Executive Pay, Hidden Compensation and Managerial Entrenchment

Camelia M. Kuhnen

Jeffrey Zwiebel*

July 2008

Abstract

We consider a "managerial optimal" framework for top executive compensation, where top management sets their own compensation subject to limited entrenchment, instead of the conventional setting where such compensation is set by a board that maximizes firm value. Top management would like to pay themselves as much as possible, but are constrained by the need to ensure sufficient efficiency to avoid a replacement. Shareholders can remove a manager, but only at a cost, and will therefore only do so if the anticipated future value of the manager (given by anticipated future performance net of future compensation) falls short of that of a replacement by this replacement cost. In this setting, observable compensation (salary) and hidden compensation (perks, pet projects, pensions, etc.) serve different roles for management and have different costs, and both are used in equilibrium. We examine the relationship between observable and hidden compensation and other variables in a dynamic model, and derive a number of unique predictions regarding these two types of pay. We then test these implications and find results that generally support the predictions of our model.

---

*Kuhnen: Department of Finance, Kellogg School of Management, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208-2001, c-kuhnen@kellogg.northwestern.edu. Zwiebel: Graduate School of Business, Stanford University, Stanford CA 94305, zwiebel@stanford.edu.

We are grateful to Elazar Berkovich, Xavier Gabaix, Yaniv Grinstein, Ilan Kremer, Mike Lemmon, Paul Oyer, Andy Skrzypacz, Dimitri Vayanos, Jaime Zender and participants at the 2006 NBER Corporate Finance meeting, the 2007 Jackson Hole Winter Finance Conference, the 13th Mitsui Life Symposium, the 2008 Western Finance Association Meeting, London Business School, London School of Economics, and the University of Notre Dame for helpful comments.

1 Introduction

The standard economic analysis of executive compensation begins with a familiar principal-agent problem: the board of directors, acting on shareholders behalf, hires a manager who is paid, as compensation for effort and inferred ability, and as an inducement for good actions. Compensation is chosen in a manner to maximize firm value, subject to informational asymmetries, noncontractible variables, and various constraints on compensation (such as a nonnegativity constraint). As such, executive compensation is envisioned to provide efficient incentives, subject to contracting restrictions.[1]

One question that this standard conception has difficulty addressing, however, is why so much executive pay appears to take a "hidden" form, seemingly disguised from direct shareholder scrutiny. Hidden compensation takes many forms, including managerial perks (Jensen and Meckling (1976), Yermack (2006a), Rajan and Wulf (2006)), lavish pension plans (Bebchuk and Fried (2004), Bebchuk and Jackson (2005)) and generous severance agreements (Yermack (2006b)). Arguably most significant and costly to the firm is the discretion that management sometimes appears to be granted to pursue inefficient pet projects (Jensen and Meckling (1976), Jensen (1986)). Even executive options, while disclosed in footnotes, often seem to be beyond the clear observation and comprehension of many shareholders, and managers have vehemently opposed regulations requiring clearer disclosure. And furthermore, certain common practices with such executive options, such as backdating (Yermack (1997), Heron and Lie (2006)) and implicit agreements to reprice options in the event of a fall in the stock price (Carter and Lynch (2001)), frequently make such options packages far more valuable than what is disclosed.

In many firms, such hidden pay comprises the large majority of compensation for top managers (Bebchuk and Jackson (2005)). Yet it is hard to understand why compensation should be hidden from shareholders if it is simply an outcome of optimal contracting subject to contracting frictions. This is especially true if hidden compensation takes an inefficient form that management would not choose to pursue out of their own pockets if they were compensated directly with cash.

[1]See, for example, Mirrlees (1974), Mirrlees (1976), Holmstrom (1979), Holmstrom (1982), Lazear and Rosen (1981), Grossman and Hart (1983), Holmstrom and Ricart i Costa (1986), Holmstrom and Milgrom (1987), and Holmstrom and Milgrom (1991), for important formative work in the voluminous literature on the principal-agent problem in the firm and its implications for executive compensation. Murphy (1999) and Prendergast (1999) provide excellent surveys of the literatures on executive compensation and incentives in firms.

An alternative conception is that of top managers setting their own pay. This notion is well-motivated by the literatures on the separation of ownership and control, on managerial entrenchment, on the performance of corporate boards, and on managerial empire-building, and is frequently expressed in the popular press.[2] Here it is envisioned that top managers have "captured" the board of directors, and that the board and management act in concert, setting one another's pay, and protecting one another from outside challenges. Indeed, in practice, while board of directors appoints top executives, management at the same time effectively selects new members of the board and sets board members' compensation. Given that the board only faces a small direct cost, if any, for paying management with firm funds, it is argued that boards are only all too happy to accede to managers pay requests, provided that shareholders do not object too strenuously.[3]

Indeed, much of the literature on corporate boards has focused on the notion that they are captured or partially captured by management.[4] In light of this extensive literature, it is all the more surprising that this notion is often readily dismissed in regard to executive compensation, and has had very little expression in formal models of executive compensation.[5]

One challenge for this perspective is the question of what does constrain executive compensation if top managers set it themselves. That is, if managers are free to set their own pay, why do they not pay themselves all of their companies' value?

In this paper, we consider a model of executive pay where the top manager sets his own pay, but is subject to the important constraint of limited entrenchment. Shareholders can remove the manager, but at a cost. The manager, in turn, would like to pay himself as much as possible over time, but is cognizant of the need to remain sufficiently efficient

---

[2]For just a couple of many recent examples in the financial press, see: "Pfizer Chief's Pay Scrutinized" in The Financial Times, April 27, 2006, and "Revolt Looms at Home Depot Over Executive Pay" in The Financial Times, May 18, 2006.

