Severance Pay, Empire Building and the Prevention of Managerial Shenanigans

Edward Dickersin Van Wesep[1]

UNC Kenan-Flagler Business School

March 2008

[1]The author would like to thank B Douglas Bernheim, Jeffrey Zwiebel, Jonathan Levin, Ilya Segal, Daniel Quint, Nageeb Ali and seminar participants at Stanford University for many helpful comments and insights. Remaining errors are attributable to the author.

Abstract

Large severance payments have become common in executive compensation packages, but there is an apparent disconnect between severance and the idea of "pay for performance." This paper introduces a novel rationale for severance: when executives of hidden quality can attempt to hide their type by taking excessive risk, a contract featuring severance can be optimal. This type of contract will be best at large and complex firms and the specified severance payment can be large both in dollar terms and relative to per-period pay. Optimal contracts also specify that CEOs engage in "empire building," increasing firm size beyond the first-best level. This paper therefore introduces a novel explanation for this widely observed phenomenon.

Nothing offends more than executives being paid for failure

-Eddie Lampert

I Introduction

The attention of the government and the popular and academic press has been increasingly devoted to the behavior of chief executive officers (CEOs) and corporate executive boards. In particular, there is concern that executive pay is too high, bears little relation to shareholder return and is deceptively designed to escape notice of shareholders and watchdogs. Among the compensation elements that have generated the most negative press is the "golden parachute," a fancy and typically pejorative term for executive severance pay. When shareholders experience good returns they tend not to be highly concerned with the size of a CEO's pay but when the CEO has clearly failed to perform to expectations and is therefore being fired, both shareholders and pundits wonder why the departing leader is rewarded with a generous cash bonus. How do we reconcile severance with the concept of "pay for performance?"

At the same time there have been a number of high profile corporate meltdowns in which managers took increasingly risky actions in the hope of overcoming as yet unreported losses. Knowing that they would be fired once their poor performance was uncovered, the personal loss from risk-taking faced by the managers was small relative to the potential personal gain. Normally managers are viewed as more risk averse than firms and contract design has focused on increasing managers' willingness to take risk. In the case where managers know they cannot meet the expectations of shareholders, however, contracts must be designed to reduce their risk taking. This paper describes an optimal contract that aligns manager risk preferences with those of the firm using severance pay and analyzes its properties.

The stylized facts surrounding severance clauses are roughly as follows. Severance clauses in executive contracts have been increasing over time, both in frequency and in size. At the largest US firms severance payments (when an executive is fired for cause) have grown to roughly 2-3 times annual pay[1], a larger ratio than at other firms in the past or present. Further, severance payments are typically not indexed to any performance measures: only in case of fraud or gross negligence does a board avoid paying severance. Finally, large retirement payouts, often called severance but not defined as such in this paper, have also become common. As we will see, the predictions arising from this model are surprisingly consistent with these stylized facts and provide a new caveat for boards and compensation committees.

There have been several strands of literature focusing on severance as a component of an optimal compensation scheme. When profits are noisy and firing of high quality executives takes place on the equilibrium path, severance payments might allow for efficient risk sharing (e.g. Kahn (1985)). Along similar lines, Manso (working paper, 2005) 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) argue that severance can also solve a commitment problem by allowing boards to only fire executives when a replacement is significantly better.[2] Severance payments can compensate managers for loss of a job in the case that their firm is acquired (Walkling and Long (1984) and Lambert and Larcker (1985), for example). Another line of research argues that severance

[1] In fact, 3 times pay is enshrined in US tax code IRC 280G and 4999: "Golden Parachute Provisions." When a severance payment exceeds triple base income and is contingent upon acquisition, executives must pay a 20% excise tax (on top of other taxes) on pay above base income and the company cannot deduct severance from its taxable income.

Schwab and Thomas (2004) find that typical severance for CEOs is 2 years salary.

[2] 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.

may result from non-arms-length bargaining where the manager is entrenched (Shleifer and Vishny (1989) or Scharfstein and Stein (2000)). Managers in this literature are seen as having captive boards and are able to effectively "write their own paychecks."

The paper making the argument that is most closely related to the one presented here is Inderst and Mueller (2005) which reasons that when CEOs can withhold information from boards as to the quality of their match with the firm (or when management can take costly actions to entrench itself), severance is a necessary part of any optimal compensation scheme. This argument is technically similar to that in this paper, though its basis and some empirical implications differ. Pay highly linked to performance and an appropriate choice of severance size can induce weak matches to leave the firm but strong matches to remain.

In this paper I offer a novel reason for severance. When board expectations are high and a CEO is not able to meet those expectations in spite of her best efforts, she has an incentive to take excessive risk. In this setting, a board is sometimes best off by lowering expectations but it is often preferable to offer severance pay to induce bad managers to simply "do their best" and be fired for it.

I model this situation with a board and a unit mass of managers, each living for two periods. Managers may be either good or bad and the board is unable to perfectly distinguish the two. The board offers a contract and, upon accepting, the agent chooses firm size and risk level. After profits are observed the board can either fire or retain the CEO. If fired a new CEO is chosen and given a new contract. In the second period firm size and risk are again chosen and profits realized.

The board would ideally design a contract that would attract only one type of manager and would decrease the value of that contract until the desired manager is indifferent between accepting and not accepting the contract. A wage and stock scheme that only is worth enough to attract bad managers would be one example. A compensation function that is highly dependant on performance would be another (it would encourage good behavior and, so long as the relationship between pay and performance is strong enough, would only attract good managers).[3]

In practice it is not always possible to design contracts this way. One common problem is limited liability: there is a lower bound to how negative pay can get so there is a limit to how sharp the relationship between pay and performance can be. With limited liability it is often the case that a board wishing to hire a good manager cannot design a pay package that screens bad managers. There are then three potential routes for the board: only hire bad managers, hire both types of manager and never fire for good behavior, or hire both types, fire bad managers and provide severance pay for good behavior. The precise definition of "good behavior" will become clear below, but for now it suffices to say that "good behavior" means taking actions specified by the board.

The third option is often the best: *Severance may be the least expensive way to induce good behavior from bad managers.* Perhaps surprisingly it is typically less expensive than promising job security once the opportunity cost of having a low quality leader is taken into account. Packages of this sort are most valuable when the additional expected profit from having a good manager is high, when ability of the board to screen candidates is high and when the quality difference between good and bad managers is high. The size of the optimal severance payment can also be large: I show that the ratio of severance size to per-period

[3] By making pay highly dependant on performance, low types would receive less pay than high types (they, on average, do not perform as well). The level of dependancy can yield any expected pay for low types given some expected pay for high types.

pay for successful management is not useful for determining whether a payment is too large. I also show that the dollar values of severance that we see in practice can be rationalized and are not necessarily a result of management entrenchment (cf. Bebchuk and Fried (2004)).

