracts that screen for high types.

Lemma 4. *The goal of hiring only high types is achievable if and only if* $\frac{u_H}{u_L} \leq \frac{1}{2}\frac{(1+\rho)}{\rho^2}$

While the contract defined in Lemma 3 is an example of a constrained optimal contract that attracts only high types, there are others. A contract that pays a performance bonus of $2u_H$ for achieving a profit level of $\pi \geq \pi_H$ twice yields the same constraints and is more intuitive. This contracting goal is only achievable when either both types have similar outside options or if CEO type has a significant effect on firm profits (i.e. it is difficult for low types to imitate high types).

As was stated above, it is best to target one type if possible. Firms that find it more profitable to hire high types may or may not be able to do so, as shown as shown in Lemma 3. If the types' outside options are similar, then we would expect $\pi_H - u_H > \pi_L - u_L$. Therefore, in practice we would expect that firms using this type of pay scheme to be firms where CEO skill is highly specific to the firm. The value to the firm of this contract is $2(\pi_H - u_H)$.

3.3. A constrained optimal contract that attracts both types and terminates low types

Now we consider the most interesting situation regarding parameter values: the one in which the board would like to hire only high types but cannot, because low types have significantly lower outside options (i.e., the case of general rather than firm-specific management skills). The two constraints on parameters that yield this situation are:

$$\pi_H - u_H > \pi_L - u_L$$
$$\frac{u_H}{u_L} > \frac{1}{2}\frac{1+\rho}{\rho^2}.$$

The board must specify actions for each type, payoffs for success and failure following those actions, and payoffs for outcomes that do not occur on the equilibrium path. There are two potentially optimal contracts that attract high types: one in which low types are fired when their actions reveal them, and one in which they are allowed to stay on. It is shown in the Appendix A that it cannot be optimal to use other schemes, but either of these can be optimal depending on parameter values. We later discuss when each is best.

If a firm wishes to dissuade low types from imitating good managers by requiring them to choose excessive size or take excessive risk, it is less expensive to require excessive size. The intuition is that any constrained optimal contract that specifies different behavior for the two types must specify socially optimal actions for the low types: they should be told to take no risk. In this case, in the neighborhood of socially optimal play for high types, a small increase in specified high type risk will result in only high types earning profits of 0 on the equilibrium path. Firing for failure is not renegotiation-proof, so the incentive for low types to deviate to taking some risk is significantly loosened: if they fail they are seen as high types and are not fired! An apparently small strengthening in the low type IC constraint by making it more difficult to imitate high types would actually result in a significant weakening in the IC constraint, making the requirement of a small level of risk for high types bad for the firm. A small increase in size, as we will see, is always good.

Before continuing, some definitions will be useful. Let $p'(s)$ be implicitly defined by $\pi(s,1,\theta_H) = \pi(s, p'(s), \theta_L)$. If a high type is assigned a target firm size s, then $p'(s)$ is the probability that a low type who chooses the same firm size can achieve the same profit. Note that $\rho = p'(s_H)$. We define empire building to be a choice of s by a CEO that is larger than first-best: i.e. $s > s_i^*$. Let s^{1H} and s^{3H} be defined by

$$\left.\frac{\partial \pi(s,1,\theta_H)}{\partial s}\right|_{s=s^{1H}} = \frac{1-\lambda}{\lambda}(2\rho u_H - u_L)\left.\frac{\partial p'(s)}{\partial s}\right|_{s=s^{1H}} \tag{1}$$

$$\left.\frac{\partial \pi(s,1,\theta_H)}{\partial s}\right|_{s=s^{3H}} = \frac{1-\lambda}{\lambda}u_H\left.\frac{\partial p'(s)}{\partial s}\right|_{s=s^{3H}}. \tag{2}$$

Note that $u_H > 2\rho u_H - u_L$, so $s_H < s^{3H} < s^{1H}$. Finally, let $\pi^{1H} = \pi(s^{1H}, 1, \theta_H)$ and $\pi^{3H} = \pi(s^{3H}, 1, \theta_H)$.

When the board wishes to hire high types and remove low types after the first period, it must induce low types to reveal themselves. If a low type reveals herself, she is paid X in period 1 and is fired, receiving u_L in period 2. If the low type imitates, then she receives w_1 in period 1 if she succeeds and 0 if she fails.[9] In the latter case, she receives u_L in period 2. In the former case, she receives w_2 in period 2 if she succeeds and zero if she fails. The probability that she successfully imitates the high type depends on the firm size that the high type is expected to maintain, s. The probability that she can successfully imitate is $p'(s)$. Her IC constraint is therefore

$$X + u_L \geq (1-p'(s))u_L + p'(s)w_1 + (p'(s))^2 w_2.$$

The way to find the optimal contract is to begin with the first-best contract, which specifies socially optimal decision-making from both types of agents and pays each agent's reservation pay. This contract will not satisfy the low type's IC constraint if

[9] We assume that the optimal contract punishes failure as much as possible. This is, of course, true, and we show a more formal derivation in the Appendix A.

$\frac{u_H}{u_L} > \frac{1}{2}\frac{1+\rho}{\rho^2}$. In order to get the constraint to hold, the required firm size s and the payouts X, w_1 and w_2 can be adjusted. In particular, X must increase, w_1 or w_2 must decrease, or $p'(s)$ must decrease (so s must increase). Note that $\frac{\partial \pi(s,1,\theta_H)}{\partial s}\Big|_{s=s_H} = 0$ (π is concave and maximized at $s = s_H$, given manager type θ_H) but $\frac{\partial p'(s)}{\partial s}\Big|_{s=s_H} < 0$.[10] Therefore, the required firm size for a high type of manager can be increased marginally, loosening the low type's IC constraint while only having a second order effect on profit. Empire building will be a part of the optimal contract.

w_1 and w_2 can only be reduced so much if the contract is still to attract high types. The high type is indifferent to receiving pay in period 1 or 2, so since $p'(s) < 1$, we can shift all pay to w_2, loosening the low type's IC without affecting high type behavior. An increase in X clearly has a first order effect on profit as well as the low type's IC, but some positive value of X will be optimal.