[3]See the book "Pay Without Performance" (Bebchuk and Fried (2004)) for a comprehensive and well-documented argument for this perspective. See also The Journal of Corporation Law (2005) for a collection of papers discussing and critiquing this book.

[4]See for example, Hallock (1997), Hermalin and Weisbach (1998), Main, O'Reilly, and Wade (1995), and Core, Holthausen, and Larcker (1999). Hermalin and Weisbach (2003) provide a survey of the literature on boards of directors.

[5]For example, Murphy (1999) states: "Based on my own observation and extensive discussions with executives, board members, and compensation consultants, I tend to dismiss the cynical scenario of entrenched compensation committees rubber-stamping increasingly lucrative pay programs with a wink and a nod."

relative to a replacement so that shareholders do not choose to remove him. Such a replacement decision depends on three elements: first, the shareholders' inference of the manager's ability relative to a replacement; second, the manager's compensation relative to a replacement; and third, the cost of removing the manager (i.e., the manager's entrenchment).

Building on free cash-flow notions of capital structure (Grossman and Hart (1982), Jensen (1986)), such a partially entrenched "managerial-optimal" perspective for *capital structure* is developed in Zwiebel (1996), Hart (2001), and Novais (2003), and tested by Garvey and Hanka (1999), among others. However, while this perspective has thus received considerable attention applied to the determination of capital structure, it is arguably even more naturally suited to questions of managerial compensation. Our model adopts the approach of Zwiebel (1996), and applies it in a simple manner to a setting where managerial compensation decisions replace capital structure decisions.

A key insight underlying our analysis is that once this perspective is adopted, it immediately suggests different roles that observable and hidden compensation serve for management, and different costs that management incurs from them, and consequently, this perspective yields a ready explanation for why both observable and hidden compensation are employed simultaneously in most firms. This distinction will be central both to the formulation of our model, and to our empirical predictions and tests.

In particular, the cost to a manager of observable pay is simply that if he announces he will pay himself too much relative to a potential replacement, he may be replaced immediately. If, for example, a typical manager (through his board) were to announce a salary of $1 Billion, shareholders would reason that it is worth challenging the manager for control, even given a substantial cost to doing so, and replacing him with a new manager who would not get this pay. Note, on the other hand, that since observable pay is fully anticipated, if the manager is not replaced, such pay will not affect the shareholders' inference of the manager's ability through the observation of the outcome. That is, if shareholders understand that the manager will be paid x dollars in salary, they will correctly anticipate that profits will be reduced by the same amount x.[6]

Hidden compensation, in contrast, is only inferred by shareholders rather than observed directly. Hence, if a manager in a given period takes more hidden compensation

[6]In our setting, there is symmetric information about managerial ability, so managers are not signaling any information with their choice of observable compensation.

than anticipated, shareholders cannot immediately react to this through a control challenge. However, since this compensation will lower overall firm profits in an unanticipated manner, it will affect shareholders' inference of the manager's ability. Intuitively, if the manager were to undertake more wasteful spending on pet projects than was anticipated in equilibrium, and shareholders do not directly observe this wasteful spending, the consequent lower returns would lead the shareholders to infer a lower ability for the manager. This will in turn increase the likelihood that shareholders will want to replace the manager in the future, and will also lower the amount of observable compensation that the manager can pay himself in the future.

Hidden compensation might also be inefficient, in the sense that a dollar cost to the firm might yield less than a dollar of benefits to the manager. While this is not necessary for most of our results, we explore this possibility, as it seems natural and gives rise to a number of additional interesting implications. Indeed, there are good reasons to think that some forms of hidden compensation (pet projects, perks, empire-building) are likely to be grossly inefficient.

Our model studies the dynamic managerial choice of observable and hidden compensation given that managers set their own compensation subject to such partial entrenchment. In particular, managers choose the pay that is best for themselves, but must take into account career concerns: their choice of the magnitude and form of pay affects the likelihood that they are retained and their perceived future ability if they are retained. In this setting we analyze how observable and hidden pay interact with one another, and how changes in exogenous variables affect these two components of pay. As such, we provide a dynamic model for both overall managerial compensation, and the composition of this compensation over time. The distinct role that observable and hidden pay have in such a model of entrenchment gives rise to a unique set of predictions regarding the magnitude and form of compensation. For example, our model predicts that observable pay will be more responsive to managerial reputation and entrenchment, whereas hidden pay will be more responsive to noise in production and uncertainty about the manager's ability. The model also yields interesting implications for managerial firings.

We exploit this distinction in empirical tests of the model and find strong support for most of our predictions. The hidden part of compensation (given by the value of options and restricted stock granted, plus that of perquisites) increases with the noise in production process (proxied by variation in the firm's industry-adjusted return on assets), with the manager's outside option (as indicated by the strength of the CEO's ties to the firm, his age and MBA education), and decreases with the noise in evaluating the manager's ability (proxied by the inverse of his tenure as CEO). The results are robust to using another set of proxies for hidden pay, namely the likelihood and magnitude of the company engaging in options backdating. All these right-hand side variables impact observable pay (salary plus bonus) much less than they impact hidden pay. At the same time, observable pay increases with inferred managerial ability (based on historical industry-adjusted stock returns and returns on assets of firms ever managed by the CEO) and with the manager's entrenchment level (proxied by the Gompers, Ishii, and Metrick (2003) governance index), as predicted by our model.