Importantly, the optimal severance contract pays fired CEOs who achieve mediocre profits *but not those who fail completely.* Appropriate compensation for mediocrity can provide an incentive for weak managers to "do their best" while pay for poor performance provides incentives for undue risk taking. Many packages we see today do not have this rule and reward failures just as they do a more middling performance. While most contracts rule out payments after fraud, more general failure is not typically penalized. Henry Yuen, for example, was paid $23MM in severance for overseeing a collapse in shareholder value at Gemstar. I would argue that outcomes like this are a result of sub-optimal contracting. I argue in this paper that severance may be necessary to attract good CEO to the firm, the design of the severance scheme must not reward poor performance. Though this model is mostly intended as descriptive, this deviation of prediction from reality can be seen as prescriptive.

It is important to understand how the predictions of this model align with reality. I do not attempt here to argue that this model is the only reason for severance payments in executive contracts: it is clear, for example, that there is a role for severance in case of being acquired quite separate from this model. I hope instead to show that, whereas in previous literature severance was seen as sometimes necessary to *in*duce risk-taking, it can also be necessary to *re*duce risk taking.

The purpose of this paper is to show that severance pay may be useful to assure good behavior from low quality CEOs when they can take detrimental actions to attempt to hide their types. Section II outlines the simple model underlying the discussion. Section III derives the set of possible optimal contracts and shows how parameters affect which out of this set is best. I show that severance is best when good managers are easier to identify and are much more valuable to the firm (net of reservation wages). Section IV highlights that when this model is extended to include more than two periods, the value of the severance package can be arbitrarily large relative to per-period compensation. Section V concludes. Note that proofs not offered in the main paper can be found in the appendix.

II The Model

There is a firm and an infinite set of potential CEOs that each live for two periods.[4] CEOs are of type $\theta \in \{\theta_L, \theta_H\}$. Agents know their type but it is hidden from others. In period one, the board nominates an agent randomly from the pool and makes a take-it-or-leave-it contract offer covering both periods of potential employment which the nominee can accept or decline. If the offer is declined then the board nominates a new candidate. This continues until a candidate accepts the offer of employment. The board has some power to screen potential candidates, but is unable to do so perfectly. With probability λ the board accidentally nominates a low quality manager when it attempts to nominate a high quality manager.

The new CEO then chooses firm size s and downside risk p, but only size s is observed by the board. This can be interpreted as follows: size, whether defined as revenue, employment, assets or any other standard measure, is easily observed by a board. It is easy to observe the breadth of the firm's businesses or the industries in which it operates. It is, however, very difficult to discern what risks are inherent in each business. Two firms that are very similar in

[4] Allowing the firm to be infinitely lived does not affect the qualitative results here.

outside appearances may face very different risks that are only observable to someone with deep knowledge of the inner workings of the business.[5] Nature chooses firm profits as either $\pi(s,p,\theta)$ (described below) with probability p or 0 with probability $1-p$ and these profits are observed by the board. The board now chooses whether to fire the CEO or retain her. If it decides to retain her, she may decide to stay or quit. In period 2, if the CEO was fired or she quit then a new one is chosen and the game continues as before with the new CEO choosing s and p etc. If the CEO was not fired and did not quit then she chooses s, p etc.

We assume that profits in case of a good outcome are decreasing in p, but that expected profits are increasing. Risk in this model is not desirable: unlike other models in which severance contracts can be optimal, risk here 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.

$$\frac{\partial \pi(s,p,\theta)}{\partial p} < 0, \; and \; \frac{\partial}{\partial p} p\pi(s,p,\theta) > 0, \;\; \forall s,p,\theta$$

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

$$\pi(s,p,\theta_H) > \pi(s,p,\theta_L), \;\; \forall s,p$$

[5] As an example, consider Capital One Financial and Providian Financial. In 2001 both firms operated in very similar areas of the sub-prime credit card industry, both had identical *professed* approaches to building risk and marketing models and both had achieved astronomical growth in the past five to ten years. However, Providian was shown in 2001 to have far inferior credit models as it failed in the face of rising charge-offs. Capital One's stock initially collapsed due to fear of imminent similar failure. However, Capital One's models were later seen to be high quality and charge-offs did not increase enough to reduce profitability. The risk inherent to the businesses was very different but even members of a corporate board would have been unlikely to know that.

Assume also that for any choice of risk there is an optimal firm size for each type of CEO with higher quality CEOs associated with larger optimal firm sizes. Let the size that maximizes expected profits be finite and equal to $s_i^*(p) = \arg\max_s \pi(s, p, \theta_i)$ and let π be concave in s.

$$\begin{aligned} \frac{\partial^2 \pi(s,p,\theta)}{\partial s^2} &< 0, \ \forall s, p, \theta \\ s_H^*(p) &> s_L^*(p), \ \forall p \end{aligned}$$

Finally, we must ensure that the bad manager can always attempt to imitate the good manager. Specifically, for any s, p, π combination that the good manager is assigned, there is a choice of $p' > 0$ for which the bad manager can also achieve s, π with probability p'.

$$\forall s, p \ \exists p' \ s.t. \, \pi(s, p', \theta_L) = \pi(s, p, \theta_H)$$

Payoffs for the board are assumed to be expected firm profits and payoffs for the CEOs are expected lifetime utility. There is no discounting, and the reservation wage for type i is u_i in each period with $u_H > u_L$. Agents know their types.

This form of game is unusual and it is worth discussing why this form is chosen rather than more traditional payoff formats. Any model that can capture the story described in the introduction must allow managers to choose a level of risk and the level of risk must be associated with a lower expected profit. Risk by definition requires uncertain payoffs, and a contract specifies a payoff for the manager for any potential outcome. By restricting the analysis to two potential outcomes, we are left with freedom to design a general contract that can be analytically derived in a simple closed form. Other modelling techniques- for example allowing profit $\pi\tilde{}F(s;p)$ where p is the risk choice by the manager, and letting $F(s;p_1)$ be a less-than-mean preserving spread of $F(s,p_2)$ when $p_1 < p_2$

would be a natural alternative- require either a restriction on contract form or numerical methods. In either case the result might seem suspect. While having risk be binary is itself a large restriction, the intuition of these results clearly applies to more complex alternative models.

III Optimal Compensation Schemes

We can now solve for the optimal contract design. The only two constraints that we will require of these contracts is that they have limited liability and that they be renegotiation-proof. Limited liability means that a board cannot pay employees negative amounts.[6] Renegotiation-proofness in this context is the requirement that after some actions have been taken and some information has been revealed there is no way to renegotiate the contract that is mutually beneficial for the participants.

The necessity of limited liability will become clear immediately below, but renegotiation-proofness has more subtle effects. It makes describing the potential optimal contracts, when the board knows with which type it is dealing, quite simple and it allows a straightforward characterization of when it is possible to hire only good managers (Lemma 3). It precisely defines contractually specified behaviors once types have revealed themselves in period one. This last effect is most important when the number of periods is increased beyond the two specified in the model above. Technical convenience is only one reason to require renegotiation; reality is the other. While firms may be able to commit to contracts for lower level workers because of reputation effects and the frequency with which the contracting game is played, the board's interaction

[6] This is actually a normalization. As long as there is some maximum penalty that is independent of choices and other parameters we can normalize this to zero.

with a CEO is quite different. It sets a long term contract only once, and there is likely to be substantial pressure to renegotiate when it can benefit both parties. Boards may be able to commit themselves using legal contracts (e.g. Almazan and Suarez (2003)), but no sensible legal contract could prevent two parties updating an ex post suboptimal contract.