Proposition 1. *A constrained optimal contract that attracts both types and fires bad managers after period one requires that low types take socially optimal actions, are fired, and receive severance pay. High types engage in empire building initially and socially optimal decision-making in period two, while being paid entirely in deferred compensation:*

1. $w_1(s_1 = s_L, \pi_1 = \pi_L) = p'(s^{1H})(2\rho u_H - u_L)$ *and* $w_1 = 0$ *for all other* $\{s_1, \pi_1\}$ *combinations.*
2. $w_2(s_1 = s^{1H}, \pi_1 = \pi^{1H}, s_2 = s_H, \pi_2 = \pi_H) = 2u_H$ *while* $w_2 = 0$ *for all other* $\{s_1, \pi_1, s_2, \pi_2\}$ *combinations.*
3. $T(s_1 = s^{1H}, \pi_1 = \pi^{1H}) = 0$ *(retain) while* $T = 1$ *(terminate) for all other* $\{s_1,\pi_1\}$ *combinations.*

Young high quality managers must choose excessive firm size to prove their mettle. When choosing expectations for a new CEO, the firm trades off requiring high types to manage a larger firm, which is costly, with simply offering more severance to low types who reveal themselves. The solution is for some of each: the firm is expected to be sub-optimally large, but there is severance pay as well.

Corollary 1. *Severance is decreasing in* u_L*, and increasing in* $p'(s)$ *and* u_H*.*

As the low type's outside option, u_L, improves, the amount of necessary severance pay decreases: low types are more willing to leave the job for greener pastures. As low types are less able to imitate high types ($p'(s)$ decreases point-wise in s) severance also decreases, because it depends on the low type's ability to gamble on a high outcome. As high type reservation utility increases, severance pay increases: high type mangers must be paid more to meet their reservation utility, but as this amount grows, severance must grow apace to induce low types to accept their own firing rather than taking risk. The amount of required empire building depends upon the precise shape of π. If high types can increase size without too much of an effect on profit, while the same is not true for low types, then more empire building is called for.

3.4. A constrained optimal contract that attracts and retains both types of manager

It can also be best to hire both types, induce separation, and retain both in period two.

Proposition 2. *A constrained optimal contract that attracts both types and fires neither after period one requires that low types take socially optimal actions in both periods, while high types empire build in period one and take socially optimal actions in period two. Both types are paid entirely in deferred compensation, with high types earning more than low types.*

1. $w_1 = 0$ *for all* $\{s_1,\pi_1\}$ *combinations*
2. $w_2(s_1 = s^{1H}, \pi_1 = \pi^{1H}, s_2 = s_H, \pi_2 = \pi_H) = 2u_H$ *and* $w_2(s_1 = s_L, \pi_1 = \pi_L, s_2 = s_L, \pi_2 = \pi_L) = 2p'(s^{1H})\rho u_H + (1 - p'(s^{1H}))u_L$ *while* $w_2 = 0$ *for all other* $\{s_1, \pi_1, s_2, \pi_2\}$ *combinations*
3. $T(s_1 = s^{1H}, \pi_1 = \pi^{1H}) = 0$ *(retain) and* $T(s_1 = s_L, \pi_1 = \pi_L) = 0$ *(retain), while* $T = 1$ *(terminate) for all other* $\{s_1, \pi_1\}$ *combinations.*

Corollary 2. *Pay for the low type's prescribed actions is increasing in* u_H *and* u_L*.*

This contract is also fairly simple. Empire building is required for high types in the first period as before because, in the neighborhood of $s = s_H$, the effect of empire building on revenue is second order while the effect on low type pay through its IC constraint is first order. All pay is deferred to generate maximal incentives to behave. The effects of parameter changes on low type pay are similar to before, except that an increase in low type outside options actually makes this contract less valuable. Low type pay is increasing in outside options because it makes being fired (and therefore imitating high types) more attractive. As before, there is no uncertainty on the equilibrium path, so outcomes not specified in the contract are punished with firing and no pay.

3.5. Globally optimal contracts

If parameters are such that the firm would ideally hire only high types but is unable to do so, there are three classes of constrained optimal contracts, each of which sometimes is globally optimal.[11] If the board offers the severance contract, then it

[10] The effect of an increase in s is second order on π but first order on $p(s)$.
[11] See the Appendix A for proofs rejecting other classes of contract.

hires a high type in period one with probability λ and receives payout $\pi^{1H} + \pi_H - 2u_H$. With probability $1 - \lambda$, it hires a low type. In this case, it receives payout $\pi_L - p'(s^{1H})(2\rho u_H - u_L)$ and then must draw a new CEO in period two, in which case it receives $\pi^{3H} - u_H$ with probability λ and $\pi_L - p'(s^{3H})u_H$ with probability $1 - \lambda$. Let V_j be the value of the firm employing contracting strategy $j \in \{severance, low, permissive\}$. Then:

$$V_{severance} = \lambda\left(\pi^{1H} + \pi_H - 2u_H\right) + (1-\lambda)\left[\pi_L - p'\left(s^{1H}\right)(2\rho u_H - u_L) + \lambda\left(\pi^{3H} - u_H\right) + (1-\lambda)\left(\pi_L - p'\left(s^{3H}\right)u_H\right)\right].$$

If the board only hires low types, then firm value is

$$V_{low} = 2(\pi_L - u_L).$$

If the board hires both types but fires neither, it gets a high type with probability λ and earns $\pi^{1H} + \pi_H - 2u_H$, and gets a low type with probability $1 - \lambda$ and earns $2\pi_L - 2p'(s^{1H})\rho u_H - (1 - p'(s^{1H}))u_L$. The value of the firm is therefore

$$V_{permissive} = \lambda\left(\pi^{1H} + \pi_H - 2u_H\right) + (1-\lambda)\left[2\pi_L - 2p'\left(s^{1H}\right)\rho u_H - \left(1 - p'\left(s^{1H}\right)\right)u_L\right].$$

We now can state our primary global optimality result.