It is worth remarking briefly on the stark distinction we make between observable and hidden pay in our model. In reality, most executive compensation falls somewhere in between these two extremes. For example, firms disclose managerial options in annual reports, but until recently, did so in a manner that made their value difficult for many shareholders to comprehend. And at the same time, certain implicit features that significantly increase the value of option plans to managers are not disclosed at all (i.e., an implicit understanding that options will be backdated or repriced under certain circumstances). Similarly, some perks are readily observable to the vigilant shareholder, whereas others are likely to be unobservable to all but corporate insiders. Information on pension plan compensation given to executives is publicly disclosed, but deciphering legal documents to understand details of this pay is rather involved and complicated (Bebchuk and Jackson (2005)).[7] We discuss compensation that is partially hidden, and how it might be incorporated into a generalization of this model in Section 6. We also discuss in Section 4 how the empirically relevant definition of hidden pay in our model is pay that is hidden from whatever mechanisms govern managerial replacement.

In Section 2 we present the basic model for this paper, and in Section 3 we analyze this model. Section 4 discusses our data, 5 provides empirical tests. Section 6 discusses implications and interpretations, with special attention paid to the questions of the recent sharp increase in executive compensation and of managerial firings. Section 7 concludes.

[7] Indeed, in the much publicized case the $140 million pension savings of Richard Grasso, chairman of the New York Stock Exchange, members of the board professed to be unaware of (and shocked by) the magnitude of the compensation that they themselves had approved. See for example "The Winding Road to Grasso's Huge Payday," in the New York Times, June 25, 2006.

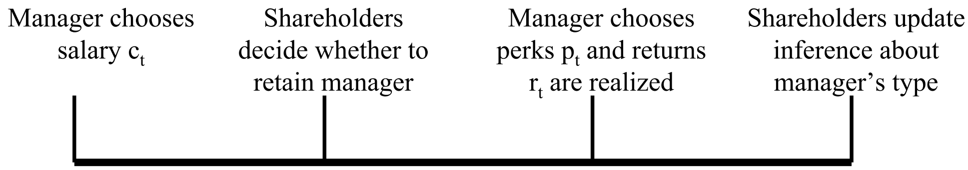

Figure 1: Timeline

2 The Model

We consider a two-period model of a firm run by a partially-entrenched manager, who sets his own compensation. In each period, compensation has two components. One is observable and committed to in advance; we will refer to this as *salary*. The other component is unobservable; we will refer to this as *perks*.[8]

Within each period, the following actions and decisions occur (see the timeline in figure 1). These decisions will be described in more detail immediately below. First, the manager announces a salary that he will receive over the period. Second, shareholders can decide whether to remove the manager, in which case the manager will not receive the announced salary or any perks. If removed, the manager is replaced by a newly drawn manager from a replacement pool of potential managers. Third, the manager chooses a level of unobservable perks and production occurs. And fourth, shareholders observe returns and update their inference of the manager's ability based on this outcome and their inferences about perk consumption.

We now detail agents' decisions and payments in each of these subperiods:

At the start of each of the two periods $t = 1, 2$, the manager is free to declare any salary c_t (up to the value of the firm).[9] As such, we envision the board of directors, who formally decides salary in the corporation, as being captured by the manager.

[8]Of course, in practice some components of salary are hard to observe and some components of perks may be readily observable. We use this nomenclature simply as a shorthand, and emphasize that the key distinction between our two forms of compensation is observability. That said, we would argue that many perks are indeed hard for shareholders to observe, or at least hard to observe that they are perks. Even for some readily observable accouterments of management – such as a jet plane, for example – it may be hard for an outsider to distinguish whether they are necessary expenditures for efficient firm operations, or inefficient benefits bestowed on management.

[9]We use subscripts on salary, perks, and inferences, to denote periods. Insofar as most of our attention will be on first period decisions, we will drop the subscript for first period decisions when the meaning is clear. Second period decisions will always be subscripted.

The manager, however, is constrained in this decision by the possibility that shareholders will remove him before he receives this pay. Specifically, after the manager announces c_t, shareholders decide whether to remove the manager or not. Shareholders remove a manager if they anticipate that the net outcome under this manager, including all anticipated compensation across all periods, will be worse than a replacement by more than a given amount E_t, that represents *entrenchment*. This involves comparing the manager's anticipated salary, perks, and the anticipated outcome to production given his inferred ability with that of a newly drawn manager, accounting for equilibrium play in all remaining periods.[10]

We call E_t entrenchment. Entrenchment can be interpreted as the cost that shareholders must face to replace a manager, due to free-rider problems, takeover defenses, legal challenges, coordination problems, search costs, firm reputation, or any other manner of replacement costs. Note that when we speak of entrenchment E_t, we are speaking of the amount by which the current manager can fall short in anticipated value of a replacement before being removed. As such, this term does not include any differences in inferred ability between the manager and a replacement (as is sometimes the case when this term is used). That is, managers in our model may have leeway for high compensation without being replaced by shareholders *both* because they are inferred to be of high quality and because they are entrenched; we separate these two terms in our notation and our nomenclature.

In each period t, entrenchment is given by

$$E_t = s_t - \beta c_t. \tag{1}$$

In this expression s_t is an exogenously given level of entrenchment that can be thought of reflecting both the general governance environment as well as firm specific institutional features (i.e., the presence of takeover defenses, the composition of the board, etc.). For much of our analysis we will take this term to be fixed over time, and denote this by s. The βc_t term, with $\beta \geq 0$, captures the possibility that entrenchment may be lessened

[10] We speak throughout this paper of "shareholders' " decisions. It should be understood that shareholders here could instead represent a raider, a subset of active shareholders, or any other group of agents that can initiate action to remove the manager at a cost. This distinction may lead to a different interpretation of "hidden" compensation – some components of compensation not easily observable or understood by a typical shareholder may be well observed and understood by a raider. We discuss the relationship between the governance mechanism and the interpretation of what is observable further in Section 4.

through higher pay.[11] This term is not critical in our analysis, and most of our results follow without it. However, this effect can interact in interesting ways with other effect in our model, so we allow for the possibility of $\beta > 0$. In Section 6 we also discuss the possibility of including random noise in the determination of E_t, which generates interesting implications for managerial firings.