For the rest of the paper, some additional notation should simplify the exposition. Let $s_i = s_i^*(1)$ be the optimal firm size for a manager of type i taking the optimal level of risk (none). Let ρ be the probability of success for a bad manager imitating a good manager by attempting the firm-optimal high level of size and profit: $\pi(s_H, \rho, \theta_L) = \pi(s_H, 1, \theta_H)$. Let the profit from firm-optimal play be $\pi_i = \pi(s_i, 1, \theta_i)$. For the rest of the paper, "firm-optimal" refers to taking no risk and playing $s = s_i$.

If the board could perfectly observe the ability of potential hires, the first-best contract would be quite simple.

Lemma 1 *Without limited liability the optimal contract would only attract one type of agent*

Proof. Consider the following 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 $-M$ otherwise where M is arbitrarily large

The first contract would only attract bad managers and firm-optimal bad manager behavior would occur on the equilibrium path. The second would attract only good managers (for some M 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

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

■

The model would not be particularly interesting if limited liability were not an issue, but in practice limited liability is a serious 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 bad managers," or "trying to hire good managers, but being content with a bad manager if you get stuck with one." For the remainder of the paper I will describe contracts that are the best given a particular goal as "constrained optimal" whereas those that are best out of all potential contracts will be called "globally optimal."

Lemma 2 *A constrained optimal contract that only hires bad managers is a flat wage and no firing in both periods*

Proof. 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 good managers because maximum pay is $w = u_L < u_H$. ■

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 the above inequality does not hold, (i.e. $\pi_L - u_L < \pi_H - u_H$) lemma 1 states that the board would prefer to only hire good managers.

Lemma 3 *A constrained optimal renegotiation-proof contract in period one that only hires good managers features*

1. *Pay of* $2u_H$ *that is deferred until retirement when* $s = s_H$ *and* $\pi = \pi_H$ *is observed in both periods*

2. *Firing and no pay when anything else is observed*

The goal of only hiring good managers is only achievable if $\frac{u_H}{u_L} \leq \frac{1}{2}\frac{(1+\rho)}{\rho^2}$

Proof. Suppose we have some contract that pays $A(s_1, \pi_1)$ in the first period and $B(s_1, s_2, \pi_1, \pi_2)$ in the second if the CEO from the first period is not fired. Since the goal is to only hire good managers and we require renegotiation-proofness, the contract must induce play of s_H, π_H in each period.[7] The simplest way to do that is to maximize the penalty from playing anything else so 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 bad manager as

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

The value of the firm is linear in $A + B$ (only good managers are hired so the total wage bill is $A + B$) so by moving pay from A to B we weaken the constraint. Because of limited liability this constraint is weakest when $A = 0$.

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

Then 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. ■

The contract specified above defers payment until period 2 in order to weaken the bad manager's IR constraint and pays only enough to hire the good manager. 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 (it is difficult for bad managers to imitate good managers). This makes intuitive sense: it is only possible for a firm to hire only good managers if there is a large amount of firm specific skill that a CEO brings.

As was stated above, it is best to target one type if possible. Firms that find it more profitable to hire good managers may or may not be able to, as shown as shown in lemma 3. The value of ρ does not imply anything about whether the firm is served best by hiring low or high, but if the types' outside options are similar then we would expect $\pi_H - u_H > \pi_L - u_L$. Therefore we would expect, in practice, to find that firms using this type of pay scheme would be firms where CEO skill is highly specific to the firm and not general. The value to the firm of this contract is $2(\pi_H - u_H)$.

Now we consider the most interesting situation regarding parameter values:

the one in which the board would like to only hire good managers but cannot because bad managers 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

$$\begin{aligned} \pi_H - u_H &> \pi_L - u_L \\ \frac{u_H}{u_L} &> \frac{1}{2}\frac{1+\rho}{\rho^2} \end{aligned}$$

The board must specify actions for each type and payoffs for success and failure following those actions as well as payoffs for actions not specified in the contract. There are two potentially optimal contracts that attract good managers: one in which bad managers are fired when their actions reveal them and one in which they are allowed to stay on. It is shown in the appendix that it cannot be optimal to use other schemes, but either of these can be optimal depending on parameter values. I will later discuss when each is best and show that it is "more likely"[8] that the former be better. For now we assume that if a firm wishes to dissuade bad managers from imitating good managers through requiring good managers to choose excessive size or take excessive risk, it is less expensive to require excessive size.[9]

Prior to stating the results of this model some definitions are in order. Let $p(s)$ be the risk taken by bad managers who imitate good managers by choosing s and $\pi(s, 1, \theta_H)$. i.e. let $p(s)$ be implicitly defined by $\pi(s, 1, \theta_H) = \pi(s, p(s), \theta_L)$.

[8] To actually discuss liklihood one needs some joint distribution of the parameters which is never specified here.

[9] The basic intuition for this is that any constrained optimal contract that specifies different behavior for the two types must specify firm optimal play for the bad manager: bad managers should be told to take no risk. In this case, in the neighborhood of firm-optimal play for good managers a small increase in specified good manager risk will result in only good managers earning profits of 0 on the equilibrium path. Firing for failure is no longer renegotiation-proof so the incentive for bad managers to deviate to taking some risk is significantly loosened: If they fail they are seen as good managers and are not fired! A small strengthening in the bad manager IC by making it more difficult to imitate good managers would actually result in a significant weakening in said IC making requiring a small level of risk for good managers bad for the firm. A small increase in size, as we will see, is always good.

For the remainder of the paper, define "empire building" as a choice of s, p by a CEO that is larger than optimal: i.e. $s > s_i^*(p)$.[10]

The following two terms will define optimal empire building to be described in proposition 1. Let s^{1H} and s^{3H} be defined by

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

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

Note that $\frac{\partial \pi(s,1,\theta_H)}{\partial s}|_{s=s_H} = 0$ (π is concave and maximized at $s = s_H$) but $\frac{\partial p(s)}{\partial s}|_{s=s_H} < 0$.[11] As s rises, the LHSs of equations 1 and 2 become increasingly negative while the RHSs do not. $p(s)$ is bounded below by 0 so there must exist some point at which equality holds for both equations. The smallest such point is the optimal firm size in each case. Because $u_H > 2\rho u_H - u_L$ we are left with

$$s_H < s^{3H} < s^{1H}$$

Proposition 1 *A constrained optimal renegotiation-proof contract that hires both types and fires bad managers after period one:*

1. *Separates types and, in period one, induces firm-optimal behavior for bad managers and empire building ($s^{1H} > s_H$) for good managers*

2. *Fires bad managers and offers severance of $p(s^{1H})(2\rho u_H - u_L)$ upon firing*

3. *Retains good managers, requires firm-optimal play in period two and makes no first period payment. After period two good managers are paid $2u_H$*

[10] 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 bob safety, perks, additional income etc. In this paper, the manager will be require to "empire build" by the board. Although the size will not be first best, it is still optimal for the firm within the model.