Proposition 3. *The contract featuring severance is globally optimal when high types of manager are much more valuable than low types for the firm, and when the board is more able to screen candidates.*

To see the intuition for this proposition, let $J \equiv \pi_H - \pi^{1H}$, $K \equiv \pi_H - \pi^{3H}$, $p_1 \equiv \rho - p'(s^{1H})$ and $p_3 \equiv \rho - p'(s^{3H})$. This parameterizes the profit function in a way that makes comparative statics manageable. The value of the firm may be higher or lower when the severance contract is used rather than the low contract (only hiring low types), depending on parameters, but there do exist situations when either is higher than the other (see the proof of Lemma 7 in the Appendix A for examples of when each can be best). Let $\overline{V}$ be the difference between firm value when the severance contract is used versus the contract where only a low type is hired. To see how parameter values determine the features of the optimal contract, we can take a derivative of $\overline{V}$ with respect to each parameter and note the sign.

(x)	$\frac{d}{dx}\overline{V}$	Sign
π_H	$\lambda(3 - \lambda)$	>0
π_L	$(1 - \lambda)(2 - \lambda) - 2$	<0
u_H	$-2\lambda - (1 - \lambda)(2\rho(\rho - p_1) + \lambda + (1 - \lambda)(\rho - p_3))$	<0
u_L	$(1 - \lambda)(\rho - p_1) + 2$	>0
λ	$-J + (3-2\lambda)(\pi_H - u_H) - \pi_L + (\rho - p_1)(2\rho u_H - u_L) - (1-2\lambda)K$	>0
ρ	$(1 - \lambda)[(\lambda - 4\rho - 1)u_H + u_L]$	<0

The severance contract is more valuable relative to the low contract when π_H and u_L are high, π_L and u_H are low, ρ is low and λ is high. In essence, severance contracts are more valuable when CEO quality is important (in particular, more important to the firm than to outsiders), when low type managers cannot easily imitate high type managers, and when the board is better able to screen candidates.

Hiring both types but firing low types when they reveal themselves entails a cost: severance must be paid to entice low types to reveal themselves, and this expense may be very high. The firm might get a high type in period two upon firing a low type in period one, but it might not. If not, then the severance was lost for no gain, but if so then the firm has a more valuable leader. Therefore, the better the firm is at screening candidate CEOs, the more likely a firing results in a superior replacement and the more valuable the severance contract becomes. In the model, this is captured by the parameter λ, which is defined to be the probability that a new CEO will be high quality.

The curvature of π is parameterized by J, K, p_1 and p_3. Without evaluating derivatives, it is clear that as J and K shrink (the profit function for high types is flatter), and p_1 and p_3 rise (the profit function for low types is steeper), the severance contract becomes more valuable. A flatter profit function for high types and a steeper one for low types have both a first and second order positive effect on profit under the severance contract.

Perhaps the most natural situation in which a severance contract is best is that of a large, complex firm. At large firms, the level of CEO pay is second or third order relative to both the level of profit and the impact of CEO quality on profit: $u_H - u_L$ is approximately equal to zero relative to $\pi_H - \pi_L$. The difference $\pi_H - \pi_L$ would be especially large at complex firms where talent is more critical. In this case, the severance contract is the optimal contract.

3.6. Optimal contracts when managers do not know their own types

It is worth investigating what would occur if potential hires were unaware of their skills prior to joining the firm. In this case, contracts cannot be used to screen candidate CEOs. Suppose that all CEOs, prior to joining the firm, perceive themselves to be high types with probability λ and low types with probability 1-λ so that their beliefs are aligned with those of the firm. Let their reservation utilities at the outset be $u = \lambda u_H + (1 - \lambda)u_L$ in each period. After joining the firm, the new CEO learns her type, chooses s and p, and the game proceeds as in the preceding sections.

The following proposition is stated without proof (the proof follows closely to that of Propositions 1, 2 and 3).

Proposition 4. *If candidate CEOs only learn their type immediately after joining the firm, the severance contract is globally optimal when π_H and u_L are high, π_L and u_H are low, ρ is low and λ is high.*

Essentially, the properties of the severance contract and the circumstances in which it is optimal are largely the same regardless of whether candidate CEOs know their types. The severance contract is only useful when screening is impossible, so assuming that potential CEOs do not know their types entails no new restriction. Indeed, screening contracts that only attract high or low types are ruled out, so the severance contract will be optimal for a larger set of parameters.

The important assumption of the model is that upon joining the firm, new CEOs are able to determine how the firm is likely to perform under their leadership given choices regarding risk and size. It is not necessary for new CEOs to see themselves as low types (i.e., it is not necessary for them to realize that anyone else could do a better job).

4. Conclusion

This paper offers a novel justification for severance pay as a method of reducing excessive managerial risk-taking. We argue that severance pay can be part of an optimal contract, but that the severance payment must depend on performance. High profit is rewarded with retention and deferred compensation, and medium profit is rewarded with firing and severance, but complete failure must be met with firing and no payment. The reason is straightforward: mediocrity results from weak managers doing their best whereas failure can only result from undesirable risk taking, the precise action that an optimal contract must discourage. In practice, most actual severance agreements depend on performance insofar as they do not pay out if termination is for cause. Whether this is equivalent to – or more than – complete failure in our model is open for debate, but most severance contracts do indeed restrict payments under the "worst case scenario," as the model predicts. While it is rare in practice for firms to make use of these provisions, their use is also never present on the equilibrium path in the model.