After a retention decision is made, whichever manager is then in place chooses the level of perks p_t to consume. Perks, in contrast with salary, may be inefficient. In particular, for perks that cost the firm p_t dollars, the manager is assumed to obtain utility given by $V(p_t)$, with $V(0) = 0$, $0 < V'(\cdot) \leq 1$, and $V'' \leq 0$. This captures the notion that hidden pay may be inefficient, with this inefficiency increasing in the amount that is hidden. Indeed, some of the more significant forms of hidden compensation (pet projects, empire-building, etc.) may be very inefficient; the expected cost to the firm may vastly exceed the benefits to the manager. Consequently, an interpretation of V' being close to 0 when p is high may be reasonable in many settings.[12]

Additionally, there is an upper bound to perk consumption each period, given by $\bar{p}$, representing a limitation to the amount of compensation that can be hidden by the manager. (Or equivalently, $V'(p) = 0$ above some upper bound $\bar{p}$.) This bound $\bar{p}$ may the thought to be represent a technical limitation on hidden pay, which may change as new forms of compensation become technically available.[13] This upper bound will be binding, trivially, in the final period, as in that period the manager has no further reputational considerations. It may or may not be binding in the first period, when the manager has a reason to care about future reputation, and this distinction will be seen to give rise to an interesting difference in implications. After perks are chosen, production occurs. Production is mechanical (there is no effort in our model), with returns to the

[11] For example, higher pay may agitate shareholder action, or might lead to negative media or regulatory attention that increases the likelihood of replacement. See for example the discussion in Bebchuk and Fried (2004).

[12] Note at the same time this specification does not rule out $V(p_t) = p_t$; i.e., no inefficiency from perks. However, some of our comparative statics results will rely on $V'(\cdot) < 1$, as without this perks may be characterized by a bang-bang solution, rather than an interior optimum.

[13] More specifically, we envision that while a board of directors may be captured and willing to allow management to dictate pay, they may not be willing to approve a form or type of pay that is not adopted elsewhere, or that has not been deemed acceptable by compensation consultants. In effect, the manager's capture of the board might be limited by board members requirement that the possibility of litigation, scandal, or impropriety be avoided. As such, $\bar{p}$ might represent "acceptable" hidden pay, that the board and/or management believes will not be subject to future legal challenges. See Bebchuk and Fried (2004) on this point as well.

shareholders in period t given by:

$$r_t = a - c_t - p_t + \eta_t; \tag{2}$$

where a gives the manager's (unknown) ability, c_t and p_t are salary and perks paid this period, and η_t is noise, distributed as $\eta_t \sim N(0, \delta^2)$, where N represents the normal distribution, parameterized by its mean and variance. Returns are observable to the manager and shareholders alike, but are not contractible.[14]

Shareholders and the manager are symmetrically informed about the manager's ability. They both start with a prior given by a_1, distributed according to $a_1 \sim N(\mu, \gamma^2)$. If the manager is removed, he is replaced by a newly drawn manager whose ability is drawn from a pool distributed according to $a_n \sim N(0, \gamma^2)$. The expected ability μ of the incumbent manager may differ from that of a new draw, based on prior (unmodeled) periods or other information. We refer to the mean μ of this prior as a manager's "type", but emphasize that type in our model is not private information. After observing the outcome r_1 in period 1, the manager and shareholders update their inference about their manager's ability, taking into account the observed salary c_1 and their inference for perks p_1.

If the manager is fired, he obtains an outside option of $q \geq 0$ in each future period. We assume that the utility a manager can get from consuming the maximal amount of perks exceeds the manager's outside option: $V(\bar{p}) > q$. This ensures that the manager's IR constrain is not binding while employed by the firm (see the discussion on this point in the introduction). Also, a manager's salary and perks are constrained to be greater than or equal to 0. Thus, a manager cannot commit to pay money into the firm and take a negative salary. We presume that shareholders can set the salary for a replacement manager in the period when she is hired. Effectively, we are assuming that a replacement manager will not be entrenched until after she is hired, and this hiring can be contingent on her acceptance of a stated initial period salary. After such hiring, however, the replacement manager will be entrenched just as the initial manager, with the same

[14]This formulation permits negative returns for shareholders, and consequently, does not incorporate limited liability. This would not be hard to address, at the cost of some notational complexity. In particular, a fixed constant could be added to returns (or equivalently, to managerial ability). This would not affect any decision, and the probability of negative returns could be made arbitrarily small with a large enough fixed constant. And with a small enough probability of negative returns, truncating shareholder returns at 0 to incorporate limited liability would then have a negligible effect. We do not undertake this detail here, but it would be straightforward to incorporate into the model.

entrenchment parameter s, and the same ability to choose perks in the current period and salary and perks in future periods. As such, entrenchment s should be interpreted as a benefit that any manager will enjoy once she is installed in the firm.[15]

Shareholders and the manager are risk neutral, and both maximize expected returns. The discount rate is 0, or equivalently, returns can be thought to be expressed in units of the initial period. The manager's utility is separable in salary and perks, and separable across periods. Hence he maximizes the expectation of the sum of all salary, the value from all perks, and any outside pay he receives if fired. All random variables in the model $\{\epsilon_t\}_t$, $\{\eta_t\}_t$, a, and a_1 are uncorrelated, both cross-sectionally and over time.