[11] The effect of an increase in s is second order on π but first order on $p(s)$.

for taking their required actions

4. *Offers pay of u_H for $s = s^{3H}, \pi = \pi^{3H}$ and $p(s^{3H})u_H$ for $s = s_L, \pi = \pi_L$ to CEOs hired in period two.*

Young high quality managers must choose a very large firm size to prove that their mettle while new managers in the second period need only choose a moderately large firm size. In both cases, empire building takes place on the equilibrium path. Note that there is no risk taken on the equilibrium path so any outcomes other than those specified are punished with firing and no pay.

This contract is fairly simple. In the first period, both types are hired and bad managers are induced to play firm-optimally while good managers are required to empire-build. good managers are kept on, induced to play firm-optimally in period two and have pay deferred until the end of period two. They are paid the minimum possible: $2u_H$. Bad managers are fired and offered severance of $p(s^{1H})(2\rho u_H - u_L)$.

As the bad manager's outside options improve, the amount of necessary severance decreases because bad managers are more willing to leave the job for greener pastures. As bad managers are less able to imitate good managers ($p(t)$ decreases for any given t) severance also decreases because their severance depends on their ability to gamble on a high outcome. As good type reservation utility increases, severance increases: high quality mangers must be paid more to meet their reservation utility, but as this amount grows severance must grow apace to induce poor managers to accept their own firing rather than taking risk. How much empire building is required depends upon the precise shape of π. If good managers can increase size without too much of an effect on profit while the same is not true for bad managers, then more empire building is called for.

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

Proposition 2 *A constrained optimal renegotiation-proof contract that hires both types and fires neither after period one:*

1. *Separates types and, in period one, induces firm-optimal behavior for bad managers and empire building ($s^{1H} > s_H$) for good managers*

2. *Induces firm-optimal behavior for both types in period two*

3. *Defers pay for both types to the end of period two and pays $2u_H$ to good managers and $2p(s^{1H})\rho u_H + (1 - p(s^{1H}))u_L$ to bad managers for taking their specified actions.*

This contract is also fairly simple. Empire building is required for good managers 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 bad manager pay are similar to before except that an increase in bad manager outside options actually makes this contract less valuable. bad manager pay is increasing in outside options because it makes being fired (and therefore imitating good managers) more attractive. As before, there is not uncertainty on the equilibrium path so outcomes not specified in the contract are punished with firing and no pay.

There are other types of potential contract one could suspect might be optimal, but in fact those suspicions would be incorrect (see the appendix for proofs rejecting other contracts). Therefore, if parameters are such that the firm would ideally hire only good managers but is unable to, there are three constrained optimal contracts. If the board offers severance then the firm value is

$$\begin{aligned} V_{seperate} &= \lambda(\pi^{1H} + \pi_H - 2u_H) + (1-\lambda)[\pi_L - p(s^{1H})(2\rho u_H - u_L) \\ &+\lambda(\pi^{3H} - u_H) + (1-\lambda)(\pi_L - p(s^{3H})u_H)] \end{aligned}$$

If the board only hires bad managers the firm value is

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

and if the board hires both types but fires neither, the value of the firm is

$$\begin{aligned} V_{permissive} &= \lambda(\pi^{1H} + \pi_H - 2u_H) + (1-\lambda)[\pi^{1L} + \pi_L \\ &-2p(s^{1H})\rho u_H - (1 - p(s^{1H}))u_L] \end{aligned}$$

For the remainder this paper (minus the appendix), I will focus on the case where it is not optimal to be permissive by allowing bad managers to remain employed at the firm. In the appendix there is a more general treatment, but there is little gained from this. It becomes clear that it is "unusual" for permissiveness to be optimal, in the sense that it is difficult to find parameters that make it so. I investigate under what conditions and how often it can be optimal but do not report results here. Nonetheless, the following proposition is correct and the following discussion informative even when permissiveness is considered.

Proposition 3 *The contract with severance is globally optimal when good managers are much more valuable than bad managers for the firm and when the board* *is more able to screen candidates*

To see the intuition for this proposition, let $\pi^{1H} = \pi_H - J, \pi^{3H} = \pi_H - K, p(s^{1H}) = \rho - p_1, p(s^{3H}) = \rho - p_3$. 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 bad managers) depending on parameters, but there do exist situations when either is higher than the other (see the proof of Lemma 9 in the appendix 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 low contract. To determine how parameter values determine the features of the optimal contract we can take a derivative of $\overline{V}$ with respect to each parameter and check the sign.

Proposition 4 *The contract featuring severance is globally optimal when CEO quality is more important and when it is easier for the board to screen candidates*

Proof. Let $u_L = 1, u_H = 2, \pi_L = 4, \pi_H = 6, \rho = .75, \lambda = .9, s^{1H} \cong s^{3H} \cong s_H$. The firm value from each of the three constrained optimal contracts is:

$$\begin{aligned} V_{seperate} &= 8.195 \\ V_{permissive} &= 7.75 \\ V_{low} &= 6 \end{aligned}$$

Therefore there exist parameter values for which the severance contract is globally optimal. To see how parameters affect its optimality consider the following table which shows derivatives of $\overline{V}$ with respect to each of the parameters in the model:

z	$\frac{d}{dz}\overline{V}$	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 which, in English, is when CEO quality is important (in particular more important to the firm than to outsiders), poor managers cannot easily imitate good managers, and the board is better able to screen candidates. ■

Hiring both types but firing bad managers when they reveal themselves entails a cost: severance must be paid to entice bad managers to reveal themselves and this can be very high. The firm might get a good manager CEO in period two upon firing a bad manager 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.

The curvature of π is parameterized by J, K, p_1 and p_2. Without evaluating derivatives, it is clear that as J and K shrink (the profit function for good managers is flatter) and p_1 and p_2 rise (the profit function for bad managers steepens) the severance contract becomes more valuable. A flatter profit function for good managers and a steeper one for bad managers 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. As a firm grows, the difference $\pi_H - \pi_L$ would grow while $u_H - u_L$ would not change (outside options are not influenced by the size of a particular firm). Therefore parameters would move to make the severance contract more valuable. Furthermore, as the complexity of a firm increases, one suspects that the importance of high quality leadership increases. In this case ρ would decrease and the relative steepness of the profit function would change to favor good managers. Both effects make the severance contract more valuable.

In this section I attempted to show that the problem of detrimental risk-taking can be solved through sophisticated compensation schemes. So long as parameters are forgiving, the firm can either pay a flat salary and attract only bad managers or offer a deferred performance bonus to attract only good managers. Each of these contracts is but one of many that satisfy the necessary ICs and IR constraints but the intuition is clear: Attracting only bad managers does not require pay for performance while attracting only good managers requires (potentially deferred) pay for performance.