Our model suggests that offering a contract with severance becomes optimal as the difference between the firm's profits and the executive's reservation wages increases, and as the importance of the executive's skill rises. It also suggests that severance payments themselves increase as profits driven by a CEO's skill also increase (outside options for high types rise and outside options for low types fall). As corporations have grown larger, profits have grown relative to pay. As firms have grown more complex and industries more competitive, returns to CEO skill may have also increased. The model therefore suggests that, over time, severance payments should have grown in size and frequency, and indeed they have. Cross-sectionally, the model predicts that severance packages should be larger and more common at larger and more complex firms that operate in more complex industries. Again, the implications of the model appear to fit well with casual observation.

Our model also makes a stark prediction regarding when success should be compensated: at retirement. When parameters are such that it is possible to target one particular type of manager in the hiring process, compensation can be paid out more evenly over a manager's lifetime, but when parameters prevent these contracts, it becomes strictly better for the firm to defer compensation as long as possible. This minimizes the probability of a low type successfully imitating a high type and therefore lowers the expected wage cost. In practice, pay deferral, while not complete, is still extensive. Lee Raymond netted $400 million in deferred compensation upon retirement from Exxon Mobil in 2005, which greatly exceeded his annual take-home pay. While pay is not necessarily deferred until retirement for all executives, pay is often in the form of long-vesting options or restricted stock and therefore is deferred significantly.

We also show that high quality CEOs should be required to expand the firm beyond the first-best optimal size. That is, what appears to be empire building is optimal. This fact is justified by an envelope argument: at the margin where the firm is the first-best optimal size, increasing the size has a second-order effect on profit, but a first-order effect on the probability that a low quality CEO can match the profit that the high quality CEO could achieve. This allows a first-order reduction in expected compensation for the low type. This argument follows for many firm attributes other than size. So long as the attribute is complementary to CEO skill in generating profit, the high quality CEO should be required to choose an excessive level of the attribute. Other examples might be firm scope (if managing multiple divisions is relatively more difficult for a lower quality CEO) or asset intangibility (if earning a high return on intangible assets is relatively more difficult for the lower quality CEO).

Appendix A

Proof of Lemma 1. Consider the following two contracts:

1. Offer a wage $w = u_L$
2. Offer a payment $w = u_H$ if $s = s_H, \pi = \pi_H$ is observed and $-X$ otherwise where X is arbitrarily large

The first contract would only attract low types and socially optimal low type behavior would occur on the equilibrium path. The second would attract only high types (for some X large enough). Firm profit from choosing one of these contracts is $\max_i(\pi_i - u_i)$ in each period.

The expected per-period profit from a contract that attracts both types is

$$\begin{aligned}&\lambda[E(\pi|\theta_H)-E(compensation|\theta_H)]+(1-\lambda)[E(\pi|\theta_L)-E(compensation|\theta_L)]\\&\leq\lambda[\pi_H-u_H]+(1-\lambda)[\pi_L-u_L]\\&\leq\max_i[\pi_i-u_i]\end{aligned}$$

■

Proof of Lemma 2. If pay is simply $w = u_L$ then it is (weakly) optimal for the CEO to choose $s = s_L$, $p = 1$ to get profit of π_L. This scheme will not attract high types because maximum pay is $2w = 2u_L < 2u_H$. ■

Proof of Lemma 3. Suppose we have some contract that pays $A(s_1,\pi_1)$ in the first period and, if the CEO from the first period is not fired, $B(s_1, s_2, \pi_1, \pi_2)$ in the second in the second period. Since the goal is to only hire high types and we require renegotiation-proofness, the contract must induce choices of s_H, π_H in each period.[12] This is best achieved by maximizing the penalty from other actions. Let $A(s_1, \pi_1) = A$ if $s_1 = s_H$ and $\pi_1 = \pi_H$ and 0 otherwise and let $B(s_1, s_2, \pi_1, \pi_2) = B$ if $s_2 = s_H$ and $\pi_2 = \pi_H$ and 0 otherwise. We need IR constraints to be satisfied as below

$$\begin{aligned}&A+B\geq 2u_H\\&\rho(A+\rho B)+(1-\rho)u_L\leq 2u_L\end{aligned}$$

Rewrite the IR constraint for the low types as

$$A+\rho B\leq\frac{(1+\rho)}{\rho}u_L$$

The value of the firm is linear in $A + B$ (only high types are hired so the total wage bill is $A + B$). By moving pay from A to B, we weaken the constraint. Because of limited liability this constraint is weakest when $A = 0$. Now, our IR constraints become

$$2u_H\leq B\leq\frac{(1+\rho)}{\rho^2}u_L$$

The set of possible values for B is only non-empty if

$$\frac{u_H}{u_L}\leq\frac{1}{2}\frac{(1+\rho)}{\rho^2}$$

Value is decreasing in B, so $B = 2u_H$ completes the contract. ■

Lemma 5. It cannot be optimal to hire both types, fire for poor performance and induce pooling.

Proof. In order for the contract to be renegotiation-proof, it must be that the pooling size is $s'' \geq s_H^*(p'')$. If the board requires pooling at some smaller size, then high types could argue that they want to choose a larger firm size. The board and CEO should be able to come to some pay agreement that allows the high types to choose the larger size and pay a little extra for it. As long as the extra pay is small enough, low types would not want to make this argument and therefore the renegotiation would take place. Given that we are constrained to have the board hire both types, fire for poor performance, and induce pooling at some firm size larger than $s_H^*(p'')$, the optimal pooling location is $s = s_H$, $\pi = \pi_H$.