3 Analysis

We look for Perfect Bayesian Equilibria (PBE) to our game. We solve the game backwards, as usual, beginning in period 2 (the final period), and then preceding to period 1. As is typical for games of career concerns or reputation, the final period is trivial, and it exhibits a number of particular "endgame" features. As such, our interest in not primarily in the outcome in the final period but rather the outcome in period 1, with the final period serving as a future period that will induce career concerns in period 1 play.

First consider the final period, period 2. Whether the initial manager or a replacement manager is in charge, maximal perks, $\bar{p}$, will be chosen in this final period, as there is no further reputational considerations after this period to constrain perks. Therefore, when shareholders are making the managerial retention decision this period, they correctly infer that all managers will consume the same perks $\bar{p}$. Hence the comparison between the incumbent and a replacement depends only on the difference in salaries, and the different inference of abilities.

Suppose that entering period 2, the market infers that the managers' expected type is μ_2. The values of μ_2 will of course be derived from inferences given period 1 outcome. Suppose that the manager has announced a salary of c_2, and recall that a replacement manager could instead be paid a salary of 0 for her initial period. Then, given entrenchment of $e_2 = s - \beta c_2$, it follows that the market will retain the incumbent manager if and

[15] Instead allowing the replacement manager to set her initial period salary immediately does not have any qualitative effect on results. We need to choose one assumption for specificity, and we chose the former assumption only because it seems reasonable to presume that initial salary may be negotiated as part of the hiring process before the new manager is entrenched.

only if:

$$\mu_2 - c_2 - \bar{p} + s - \beta c_2 \geq -\bar{p};$$

that is, provided that

$$c_2 \leq \frac{\mu_2 + s}{1+\beta}.$$

At the beginning of this period, the incumbent manager recognizes that his perk consumption is unaffected by his salary this period, provided that he is not fired. And since reputation does not matter after this period, all he cares about is salary plus perks. If instead, he is fired, he will not get either. Consequently, he sets salary to the highest level possible subject to not being fired; that is,

$$c_2 = \frac{\mu_2 + s}{1+\beta}, \tag{3}$$

for $\mu_2 + s \geq 0$. If instead $\mu_2 + s < 0$, then the manager cannot prevent firing in period 2 (recall that salary cannot be set to be negative).

Overall, then, the manager will realize benefits of

$$\frac{\mu_2 + s}{1+\beta} + V(\bar{p}) \tag{4}$$

in the final period provided that $\mu_2 \geq -s$, and benefits of q (the outside option) if this does not hold.

While this final period is trivial, it serves as the basis for decisions in the first period, and therefore deserves some comment. First, note the form this compensation takes as a function of the manager's inferred ability. The lowest type manager that is retained receives a surplus $V(\bar{p}) - q > 0$ over the outside option. That is, the manager's ability to extract perks ensures that his IR constraint will not be binding. The marginal CEO that is retained in our model is better off than one who is fired; he is not indifferent as in many standard executive compensation models.

Second, managerial compensation increases linearly with the market's inference μ_2 of the manager's ability. The lowest type manager that is retained is determined by entrenchment; that is, the lowest retained type is $\mu_2 = -s$. This manager receives net pay relative to the outside option of $V(\bar{p}) - q$. Higher types receive higher pay, due to their ability to extract a higher salary without getting fired.

Third, note that the shape of this final period compensation is as in Zwiebel (1996), and for similar reasons. In that model, managerial pay was determined through manager-firm bargaining subject to a replacement cost. The manager's ability to extract some of this replacement cost through the threat of leaving serves a role similar to the entrenchment here, resulting in managerial pay that exceeds an outside option. Also, in this paper, as well as in the present one, higher inferred ability translates into higher managerial compensation due to the higher anticipated outcome associated with a manager with higher ability.

Now consider period 1. Working backwards we first solve for the manager's perk decision. At the time of this decision, either the incumbent manager or a replacement manager may be present. The two differ only in that the inferred ability of the incumbent is distributed according to $N(\mu, \gamma^2)$, while that of a replacement manager is distributed according to $N(0, \gamma^2)$. We derive the perks for the former; the corresponding term for the latter follows immediately by setting $\mu = 0$.

In choosing perks, a manager must take into account the effect that this will have on the shareholders' inference of his ability, and consequently his pay and retention in the final period. Shareholders observe the returns r as given by equation (2) and salary c, and have an inference $\hat{p}$ about period 1 perks. Given normality assumptions for a and $\tilde{\eta}$, it then follows from a standard normal learning model that shareholders' posterior inference a_2 for the manager's ability will be normally distributed, with a mean given by:[16]

$$\mu_2 = \frac{\delta^2}{\delta^2 + \gamma^2}\mu + \frac{\gamma^2}{\delta^2 + \gamma^2}[r + c + \hat{p}], \tag{5}$$

and precision (the reciprocal of variance) given by:

$$\frac{1}{\delta^2} + \frac{1}{\gamma^2}. \tag{6}$$

Intuition for equations (5) and (6) is straightforward. Upon observing outcome r, the shareholders update their inferred mean ability of the manager to be a weighted average of their prior μ and what their best estimate of ability would be if they only observed r and had a diffuse prior. (This is given simply by the inference that would follow with no noise, $r + c + \hat{p}$, and is also the maximum likelihood estimator.) The weights that

[16]see DeGroot (1970), Chapter 9 for details on normal learning models and more generally, on conjugate prior distributions.

are put on the prior versus the new estimate depend on the signal to noise ratio – if the prior is very noisy (high γ^2), more weight is put on the new information, while if the new information is very noisy (high δ^2), more weight is put on the prior. With the new observation, the posterior is more precise than the prior, with precision increasing (and hence variance decreasing). The fact that precisions of the initial noise and the signal are additive for the new prior, and that the posterior follows a normal distribution, follows from normality assumptions of both noise terms.