If parameters are not forgiving (i.e. if the firm would prefer to screen in favor of good managers but is unable to do so), then the firm has three options: only hire bad managers, hire both and offer severance as an inducement for bad managers to leave or hire both and allow both to stay. Regardless of whether the firm fires bad managers, it is always optimal to induce good managers to engage in empire building to prove their worth.

I have specified, then, conditions under which severance is optimal and these conditions conform well to intuition. The introduction to this paper suggested that it would discuss golden parachutes, a pejorative term for severance that usually implies that the value of the payment is unjustifiably large. We now turn to the question of whether the severance suggested in this model could plausibly be called a golden parachute.

IV The size of optimal severance packages

In the model above I show that under certain conditions on parameters, if the firm wishes to hire good managers as CEO the optimal contract requires offering severance. Severance is a standard contract component for many exempt employees of a typical firm and, far from receiving negative press, is usually seen as necessary in a "fair" contract. Firms may offer job placement counselling, retirement bonuses or simply cash as a way to soften the blow of layoffs or more general firing. Many unions demand severance as a component in their collective bargaining agreements and some countries mandate severance for all workers.[12]

Less popular is the type of massive severance package offered to executives who have driven their firms into the ground, often significantly reducing the net worth of current (or soon to be former) employees of their firm. These severance packages are often referred to as "golden parachutes" and are usually seen as a drain on society rather than the other way around.

In the preceding model I derived the optimal contract and showed that severance may be a part of it. The optimal contract clearly identified when severance should accompany firing: complete failure of a CEO (profit of zero) should not be rewarded with severance, but should be punished with the maximum possible force. However, an inability of the CEO to achieve top notch results, while correctly resulting in firing, should be accompanied by severance to keep the CEO from taking undue risks in an effort to keep her job. In this section I answer the question of how big a severance package is implied by this model.

For simplicity, for the rest of the paper we can reduce the available choices of the types of CEO to $s \in \{s_L, s_H\}, p \in \{0, \rho\}$. This will allow us to avoid

[12] Australia, for example, requires severance for any worker who is part of a layoff of over 15 people. The size of the severance is court mandated and dependant on tenure.

questions about board beliefs when off-path actions are observed. Setting aside the question of whether executive pay is, in general, too high, we can ask the question of whether severance is too high relative to pay: are the severance payments we see too large to be optimal? I interpret the question as follows: how large is the severance payment relative to per-period pay for CEOs who are not fired. This ratio is

$$\begin{aligned}\frac{A_L}{\frac{1}{2}B_{HH}} &= \frac{2\rho^2 u_H - \rho u_L}{u_H} \\ &= 2\rho^2 - \rho\frac{u_L}{u_H}\end{aligned}$$

Because bad managers cannot be excluded from the joining the firm, we know $\frac{u_H}{u_L} > \frac{1}{2}\frac{(1+\rho)}{\rho^2}$. Then

$$\frac{A_L}{\frac{1}{2}B_{HH}} > \frac{2\rho^2}{1+\rho}$$

which can be arbitrarily close to unity. Perhaps surprisingly, pay in case of mediocrity and firing can be almost as large as pay in case of success and retention.

In our model the number of periods in which a CEO lives is set at two. Until now this has not hindered us in answering questions of interest. Here, the number of periods has a significant effect on just how high severance pay can be.

Consider the following model: firms are infinitely lived and CEOs live for $N \geq 2$ periods. After a CEO is fired or retires, the firm nominates a new CEO who is in period one of her life. Let other elements of the earlier model be the same. Let the discount rate for the firm and for agents be $\beta \in (0, 1)$.

Lemma 1 holds as before and the board would like to hire either only low or only good managers. As before, it may be the case that the firm would only like to hire good managers but is unable to because the bad manager reservation utility is too low.

Lemma 4 *The board wishes to hire only good managers but is unable to do so when*

$$\begin{aligned} \pi_H - \pi_L &> u_H - u_L \\ \frac{u_H}{u_L} &> \frac{1-(1-\rho)\sum_{i=1}^{N}\frac{(\beta^i-\beta^N)}{(1-\beta^N)}\rho^{i-1}}{\rho^N} \end{aligned}$$

The set implied by these inequalities is always non-empty

Intuitively, so long as bad manager outside options are small enough we should expect the firm to be unable to use contract terms to screen them. If $u_L = 0$ so that bad managers have no outside options, there will clearly be no way to keep them from joining the firm. By assumption there is always some probability they could imitate good managers for N periods and that probability multiplied by a positive number is always greater than zero.

I will not go through the process, as I did in the main model, of showing that often a globally optimal contract that hires good managers when bad managers cannot be weeded out can include severance. I also state without proof that the optimal contract with severance severs immediately after period one rather than at a later time. For brevity I simply calculate the optimal contract with severance and evaluate its size. Let $\gamma(N) = \frac{\beta^0-\beta^{N-1}}{1-\beta}$.

Lemma 5 *The optimal severance contract features severance of* $\gamma(N)(\rho^N u_H - u_L) + u_L\left[1+(1-\rho)\sum_{i=1}^{N-1}\frac{\beta^i-\beta^N}{1-\beta}\rho^{i-1}\right]$

Proof. Clearly the optimal contract has the board paying as little as possible to get severance. Therefore the utility to a bad manager from accepting the severance package equals the utility from imitating

$$\begin{aligned} A_L + \frac{\beta - \beta^{N-1}}{1-\beta} u_L &= \rho^N \gamma(N) u_H \\ &+ (1-\rho) u_L \sum_{i=1}^{N-1} \frac{\beta^i - \beta^{N-1}}{1-\beta} \rho^{i-1} \\ so \;\; A_L &= \gamma(N)(\rho^N u_H - u_L) \\ &+ u_L \left[1 + (1-\rho) \sum_{i=1}^{N-1} \frac{\beta^i - \beta^{N-1}}{1-\beta} \rho^{i-1} \right] \end{aligned}$$

■

In this section I am not assuming that CEOs are really "living longer;" rather, I am assuming that their lives can be partitioned into more sections. Therefore, let lifetime utility be fixed, regardless of N and let the probability of success in imitating good managers be fixed:

$$\begin{aligned} \gamma(N) u_j &= \overline{u_j}, j \in \{L, H\} \\ \rho^N &= C \end{aligned}$$

Then we get

$$A_L = C\overline{u_H} - \overline{u_L} + \gamma(N)\overline{u_L} \left[1 + (1 - C^{\frac{1}{N}}) \sum_{i=1}^{N-1} \frac{\beta^i - \beta^{N-1}}{1-\beta} C^{\frac{1}{N}(i-1)} \right]$$

It may be difficult to interpret these numbers, so consider an example

Example 1 *Let* $u_L = \$2,000,000, u_H = \$6,000,000, C = \frac{9}{16}, \beta = .95.$ *These are per year income numbers. Let the number of years for a typical CEO be*