Regardless of the initial contract, the board could alter the contract in period 2 to the benefit of the firm and low types by offering a contract that induces them to choose $s = s_L$, $\pi = \pi_L$. This contract could have the same expected wage bill for the firm, same expected pay for the low types and high types, and induce socially optimal choices. Therefore, for a contract to be renegotiation-proof it must make that offer in period two. The contract then would specify a payment of u_H if $s = s_H$, $\pi = \pi_H$ is achieved in the first period with a payment of 0, and firing if anything else occurs. In period two the contract would specify a payment of u_H if $s = s_H$, $\pi = \pi_H$ and ρu_H if $s = s_L$, $\pi = \pi_L$ is realized and zero otherwise.

If the board requires pooling in period one at $s = s_H$, $\pi = \pi_H$, then with probability 1-ρ low types will fail and be fired. The board could instead write a contract which specifies a probabilistic response to a CEO choice of $s = s_L$, $\pi = \pi_L$ in the first period. They could fire with probability 1-ρ and not pay money, or keep on as CEO with probability ρ and pay u_H. Expected revenue to the firm is higher since $s = s_L$, $\pi = \pi_L$ is socially optimal, and low types are indifferent; thus, they would be willing to choose $s = s_L$, $\pi = \pi_L$. High types clearly would still choose $s = s_H$, $\pi = \pi_H$. ■

[12] If any other actions are specified then, after the agent is hired, she and the board will be able to renegotiate to optimal play.

Proof of Proposition 1. The board proposes a contract specifying period one high and low type actions: ($s = s^{1L}, \pi = \pi^{1L}$) and ($s = s^{1H}, \pi = \pi^{1H}$). The contract will pay 0 if any other profit number or firm size is observed because both types have probability 0 of failing on the equilibrium path and the board wants to create the largest possible disincentives to deviate.

To find the optimal contract that fires low types and induces separation, the board will set pay levels A_i, B_{ij}, where A_i is paid at the end of the first period if $s = s^{1i}, \pi = \pi^{1i}$ is observed, and B_{ij} is paid at the end of the second period if $s = s^{1i}, \pi = \pi^{1i}$ was observed in period one and $s = s^{2j}, \pi = \pi^{2j}$ was observed in period two. In the second period, if a new CEO is to be hired, the board will offer C_k if $s = s^{3k}, \pi = \pi^{3k}$ where $i, j, k \in \{L, H\}$.

We then have IR and IC constraints in the first period of:

$$
\begin{aligned}
&A_H + B_{HH} \geq 2u_H \\
&B_{HH} \geq u_H \\
&A_H + B_{HH} \geq A_L + u_H \\
&A_H + B_{HH} \geq A_H + B_{HL} \\
&A_L + u_L \geq 2u_L \\
&A_L + u_L \geq p'\left(s^{1H}\right)(A_H + B_{HL}) + \left(1 - p'\left(s^{1H}\right)\right)u_L \\
&A_L + u_L \geq p'\left(s^{1H}\right)\left(A_H + p'\left(s^{2H}\right)B_{HH}\right) + \left(1 - p'\left(s^{1H}\right)\right)u_L
\end{aligned}
$$

Shifting from A_H to B_{HH} weakens some constraints without strengthening any, so we can set $A_H = 0$. We can also set $B_{HL} = 0$ without tightening any constraints. Thus, we obtain

$$
\begin{aligned}
&B_{HH} \geq 2u_H \\
&B_{HH} \geq A_L + u_H \\
&A_L \geq u_L \\
&A_L \geq p'\left(s^{1H}\right)p'\left(s^{2H}\right)B_{HH} - p'\left(s^{1H}\right)u_L
\end{aligned}
$$

Then $A_L = \min\{u_L, p'(s^{1H})(p'(s^{2H})B_{HH} - u_L)\}$, $B_{HH} = \min\{2u_H, A_L + u_H\}$. Since it is not possible to only hire high types we get

$$
\begin{aligned}
&B_{HH} = 2u_H \\
&A_L = p'\left(s^{1H}\right)\left(2p'\left(s^{2H}\right)u_H - u_L\right)
\end{aligned}
$$

Let $s = s^{3i}$, $i \in \{L,H\}$ be specified for new CEOs hired in period two. Then in the second period we have the following constraints

$$
\begin{aligned}
&C_H \geq u_H \\
&C_H \geq C_L \\
&C_L \geq u_L \\
&C_L \geq p'\left(s^{3H}\right)C_H
\end{aligned}
$$

Then we get $C_H = u_H$, $C_L = p'(s^{3H})u_H$. The final thing we must do is choose s^{1H}, s^{2H} and s^{3H}. This contract has defined what the firm must pay to get specified actions, but has not said anything yet about what those actions should be. The value of the firm is

$$
\begin{aligned}
V_{severance} = {} & \lambda\left(\pi^{1H} + \pi^{2H} - 2u_H\right) \\
& + (1-\lambda)\left[\pi^{1L} - p'\left(s^{1H}\right)\left(2p'\left(s^{2H}\right)u_H - u_L\right) + \lambda\left(\pi^{3H} - u_H\right) + (1-\lambda)\left(\pi^{3L} - p'\left(s^{3H}\right)u_H\right)\right]
\end{aligned}
$$

We clearly want to require low types to take actions that maximize π^{iL} because doing so does not affect the value function negatively anywhere. Given that we are requiring types to take no risk, $\pi^{iL} = \pi_L$, $i \in \{1,3\}$.[13] The same cannot be said for high types. Requiring them to choose sizes larger than socially optimal reduces the probability with which low types can successfully imitate them and therefore reduces the severance that must be offered.

We now ask what the optimal levels of these three variables are. First, a renegotiation-proof contract must set $s^{2H} = s_H$ because once high types have credibly revealed themselves in period one they could otherwise always renegotiate with the board

[13] We could allow low types to take risk and get somewhat messier functions. The result, however, would be identical.

to get a mutually beneficial change in terms. Therefore there cannot be empire-building in period two given that the CEO is a carry-over from period one.