The outcome $\tilde{r}$ is of course stochastic for both the manager and shareholders, as ability a and the outcome of $\tilde{\eta}$ are both unknown. Furthermore the manager may choose a different level of perks than that anticipated by shareholders. Substituting (2) into (5), it follows that:

$$\mu_2 = \frac{\delta^2}{\delta^2+\gamma^2}\mu + \frac{\gamma^2}{\delta^2+\gamma^2}[a + (\hat{p} - p) + \tilde{\eta}].$$

And now, noting that $a \sim N(\mu, \gamma^2)$, $\eta \sim N(0, \delta^2)$, and these two terms are uncorrelated, it follows that this can be rewritten as:

$$\mu_2 = \mu + \frac{\gamma^2}{\delta^2+\gamma^2}[(\hat{p} - p) + \tilde{w}], \tag{7}$$

where $\tilde{w}$ is a random variable distributed according to $N(0, \delta^2 + \gamma^2)$. Equation (5) indicates that the ex-ante distribution of the mean of the shareholders' posterior at the end of period 1 will be given by their prior, plus a term that is positive (negative) if they inferred a perk that was higher (lower) than the actual perk, plus a 0-mean noise term. Intuitively, if the manager consumes more perks than anticipated, this will lower returns, and will be attributed in part by the shareholders to a lower ability and in part to a bad draw of noise η. Conversely, if the manager consumes less perks than anticipated, this will lead to a higher inference by shareholders. In equilibrium, of course, the inferred perks will be correct, and on average, the posterior mean will be the same as that for the prior. It follows immediately from equation (7) that if the manager consumes perks p, then the ex-ante distribution of the shareholders' mean inference of this manager's ability will be given by:

$$\mu_2 \sim N\left(\mu + \frac{\gamma^2}{\delta^2+\gamma^2}(\hat{p} - p), \frac{\gamma^4}{\delta^2+\gamma^2}\right). \tag{8}$$

The manager chooses p in order to maximize his combined expected period 1 and period

2 returns. Period 1 salary is already sunk, and period 2 returns are given by equation (3) if he is retained, and by q is he is fired. Hence, p is chosen to maximize

$$V(p) + \int_{-s}^{\infty} \left(V(\bar{p}) + \left(\frac{\mu_2 + s}{1+\beta} \right) \right) g(\mu_2) d\mu_2 + \int_{-\infty}^{-s} q g(\mu_2) d\mu_2; \tag{9}$$

where g is the density of μ_2 as given in equation (8).

We let $u(p)$ and σ^2 represent the mean and variance of the ex-ante distribution of μ_2 in equation (8); that is,

$$u(p) \equiv \mu + \frac{\gamma^2}{\delta^2 + \gamma^2}(\hat{p} - p),$$

$$\sigma^2 \equiv \frac{\gamma^4}{\delta^2 + \gamma^2};$$

and we let ϕ and Φ represent the density function and the cdf of the standard normal distribution respectively. Then, after a substitution of variables, evaluating the integrals, and rearranging terms, expression (9) can be rewritten as:

$$V(p) + \left(1 - \Phi\left(\frac{-s - u(p)}{\sigma}\right)\right)\left(V(\bar{p}) + \frac{u(p) + s}{1+\beta}\right) + q\Phi\left(\frac{-s - u(p)}{\sigma}\right) + \frac{\sigma}{1+\beta}\phi\left(\frac{-s - u(p)}{\sigma}\right). \tag{10}$$

Differentiating this expression with respect to p, and rearranging terms, yields the following first order condition for p:

$$\begin{aligned} V'(p) &= \frac{\gamma^2}{\delta^2 + \gamma^2}\left[\frac{1}{\sigma}\phi\left(\frac{-s - u(p)}{\sigma}\right)\left(V(\bar{p}) + \frac{u(p) + s}{1+\beta} - q\right)\right. \qquad (11) \\ &\left. + \frac{1}{1+\beta}\left(\phi\left(\frac{-s - u(p)}{\sigma}\right)\left(\frac{-s - u(p)}{\sigma}\right) + \left(1 - \Phi\left(\frac{-s - u(p)}{\sigma}\right)\right)\right)\right]. \end{aligned}$$

Equation (11) is readily interpreted. The left hand side gives the first period benefits of added consumption from increasing p. The right hand side gives the reputational costs of increasing p in the second period. The term on the first line of this expression accounts for the additional probability of being fired if p is increased, as the last bracketed term of this expression gives the difference in expected second period compensation between a manager with inferred type $u(p)$ and one that is fired. The coefficient on this term gives the increased likelihood of being fired if p is increased. This term is always positive; the greater the perks p consumed, the lower the inference and the greater the probability of being fired for all types.

The second line on the right hand side accounts for the cost of lower reputation independent of firing that follows from an increase of p. Second period compensation increases linearly in reputation as per equation (3), and consequently, a lower reputation translates into lower expected second period compensation even if the manager is not fired. Dividing this term by $1 - \Phi\left(\frac{-s-u(p)}{\sigma}\right)$, gives 1 plus the hazard rate of the standard normal distribution evaluated at this point. From this, it is immediate to show that the second line on the right hand side is always positive (representing a net cost), and is a decreasing function of the argument $\left(\frac{-s-u(p)}{\sigma}\right)$, which takes values from 1 (as $\left(\frac{-s-u(p)}{\sigma}\right) \to -\infty$) to 0 (as $\left(\frac{-s-u(p)}{\sigma}\right) \to \infty$).