20. Then $\overline{u_H} = \$38,490,800, \overline{u_L} = \$12,830,267$

$$\begin{aligned} A_L &= C\overline{u_H} - \overline{u_L} + \gamma(N)\overline{u_L}\left[1 + (1 - C^{\frac{1}{N}})\sum_{i=1}^{N-1}\frac{\beta^i - \beta^N}{1-\beta}C^{\frac{1}{N}(i-1)}\right] \\ &\approx \$14,000,000 \end{aligned}$$

CEOs who fail in the first year to achieve high profits are fired and given a severance payment worth 2 1/3 times annual pay whereas those who succeed are not immediately paid anything. Even when they are paid, the per annum pay is less than half of the severance. These numbers were not chosen to maximize the value of the severance; they were similar (but scaled) to the ones chosen in the proof of proposition 3, normalized to the typical compensation level of a Fortune 500 CEO in 2005. A severance of almost $\$14,000,000$ *would be considered by many to be a "golden parachute" given that it is awarded for one year of mediocre work. In fact it is greater than the present value of bad manager lifetime earnings. One could easily get more extreme numbers by a more focused choice of parameter values.*

Finally we return to the question of how large the ratio of severance to "per-period" pay for success can be. Let $\beta^{N-1} = \overline{\beta}$. That ratio is given by

$$\begin{aligned} \lim_{N\to\infty}\frac{A_L}{\gamma(N)\overline{u_H}} &= \lim_{N\to\infty}\left[C\gamma(N) - \gamma(N)\frac{\overline{u_L}}{\overline{u_H}}\right] \\ &\quad + \lim_{N\to\infty}\frac{\overline{u_L}}{\overline{u_H}}\left[1 + (1 - C^{\frac{1}{N}})\sum_{i=1}^{N-1}\frac{\overline{\beta}^{\frac{1}{N}i} - \overline{\beta}}{1 - \overline{\beta}^{\frac{1}{N}}}C^{\frac{1}{N}(i-1)}\right] \\ &> \lim_{N\to\infty}\left[\gamma(N)(C - \frac{\overline{u_L}}{\overline{u_H}}) + \frac{\overline{u_L}}{\overline{u_H}}\right] \\ &= \left[\lim_{N\to\infty}\gamma(N)\right](C - \frac{\overline{u_L}}{\overline{u_H}}) + \frac{\overline{u_L}}{\overline{u_H}} \end{aligned}$$

Because of the fact that bad managers cannot be excluded if good managers are induced to work for the firm, we know that $C - \frac{\overline{u_L}}{\overline{u_H}} > 0$.

Lemma 6 *The ratio of severance to "per-period" pay for success can be made arbitrarily large by increasing the number of evaluative periods*

Proof. The ratio is

$$\begin{aligned} \lim_{N\to\infty} \frac{A_L}{\gamma(N)\overline{u_H}} &> \left[\lim_{N\to\infty} \gamma(N)\right](C - \frac{\overline{u_L}}{\overline{u_H}}) + \frac{\overline{u_L}}{\overline{u_H}} \\ &= \infty \end{aligned}$$

because $C - \frac{\overline{u_L}}{\overline{u_H}} > 0$ (the firm cannot prevent bad managers from accepting offers). ∎

The amount of severance required to get bad managers to reveal themselves is the difference between lifetime utilities from imitating good managers and taking an outside offer. By assumption, that difference has a positive limit as N increases whereas per-period pay goes to zero. That the ratio of a (positive) stock and a flow goes to zero as the length of evaluative period goes to zero should be no surprise, but there is an important point here. One popular argument against severance packages is that they are so valuable relative to pay for success. In this setting, that ratio can be arbitrarily large in optimal packages.

Hopefully this section has illuminated how the size of severance described in this paper compares to the "golden parachutes" seen in the real world. By increasing the number of periods in which CEOs are evaluated, we see that the value of these severance clauses can be very high, both absolutely and relative to period pay for success.

V Conclusion

In this paper I have introduced a novel justification for severance pay and attempted to analyze how the implications of the story compare to real world stylized facts. As stated in the introduction, the facts appear to be as follows. The incidence and value of severance clauses in executive contracts has been increasing over time. Severance payments for cause are typically 3 times annual pay at Fortune 500 firms though the ratio has been constant neither over time nor across firms. They are not usually indexed to performance except insofar as the executive's actions are criminal. Retirement payouts (deferred pay, early vesting options, pensions etc.), often called severance but not defined as such in this paper, have also become common. The predictions arising from this model are mostly consistent with these stylized facts and provide a new caveat for boards and compensation committees.

We can ask why the frequency of severance packages and the size of the prescribed payments have increased. The model suggests that offering a contract with severance becomes optimal as profits grow relative to reservation wages (where the difference and not the ratio is the relevant measure) and as the return to skill rises. It also suggests that severance payments themselves increase as the return to skill rises (outside options for good managers rise and outside options for bad managers fall). As corporations have grown larger, profits have grown relative to pay. As firms have grown more complex and industry more competitive, the return to CEO skill may have increased as well. The model suggests, in line with observation, that severance payments should have grown in size and frequency.

The discussion of the predicted time series of severance packages highlights a further point: at any time 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 jibe well with casual observation.

The size of the optimal severance payment depends on parameters, but the model implies that it is not necessarily bad for severance payments to be large relative to per-period pay for success. If CEO tenure is typically long then optimal severance payments can be quite large. A typical ratio of 3 is certainly not out of the question.[13] As for whether the dollar values we see are appropriate, plugging "reasonable" numbers into the model can provide an optimal severance payment in excess of the present value of lifetime reservation compensation for a bad manager! When the typical Fortune 500 CEO is paid $6MM per year and fired executives maintain roles on boards worth $2MM per year, the optimal severance payment is $14MM. This model corresponds to the dollar values of severance we often see in the real world.

The 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 an agent's lifetime. Under parameters preventing these contracts, it becomes strictly better for the firm to defer compensation for success as long as possible. This minimizes the probability of success for a bad manager imitating a good manager and therefore lowers the expected wage cost. In the real world, pay deferral, while not complete, is still extensive. Lee Raymond recently netted $400MM upon retirement from Exxon Mobile, which greatly exceeded his annual 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. That pay should be deferred is, to be clear, not a prediction exclusive to this model. Similar agency problems to the one described here will always have some degree of pay deferral (or require executives to make large initial payments to the firm to bond

[13] Carly Fiorina was awarded severance of $21MM upon her departure from HP in 2005 after 6 years. This was roughly 2.5 times her annual pay over that span.

themselves).

There is one major area in which the model does not fit reality. I argue here that severance pay can be optimal, but it must depend on performance. High profits are rewarded with retention and medium profits are rewarded with firing and severance, but low profits must be met with firing and no payment. The reason is straightforward: mediocrity results from weak managers doing their best whereas failure results from unprofitable risk taking, just the activity severance payments are designed to discourage. Most actual severance agreements only depend on performance insofar as they do not pay out in case of negligence or fraud. This paper provides the following recommendation to corporate boards: while pay and performance should not necessarily be perfectly linked, complete failure should not be rewarded.