As for s^{1H} and s^{3H}, these are both greater than s_H. They are apparently not less, because in that case high type profits could increase while low type pay decreases when they are raised to s_H. At $s = s_H$, the effect of an increase of s on π is, by assumption, second order (π is maximized and differentiable at $s = s_H$). The effect on low type wage is, however, first order. Pay is $A_L = p'(s^{1H})(2p'(s^{2H})u_H - u_L)$ or $C_L = p'(s^{3H})u_H$ and $p(t)$ is strictly monotone by assumption. Therefore, $s^{1H}, s^{3H} > s_H$. ■

Lemma 6. *It cannot be optimal to keep low types on as CEO with some probability $x \in (0,1)$.*

Proof. Let's find the optimal contract for any given x. Let other definitions be as in Proposition 1.

The constraints are

$$\begin{aligned}
&A_H + B_{HH} \geq 2u_H \\
&B_{HH} \geq u_H \\
&A_H + B_{HH} \geq A_L + (1-x)u_H + xB_{LH} \\
&A_H + B_{HH} \geq A_L + (1-x)u_H + xB_{LL} \\
&A_H + B_{HH} \geq A_H + B_{HL} \\
&A_L + (1-x)u_L + xB_{LL} \geq 2u_L \\
&B_{LL} \geq u_L \\
&A_L + (1-x)u_L + xB_{LL} \geq A_L + (1-x)u_L + xB_{LH} \\
&A_L + (1-x)u_L + xB_{LL} \geq \left(1-p'\left(s^{1H}\right)\right)u_L + p'\left(s^{1H}\right)\left(A_H + p'\left(s^{2H}\right)B_{HH}\right) \\
&A_L + (1-x)u_L + xB_{LL} \geq \left(1-p'\left(s^{1H}\right)\right)u_L + p'\left(s^{1H}\right)(A_H + B_{HL}) \\
&C_H \geq u_H \\
&C_H \geq C_L \\
&C_L \geq u_L \\
&C_L \geq p^{3H}C_H.
\end{aligned} \tag{3}$$

Profits for the firm are

$$V(p) = \lambda\left(2\pi^{1H} - A_H - B_{HH}\right) + (1-\lambda)\left[\pi^{1L} + x\pi^{2L} + (1-x)\left(\lambda\pi^{3H} + (1-\lambda)\pi^{3L}\right) - A_L - xB_{LL} - (1-x)(\lambda C_H + (1-\lambda)C_L)\right].$$

We can, as before, shift payment between A_H and B_{HH} without affecting profit. We set $A_H = 0$ and raise B_{HH} accordingly to loosen the two constraints. We can also set $B_{HL} = B_{LH} = 0$ without affecting profit and loosening constraints: because the contract identifies types and induces optimal behavior it is best to minimize payment to CEOs that behave sub-optimally in one period. We can also set B_{HH} as small as possible to still hire high types. As long as the offer to low types is not too enticing, high types will not deviate so we, for now, ignore that constraint and will check that it is satisfied later. Then $B_{HH} = 2u_H$. We can also let $C_H = u_H$ and $C_L = \max\{u_L, p^{3H}u_H\}$. Decreasing p^{3H} also decreases profits from high types, and therefore will only be decreased if there is sufficient value gained from paying low types less. Therefore, $\max\{u_L, p^{3H}u_H\} = C_L = p^{3H}u_H$. We are left with

$$\begin{aligned}
&A_L + xB_{LL} \geq (1+x)u_L \\
&B_{LL} \geq u_L \\
&A_L + \left(p'\left(s^{1H}\right) - x\right)u_L + xB_{LL} \geq 2p'\left(s^{1H}\right)p'\left(s^{2H}\right)u_H.
\end{aligned}$$

If we decrease A_L by 1 and increase B_{LL} by $\frac{1}{x}$ then profits are left unchanged. One constraint is loosened while the other two remain unchanged. This remains valid as long as $x > 0$, so we assume $x > 0$ and $A_L = 0$. Subsequently,

$$\begin{aligned}
&B_{LL} \geq \frac{(1+x)}{x}u_L \\
&B_{LL} \geq \frac{2p'\left(s^{1H}\right)p'\left(s^{2H}\right)}{x}u_H - \frac{p'\left(s^{1H}\right) - x}{x}u_L.
\end{aligned}$$

If the first constraint binds, we get a value of

$$V = \lambda\left(\pi^{1H} + \pi^{2H} - 2u_H\right) + (1-\lambda)\left[\pi^{1L} + x\pi^{2L} + (1-x)\left(\lambda\pi^{3H} + (1-\lambda)\pi^{3L}\right) - x\frac{(1+x)}{x}u_L - (1-x)\left(\lambda u_H + (1-\lambda)p^{3H}u_H\right)\right],$$

which is linear in x. If the second constraint binds, we get

$$V=\lambda\left(\pi^{1H}+\pi^{2H}-2u_H\right)$$
$$+(1-\lambda)\left[\pi^{1L}+x\pi^{2L}+(1-x)\left(\lambda\pi^{3H}+(1-\lambda)\pi^{3L}\right)-2p'\left(s^{1H}\right)p'\left(s^{2H}\right)u_H+\left(p'\left(s^{1H}\right)-x\right)u_L-(1-x)\left(\lambda u_H+(1-\lambda)p^{3H}u_H\right)\right],$$

which is also linear in x. In either case it cannot be optimal to have $x\in(0,1)$. The final case would be where both bind:

$$\frac{(1+x)}{x}u_L=\frac{2p'\left(s^{1H}\right)p'\left(s^{2H}\right)}{x}u_H-\frac{p'\left(s^{1H}\right)-x}{x}u_L$$

$$\frac{u_H}{u_L}=\frac{\left(1+p'\left(s^{1H}\right)\right)}{2p'\left(s^{1H}\right)p'\left(s^{2H}\right)}.$$