Both terms on the right hand side of equation (11) have the coefficient $\frac{\gamma^2}{\delta^2+\gamma^2}$, signifying that the extent to which an increase in perks will affect second period reputation depends on the signal to noise ratio. In particular if δ is large relative to γ, then because there is a lot of noise in outcome, inferences will not change as much with a change in perks, as variations in outcome will mainly be attributable to production noise. Consequently, the reputational cost of consuming perks will be lower. Conversely, if γ is very large relative to δ, increasing perks will translate almost one to one with an expected lower inference.

Equation (11) can readily be simplified by canceling the $(u(p) + s)/(1 + \beta)$ terms and imposing the equilibrium condition that $\hat{p} = p$ (and therefore $u(p) = \mu$), i.e, that shareholders' inferences about p are correct. This gives:

$$V'(p) = \frac{\gamma^2}{\delta^2 + \gamma^2}\left[\frac{1}{\sigma}\phi\left(\frac{-s-\mu}{\sigma}\right)(V(\bar{p}) - q) + \frac{1}{1+\beta}\left(1 - \Phi\left(\frac{-s-\mu}{\sigma}\right)\right)\right]. \tag{12}$$

Note that the right hand side of equation expression (12) is always positive, as $V(\bar{p}) > q$. We will explore this expression in detail in our comparative statics below, after first turning to the manager's first period choice of salary c.

Equation (12) gives equilibrium first period perk consumption. If parameters are such that the value of the right hand side of equation (12) falls between $V'(0)$ and $V'(\bar{p})$, then there will be an interior level of perk consumption. If instead the value of the right hand side of equation (12) is less than $V'(\bar{p})$ then $\bar{p}$ will be consumed, and if it is greater than $V'(0)$, then no perks will be consumed. One special case is if the value of perks V is linear in perks p (our assumption of weak concavity did not rule this out). Then generically, all types will choose one of $p = \bar{p}$ (if the right hand side of equation (12) is less than the constant of linearity for V) and $p = 0$ (if the converse holds).

Let $P_1(\mu)$ represent the equilibrium first period perks chosen by a manager of type μ. We now consider first period salary c. First consider the case when $\beta = 0$. In such an event, the expected return r_2 that shareholders get in the second period does not depend on the manager's type μ, since the manager extracts any surplus in salary. Consequently, shareholders receive expected returns of

$$\mu - c_2 - p_2 = -s - \bar{p}$$

for all types. If instead the shareholders fire the manager, they receive expected returns of $-\bar{p}$ (since a new manager will consume maximum perks in the final period and will be of type $\mu = 0$), and they will have to pay the replacement cost of s, yielding once again the same net expected benefits.[17]

Consequently, it follows that for $\beta = 0$, when shareholders make firing decisions in the first period, they do not need to consider the benefits from type in the second period, as their expected returns will be the same in equilibrium for all types. Hence, they only need to consider whether first period expected returns for a manager given the announced salary will come within $E = s$ of expected returns in the first period for a newly drawn manager. Since a new manager receives salary of 0, and is of type $\mu = 0$, it follows that expected first period returns of a newly drawn manager is simply given by $-P_1(0)$. And since a manager of type μ will set salary as high as possible without triggering replacement, it follows that this manager will choose period 1 salary to satisfy

$$\mu - c - P_1(\mu) + s = -P_1(0),$$

or simply

$$c = \mu + s + (P_1(0) - P_1(\mu)). \tag{13}$$

When instead $\beta > 0$, things are more complicated for several reasons. First, since higher salary lowers managerial entrenchment, it is no longer the case that in the final period all managers extract the full benefit of their inferred value through salary. And since shareholders realize some surplus, they are no longer indifferent between who will manage in the second period. High type managers can in turn extract some of this second period surplus through higher first period salary. Second, higher first period salary affects

[17] The fact that these expected benefits are negative is once again only a scaling convention. If returns for all types also included a constant independent of type, this constant would go to shareholders.

first period entrenchment, thereby dampening the extra first period salary that higher types extract.

Considering the first effect, it follows from equation (3) that the equilibrium value to shareholders of a type μ manager in period 2 is given by:

$$\mu - c_2 - p_2 = \frac{\beta}{1+\beta}\mu - \frac{s}{1+\beta} - \bar{p}.$$

Consequently, the expected difference in value in period 2 between a type μ manager and a type 0 manager (hired in period 1) will be

$$\frac{\beta}{1+\beta}\mu.$$

Turning now to the second effect, it then follows that at the beginning of period 1, a manager of type μ will choose period 1 salary c to satisfy

$$\mu - c - P_1(\mu) + s - \beta c + \frac{\beta}{1+\beta}\mu = -P_1(0);$$

or equivalently,

$$c = \frac{1}{1+\beta}\left[\frac{1+2\beta}{1+\beta}\mu + s + (P_1(0) - P_1(\mu))\right]. \tag{14}$$

We summarize our results for perks and salary in the following Proposition.