VI Appendix

Lemma 7 *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 $s' \geq s_H^*(p')$. This follows for the same reason as before: if the board requires pooling at some smaller size then the good managers 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 good managers to choose the larger size and pay a little extra for it. As long as the extra pay is small enough, bad managers would not want to make this argument and therefore the renegotiation would take place. Given that we are restricted within this section 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 the bad managers by offering a contract that induces bad managers 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 bad manager and good manager and induce firm-optimal play. 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 $1 - \rho$ the bad managers fail and are 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 $1 - \rho$ 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 firm optimal and bad managers are indifferent so would be willing to play $s = s_L, \pi = \pi_L$. good managers clearly would still play $s = s_H, \pi = \pi_H$. ■

Proof. of proposition 1 The board proposes a contract specifying period one

high and bad manager actions: $(s = s^{1L}, \pi = \pi^{1L})$ and $(s = s^{1H}, \pi = \pi^{1H})$. It clearly 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 bad managers and induces separation, let 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:

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

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. Then we get

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

Then $A_L = \min\left\{u_L, p(s^{1H})(p(s^{2H})B_{HH} - u_L)\right\}, B_{HH} = \min\left\{2u_H, A_L + u_H\right\}$.

Since it is not possible to only hire good managers we get

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

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

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

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

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

We clearly want to require bad managers 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, $\pi^{iL} = \pi_L, i \in \{1, 3\}$.[14] However, the same cannot be said for good managers. Requiring them to choose sizes

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

larger than firm-optimal reduces the probability with which bad managers 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 good managers have credibly revealed themselves in period one they could otherwise always renegotiate with the board 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 good manager profits could increase while bad manager 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 bad manager pay 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 8 *It cannot be optimal to keep bad managers 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

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

Profits for the firm are

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

We can, as before, shift payment between A_H and B_{HH} without affecting profit. Setting $A_H = 0$ and raising B_{HH} accordingly loosens two constraints so we do that. 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 play sub-optimally in one period. We can also set B_{HH} as small as possible to still hire good managers. So long as the offer to bad managers is not too good, good managers 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 $C_L = \max\left\{u_L, p^{3H}u_H\right\}$. Decreasing p^{3H} also decreases profits from good managers and therefore it will only be decreased if there is sufficient value gained from paying bad managers less. Therefore, $\max\left\{u_L, p^{3H}u_H\right\} = C_L = p^{3H}u_H$. We are left with

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

If we decrease A_L by 1 and increase B_{LL} by $\frac{1}{x}$ then profits are left unchanged. One constraint is loosened and the other two are unchanged. So long as $x > 0$ this works so assume $x > 0$ and $A_L = 0$. We then get

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

If the first constraint binds we get a value of

$$\begin{aligned} V = {} & \lambda(\pi^{1H} + \pi^{2H} - 2u_H) + (1-\lambda)[\pi^{1L} + x\pi^{2L} \\ & +(1-x)(\lambda\pi^{3H} + (1-\lambda)\pi^{3L}) - x\frac{(1+x)}{x}u_L \\ & -(1-x)(\lambda u_H + (1-\lambda)p^{3H}u_H)] \end{aligned}$$

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

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

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:

$$\begin{aligned}
\frac{(1+x)}{x}u_L &= \frac{2p(s^{1H})p(s^{2H})}{x}u_H - \frac{p(s^{1H}) - x}{x}u_L \\
\frac{u_H}{u_L} &= \frac{(1+p(s^{1H}))}{2p(s^{1H})p(s^{2H})}
\end{aligned}$$

This cannot happen because we have shown that $p(s^{2H}) = \rho$, $p(s^{1H}) < \rho$ and we only need to consider these fancy 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 bad managers are never fired) we get a firm value of

$$\begin{aligned}
V_{permissive} &= \lambda(\pi^{1H} + \pi_H - 2u_H) + (1-\lambda)[2\pi_L \\
&\quad -2p(s^{1H})\rho u_H - (1 - p(s^{1H}))u_L]
\end{aligned}$$

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 bad managers:

$$\begin{aligned}
\frac{dV_{permissive}}{ds} \mid_{s=s_H} &= \lambda\frac{d\pi}{ds}\mid_{s=s_H} + (1-\lambda)(u_L - 2\rho u_H)\frac{dp(s)}{ds}\mid_{s=s_H} \\
&\simeq (1-\lambda)(u_L - 2\rho u_H)\frac{dp(s)}{ds}\mid_{s=s_H} \\
&> 0
\end{aligned}$$

where the last inequality follows because $\frac{dp(s)}{ds} \mid_{s=s_H} < 0$ and $\frac{u_H}{u_L} > \frac{1}{2}\frac{(1+\rho)}{\rho^2}$. Therefore, as before, empire building by good managers is optimal.

Recall that

$$\begin{aligned} V_{seperate} &= \lambda(\pi^{1H} + \pi_H - 2u_H) + (1-\lambda)[\pi_L - p(s^{1H})(2\rho u_H - u_L) \\ &\quad + \lambda(\pi^{3H} - u_H) + (1-\lambda)(\pi_L - p(s^{3H})u_H)] \end{aligned}$$

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

Lemma 9 *Any of the three schemes: only hire low, hire both and fire and hire both and retain 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

$$\begin{aligned} V_{seperate} &= 5.795 \\ V_{permissive} &= 5.75 \\ V_{low} &= 6 \end{aligned}$$

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

$$\begin{aligned} V_{seperate} &= 8.195 \\ V_{permissive} &= 7.75 \\ V_{low} &= 6 \end{aligned}$$

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

$$\begin{aligned} V_{seperate} &= 10.72028 \\ V_{permissive} &= 10.780536711 \\ V_{low} &= 10.634 \end{aligned}$$

■

Proof. *of proposition 3* The value of the firm under the three potential contract schemes are

$$\begin{aligned} V_{seperate} &= \lambda(\pi^{1H} + \pi_H - 2u_H) + (1-\lambda)[\pi_L - p(s^{1H})(2\rho u_H - u_L) \\ &\quad + \lambda(\pi^{3H} - u_H) + (1-\lambda)(\pi_L - p(s^{3H})u_H)] \end{aligned}$$

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

and

$$\begin{aligned} V_{permissive} &= \lambda(\pi^{1H} + \pi_H - 2u_H) + (1-\lambda)[\pi^{1L} + \pi_L \\ &\quad - 2p(s^{1H})\rho u_H - (1 - p(s^{1H}))u_L] \end{aligned}$$

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

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

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

	$V'_{seperate}$
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	$V'_{seperate} > V'_{permissive} > V'_{low}$
π_L	$V'_{low} > V'_{permissive} > V'_{seperate}$
u_H	$V'_{low} > V'_{permissive} > V'_{seperate}$
u_L	$V'_{seperate} > V'_{permissive} > V'_{low}$
λ	$V'_{seperate} > V'_{permissive} > V'_{low}$
ρ	$V'_{low} > V'_{permissive} > V'_{seperate}$

$\blacksquare$

The above proof shows that any of the three schemes- only hire low, hire both and fire and hire both and do not fire- can be optimal depending on parameters. 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.