This cannot happen because we have shown that $p'(s^{2H})=\rho$, $p'(s^{1H})<\rho$ and we only need to consider these contracts when $\frac{u_H}{u_L}>\frac{1}{2}\frac{(1+\rho)}{\rho^2}$. Therefore, we must have a corner solution and $x\notin(0,1)$. ■

Plugging in $x=1$ (where low types are never fired) we get a firm value of

$$V_{permissive}=\lambda\left(\pi^{1H}+\pi_H-2u_H\right)+(1-\lambda)\left[2\pi_L-2p'\left(s^{1H}\right)\rho u_H-\left(1-p'\left(s^{1H}\right)\right)u_L\right].$$

As before, the effects of an increase in s in the region of s_H are second order on π and first order on pay for low types:

$$\left.\frac{dV_{permissive}}{ds}\right|_{s=s_H}=\lambda\left.\frac{d\pi}{ds}\right|_{s=s_H}+(1-\lambda)(u_L-2\rho u_H)\left.\frac{dp'(s)}{ds}\right|_{s=s_H}$$
$$\simeq(1-\lambda)(u_L-2\rho u_H)\frac{dp'(s)}{ds}\bigg|s=s_H>0,$$

where the last inequality follows because $\left.\frac{dp'(s)}{ds}\right|_{s=s_H}<0$ and $\frac{u_H}{u_L}>\frac{1}{2}\frac{(1+\rho)}{\rho^2}$. Therefore, as before, empire building by high types is optimal.

Recall that

$$V_{severance}=\lambda\left(\pi^{1H}+\pi_H-2u_H\right)+(1-\lambda)\left[\pi_L-p'\left(s^{1H}\right)(2\rho u_H-u_L)+\lambda\left(\pi^{3H}-u_H\right)+(1-\lambda)\left(\pi_L-p'\left(s^{3H}\right)u_H\right)\right]$$
$$V_{low}=2(\pi_L-u_L).$$

Lemma 7. *Any of the three schemes: only hire low, hire both and fire, and hire and retain both can be optimal.*

Proof. Let $u_L=1$, $u_H=2$, $\pi_L=4$, $\pi_H=6$, $\rho=.75$, $\lambda=.1$, $s^{1H}\cong s^{3H}\cong s_H$. Then we get

$$V_{severance}=5.795$$
$$V_{permissive}=5.75$$
$$V_{low}=6.$$

V_{low} is highest. We know that raising the fraction of high types in the population makes hiring both types and firing low types more valuable. Let all parameters be set as before except $\lambda=.9$.

$$V_{severance}=8.195$$
$$V_{permissive}=7.75$$
$$V_{low}=6.$$

Now let parameters have middling values: $u_L=1.453$, $u_H=4.453$, $\pi_L=6.797$, $\pi_H=10$, $\rho=.5694$, $\lambda=.6906$, $s^{1H}\cong s^{3H}\cong s_H$. Then we get

$$V_{severance}=10.72028$$
$$V_{permissive}=10.780536711$$
$$V_{low}=10.634.$$

■

Proof of Proposition 3. The values of the firm under the three potential contract schemes are

$$V_{severance} = \lambda\left(\pi^{1H} + \pi_H - 2u_H\right) + (1-\lambda)\left[\pi_L - p'\left(s^{1H}\right)(2\rho u_H - u_L) + \lambda\left(\pi^{3H} - u_H\right) + (1-\lambda)\left(\pi_L - p'\left(s^{3H}\right)u_H\right)\right]$$
$$V_{low} = 2(\pi_L - u_L)$$
$$V_{permissive} = \lambda\left(\pi^{1H} + \pi_H - 2u_H\right) + (1-\lambda)\left[\pi^{1L} + \pi_L - 2p'\left(s^{1H}\right)\rho u_H - \left(1 - p'\left(s^{1H}\right)\right)u_L\right].$$

We can use the same superscripts in $V_{severance}$ and $V_{permissive}$ because the first order conditions are the same

$$\frac{dV_{permissive}}{ds^{1H}} = \frac{dV_{severance}}{ds^{1H}} = \lambda\frac{d\pi}{ds^{1H}} + (1-\lambda)(u_L - 2\pi u_H)\frac{dp}{ds^{1H}} = 0.$$

Let $\pi^{1H} = \pi_H - J$, $\pi^{3H} = \pi_H - K$, $p'(s^{1H}) = \rho - p_1$, $p'(s^{3H}) = \rho - p_3$.

	$V'_{severance}$
π_H	$\lambda(3 - \lambda)$
π_L	$(1 - \lambda)(2 - \lambda)$
u_H	$-2\lambda - (1 - \lambda)(2\rho(\rho - p_1) + \lambda + (1 - \lambda)(\rho - p_3))$
u_L	$(1 - \lambda)(\rho - p_1)$
λ	$-J + (3 - 2\lambda)(\pi_H - u_H) - \pi_L + (\rho - p_1)(2\rho u_H - u_L) - (1 - 2\lambda)K$
ρ	$(1 - \lambda)[(\lambda - 4\rho - 1)u_H + u_L]$
	V'_{low}
π_H	0
π_L	2
u_H	0
u_L	-2
λ	0
ρ	0
	$V'_{permissive}$
π_H	2λ
π_L	$2(1 - \lambda)$
u_H	$-2\lambda - 2(1 - \lambda)\rho(\rho - p_1)$
u_L	$-(1 - \lambda)(1 - (\rho - p_1))$
λ	$-J + 2\pi_H - 2(1 - \rho(\rho - p_1))u_H - 2\pi_L + (1 - (\rho - p_1))u_L$
ρ	$(1 - \lambda)((-4\rho + 2p_1)u_H + u_L)$