To say "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, and it is not especially productive given the work it would entail. 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. Letting $\pi_H = 1$, 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.

Proof. of Lemma 4 The first condition states that the firm would prefer to only hire good managers. The second shows when this is possible. The minimum good managers must be paid is $\frac{1-\beta^{N-1}}{1-\beta}u_H$ because this is their reservation utility. The way to reduce the desirability, for bad managers, of imitating good managers is to defer pay as long as possible and maximize the risk to bad managers that they will be discovered. Therefore, the best package for deterring bad managers is to offer $\frac{1-\beta^{N-1}}{1-\beta}u_H$ after observing $s = s_H, \pi = \pi_H$ for all N periods and offer no pay plus firing in all other cases (recall that $s \in \{s_H, s_L\}$ for the duration of the paper).

If a bad manager imitates a good manager, then in period one she is not discovered with probability ρ and continues as CEO in period two. She is discovered with probability $1-\rho$, is fired with no pay, and receives her reservation utility of $\frac{\beta^1-\beta^{N-1}}{1-\beta}u_L$. If the passes successfully to period two, she again is discovered with probability $(1-\rho)$ and receives reservation utility $\frac{\beta^2-\beta^{N-1}}{1-\beta}u_L$, etc. Let $\gamma(N) = \frac{\beta^0-\beta^{N-1}}{1-\beta}$. Then her utility from attempting to imitate is

$$\begin{aligned} u_{imitate} &= \rho^N \frac{\beta^0-\beta^{N-1}}{1-\beta}u_H + (1-\rho)\frac{\beta^1-\beta^{N-1}}{1-\beta}u_L + (1-\rho)\rho\frac{\beta^2-\beta^{N-1}}{1-\beta}u_L \\ &\quad +... + (1-\rho)\rho^{N-1}\frac{\beta^{N-1}-\beta^{N-1}}{1-\beta}u_L \\ &= \rho^N\gamma(N)u_H + (1-\rho)u_L\sum_{i=1}^{N-1}\frac{\beta^i-\beta^{N-1}}{1-\beta}\rho^{i-1} \end{aligned}$$

The board is unable to hire only good managers if this utility is higher than bad manager reservation utility

$$\rho^N\gamma(N)u_H + (1-\rho)u_L\sum_{i=1}^{N-1}\frac{\beta^i-\beta^{N-1}}{1-\beta}\rho^{i-1} > \gamma(N)u_L$$

$$\rho^N u_H > u_L\left[1-(1-\rho)\sum_{i=1}^{N-1}\frac{\beta^i-\beta^{N-1}}{1-\beta^N}\rho^{i-1}\right] \tag{4}$$

To complete the derivation, note that if a bad manager has optimally imitated in one period then it is even more in her interest to imitate in future periods. Her probability of getting the "brass ring" of high pay increases with each passing period while the value of revealing herself and leaving the firm decline. ■

VII References

Almazan, Andres and Javier Suarez, 2003, Entrenchment and Severance Pay in Optimal Governance Structures. *Journal of Finance 63 (2),* 519-547

Bebchuk, Lucien and Jesse Fried, 2004, Pay Without Performance. Cambridge, MA: Harvard University Press

Inderst, Roman and Holger M. Mueller, 2005, Keeping the Board in the Dark: CEO Compensation and Entrenchment. *CEPR working paper*

Kahn, Charles M., 1985, Optimal Severance Pay with Incomplete Information. *Journal of Political Economy 93(3),* 435-451

Lambert, Richard A. and David F. Larcker, 1985, Golden Parachutes, Executive Decision Making, and Shareholder Wealth. *Journal of Accounting and Economics, 7,* 179-203

Manso, Gustavo, 2005, Motivating Innovation. *MIT Sloan School of Management working paper*

Scharfstein, David and Jeremy Stein, 2000, The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment, *Journal of Finance, 55,* 2537-2564

Schleifer, Andrei and Robert W. Vishny, 1989, Managerial Entrenchment: The Case of Manager-Specific Investments. *Journal of Financial Economics 25,* 123-139

Schwab, Stewart J. and Randall S. Thomas, 2004, What Do CEOs Bargain For? An Empirical Study of Key Legal Components of CEO Contracts, *Vanderbildt Law and Economics working paper*

Walkling, Ralph A. and Michael S. Long, 1984, Agency Theory, Managerial Welfare, and Takeover Bid Resistance. *Rand Journal of Economics 15,* 54-68

Notes

[1] In fact, 3 times pay is enshrined in US tax code IRC 280G and 4999: "Golden Parachute Provisions." When a severance payment exceeds triple base income and is contingent upon acquisition, executives must pay a 20% excise tax (on top of other taxes) on pay above base income and the company cannot deduct severance from its taxable income.

Schwab and Thomas (2004) find that typical severance for CEOs is 2 years salary.

[2] 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.

[3] By making pay highly dependant on performance, low types would receive less pay than high types (they, on average, do not perform as well). The level of dependancy can yield any expected pay for low types given some expected pay for high types.

[4] Allowing the firm to be infinitely lived does not affect the qualitative results here.

[5] As an example, consider Capital One Financial and Providian Financial. In 2001 both firms operated in very similar areas of the sub-prime credit card industry, both had identical *professed* approaches to building risk and marketing models and both had achieved astronomical growth in the past five to ten years. However, Providian was shown in 2001 to have far inferior credit models as it failed in the face of rising charge-offs. Capital One's stock initially collapsed due to fear of imminent similar failure. However, Capital One's models were

later seen to be high quality and charge-offs did not increase enough to reduce profitability. The risk inherent to the businesses was very different but even members of a corporate board would have been unlikely to know that.

[6] This is actually a normalization. As long as there is some maximum penalty that is independent of choices and other parameters we can normalize this to zero.

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

[8] To actually discuss liklihood one needs some joint distribution of the parameters which is never specified here.

[9] The basic intuition for this is that any constrained optimal contract that specifies different behavior for the two types must specify firm optimal play for the bad manager: bad managers should be told to take no risk. In this case, in the neighborhood of firm-optimal play for good managers a small increase in specified good manager risk will result in only good managers earning profits of 0 on the equilibrium path. Firing for failure is no longer renegotiation-proof so the incentive for bad managers to deviate to taking some risk is significantly loosened: If they fail they are seen as good managers and are not fired! A small strengthening in the bad manager IC by making it more difficult to imitate good managers would actually result in a significant weakening in said IC making requiring a small level of risk for good managers bad for the firm. A small increase in size, as we will see, is always good.

[10] 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 bob safety, perks, additional income etc. In this paper, the manager will be require to "empire build" by the board. Although the size will not be first best, it is still optimal for the firm within the model.

[11] The effect of an increase in s is second order on π but first order on $p(s)$.

[12] Australia, for example, requires severance for any worker who is part of a layoff of over 15 people. The size of the severance is court mandated and dependant on tenure.

[13] Carly Fiorina was awarded severance of $21MM upon her departure from HP in 2005 after 6 years. This was roughly 2.5 times her annual pay over that span.

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