We can now rank the size of the derivatives with respect to each variable: ■

	Rank
π_H	$V'_{severance} > V'_{permissive} > V'_{low}$
π_L	$V'_{low} > V'_{permissive} > V'_{severance}$
u_H	$V'_{low} > V'_{permissive} > V'_{severance}$
u_L	$V'_{severance} > V'_{permissive} > V'_{low}$
λ	$V'_{severance} > V'_{permissive} > V'_{low}$
ρ	$V'_{low} > V'_{permissive} > V'_{severance}$

The above proof shows that, depending on parameters, any of the three schemes – only hire low, hire both and fire, and hire both and do not fire – can be optimal. The derivatives of the value functions also show that the permissive contract will only be optimal for "middling" values of all parameters. In fact, it is difficult to find parameters for which it is best. Note that in the above three examples, permissiveness is only best by a very small amount if at all. This is not because of some particular choice of parameters: this is, in general, true. Because it is best for middling values and because the size of the set of those values is small, it will never be significantly better than either other type of contract.

In order to state "how often" each contract would be best, we would need some distribution of parameters $u_L, u_H \in \mathbb{R}, \lambda \in (0, 1)$ and some distribution of the function π in the space of functions satisfying out assumptions. We do not do this here. However, just to get some idea of how unusual it is for permissiveness to be optimal, we can assume $s^{1H} \cong s^{3H} \cong s_H$ and therefore parameterize π by stating π_L, π_H and ρ. If we let $\pi_H = 1$, then let ρ and λ be distributed uniformly over $(0, 1)^2$, let π_L be uniform over $(\rho, 1)$, let u_L be uniform over $(0, \pi_L)$ and let u_H be uniform over $(\pi_L, 1)$, Monte Carlo simulations from this distribution yields permissiveness being optimal roughly .48% of the time.

References

Almazan, Andres, Suarez, Javier, 2003. Entrenchment and severance pay in optimal governance structures. J. Finance 63 (2), 519–547.
Baber, William, Fairfield, Patricia, Haggard, James, 1991. The effect of concern about reported income on discretionary spending decisions: the case of research and development. Account. Rev. 66, 818–829.
Bartov, Eli, 1993. The timing of asset sales and earnings manipulation. Account. Rev. 68, 840–855.
Bushee, Brian, 1998. The influence of institutional investors on myopic R&D investment behavior. Account. Rev. 73, 305–333.
Chevalier, Judith, Ellison, Glenn, 1997. Risk taking by mutual funds as a response to incentives. J. Polit. Econ. 105, 1167–1200.
Coles, Jeffrey, Daniel, Naveen, Naveen, Lalitha, 2006. Managerial incentives and risk-taking. J. Financ. Econ. 79, 431–468.
Core, John, Guay, Wayne, 1999. The use of equity grants to manage optimal incentive levels. J. Account. Econ. 28, 151–184.
Core, John, Guay, Wayne, 2002. Estimating the value of employee stock option portfolios and their sensitivites to price and volatility. J. Account. Res. 40, 613–630.
Dechow, Patricia, Sloan, Richard, 1991. Executive incentives and the horizon problem: an empirical investigation. J. Account. Econ. 14, 51–89.
Graham, John, Harvey, Campbell, Rajgopal, Shivaram, 2005. The economic implications of corporate financial reporting. J. Account. Econ. 40, 3–73.
Green, Jeremiah, 2009. Discretion, managerial incentives, and market conditions: the misreporting of hedge fund returns. University of North Carolina Working Paper.
Guay, Wayne, 1999. The sensitivity of CEO wealth to equity risk: an analysis of the magnitude and determinants. J. Financ. Econ. 53, 385–415.
Inderst, Roman, Mueller, Holger M., 2010. CEO Replacement Under Private Information. Review of Financial Studies 23 (8), 2935–2969.
Kahn, Charles M., 1985. Optimal severance pay with incomplete information. J. Polit. Econ. 93 (3), 435–451.
Lambert, Richard A., Larcker, David F., 1985. Golden parachutes, executive decision making, and shareholder wealth. J. Account. Econ. 7, 179–203.
Laux, Volker, 2008. Board independence and CEO turnover. J. Account. Res. 46, 137–171.
Levitt, Steven, Snyder, Christopher, 1997. Is no news bad news? Information transmission and the role of "early warning" in the principal-agent model. RAND J. Econ. 28, 641–661.
Lys, Thomas, Rusticus, Tjomme O., Sletten, Ewa, 2007. Are large CEO severance packages justified? The underlying factors of CEO severance. *Kellogg Insight, Kellogg School of Management*. Retrieved January 10, 2010, from http://insight.kellogg.northwestern.edu/index.php/Kellogg/article/are_large_ceo_severance_packages_justified.
Manso, Gustavo, 2011. Motivating Innovation. J. Finance 66, 1823–1860.
Roychowdhury, Sugata, 2006. Earnings management through real activities manipulation. J. Account. Econ. 42, 335–370.
Scharfstein, David, Stein, Jeremy, 2000. The dark side of internal capital markets: divisional rent-seeking and inefficient investment. J. Finance 55, 2537–2564.
Schleifer, Andrei, Vishny, Robert W., 1989. Managerial entrenchment: the case of manager-specific investments. J. Financ. Econ. 25, 123–139.
Smith, Clifford, Stulz, Rene, 1985. The determinants of firms' hedging policies. J. Financ. Quant. Anal. 20, 391–405.
Walkling, Ralph A., Long, Michael S., 1984. Agency theory, managerial welfare, and takeover bid resistance. RAND J. Econ. 15, 54–68.
Wang, Sean, D'Souza, Julia, 2008. Earnings management: the effect of accounting flexibility on R&D investment choices. Cornell University Working Paper.
Yermack, David, 2006. Golden handshakes: separation pay for retired and dismissed CEOs. J. Account. Econ. 41, 237–256.