Proposition 1 *In period 2, salary and perks are given by:*

$$p_2 = \bar{p};$$

$$c_2 = \frac{\mu_2 + s}{1+\beta}.$$

In period 1, perks are given by the solution to:

$$V'(p) = \frac{\gamma^2}{\delta^2+\gamma^2}\left[\frac{1}{\sigma}\phi\left(\frac{-s-\mu}{\sigma}\right)(V(\bar{p}) - q) + \frac{1}{1+\beta}\left(1 - \Phi\left(\frac{-s-\mu}{\sigma}\right)\right)\right]$$

if this condition is satisfied for some $0 \leq p \leq \bar{p}$, and otherwise by $p = \bar{p}$ or $p = 0$ if right hand side of this expression is less than $V'(\bar{p})$ or greater than $V'(0)$ respectively. Letting this solution for perks as a function of type be given by $P(\mu)$, first period salary is given *by:*

$$c = \frac{1}{1+\beta}\left[\frac{1+2\beta}{1+\beta}\mu + s + (P_1(0) - P_1(\mu))\right].$$

Comparative statics results for salary and perks follow from this equilibrium determination. Salary and perks in the second period are a consequence of "endgame" effects, when there are no further career concerns. Consequently, we focus on the period 1 results, which are more relevant for any setting where managers must trade off current compensation with considerations of future reputation. The next two propositions summarize a number of comparative statics results for first period salary and perks, and form the basis for many of our empirical tests.[18]

Proposition 2 *Provided that first period perks has an interior solution (i.e., perks are not either* 0 *or* $\bar{p}$*), first period perks* p *are increasing in* δ *(noise in the production process),* q *(the manager's outside option), and* β *(the loss of entrenchment due to high pay), and are decreasing in* γ *(uncertainty in the manager's ability), and* $\bar{p}$ *(maximum perks). If perks are instead given by a corner solution, all these results hold weakly, save for the effect of changes in* $\bar{p}$*. When* $p = \bar{p}$ *is binding, perks increase in* $\bar{p}$*.*

To understand the intuition for these results in Proposition 2 note that perks are costly to managers because they lower their inferred ability, and consequently, their future pay. Perk consumption lowers output, and when shareholders observe a lower output than expected, they must decide how much of this to attribute to a lower than expected managerial ability, and how much to attribute instead to unrelated noise. The more uncertainty there is about managerial type, and the less production noise there is in outcome, the more weight will be placed by shareholders on lower managerial ability, and consequently, the greater the expected negative impact on reputation of increasing perk consumption. Perks increase in the manager's outside option, because this places a bound on the maximum reputational cost that higher perk consumption can induce. Perks increase in the loss of entrenchment that comes about from high pay because this term lessens the amount that managers with a higher reputation can extract, and therefore lessens the importance of future reputation. Finally, there are two effects of an increase in the upper bound of $\bar{p}$ on perk consumption in the first period. If this bound is binding,

[18] Proofs are in the appendix.

then trivially, increasing this bound will lead to greater perk consumption. However, if it is not binding, perk consumption will decline, because increasing $\bar{p}$ increases the value of managing in the second period (because $\bar{p}$ in perks will always be extracted in the second period), which makes the manager more adverse to being fired and consequently more sensitive to reputation.

The effect that these variables have on first period salary is more subtle, and depends on whether the variable changes just for the manager in question or for all potential managers of the firm. Note that none of these direct considerations with perks are relevant to first period salary. Since salary is observable, it does not affect market inferences about the manager's ability, and hence does not affect reputation. However, all of these variables do effect first period salary indirectly through their effect on perks. Recall that managers will pay themselves the largest salary that they can without getting fired, and this will depend on shareholders' inferences about the value of a manager relative to a replacement. Given this, if shareholders infer that perk consumption will be higher in the first period for a given type manager, it follows that the manager will have less leeway to extract a salary in this period as well. This will cause all of these variables to affect salary in the opposite direction as in Proposition 2.

However, at the same time, the overall benefits that an incumbent manager can extract depends on what a potential replacement manager will get as well. If the variable in question also changes for the replacement manager, this will have an opposite effect on the incumbent. That is, any increase (decrease) in perks that the replacement manager obtains in turn slackens (tightens) the constraint on first period salary faced by the incumbent. Hence, if a change in one of the exogenous variables of Proposition 2 applies to both the current manager and a replacement manager, the net effect on the salary of the current manager will depends on whether second period perk consumption increases more for the type in question or for the replacement type ($\mu = 0$). In general either is possible, however, in general the net effect will typically be that changes in these exogenous variables affect salary less than perks.

These effects are formalized in the following two propositions. Proposition 3 considers the easy case, where exogenous variables change only for the incumbent and not for a replacement manager. The result is trivial, following immediately from equation (14). Proposition 4 considers the more complicated case where exogenous variables change for both the incumbent and the replacement manager. Both of the propositions are of empirical interest, as in some settings changes are likely to be manager specific, while in other settings, changes are likely to affect all managers a firm might hire.

Proposition 3 *Suppose that changes in δ (noise in the production process), q (the manager's outside option), γ (uncertainty in the manager's ability), and $\bar{p}$ (maximum perks) are changes only for the current manager and not a replacement manager. Then, the effect of a change in any of these variables on first period salary is given by $-\frac{1}{1+\beta}$ times the effect that the change of this variable had on first period perks.*

Proposition 4 *Consider changes in δ (noise in the production process), q (the manager's outside option), γ (uncertainty in the manager's ability), and $\bar{p}$ (maximum perks) that are changes for both the current manager and any potential replacement manager. Then, provided that first period perks has an interior solution (i.e., perks are not 0 or $\bar{p}$),*

1. *For any type μ, the effect of a change in δ, q, γ, and $\bar{p}$ on first period salary c will be of the opposite sign as the effect of the change of this variable on perks, if and only if the change in perks for this type μ is larger than the change in perks for the replacement type 0.*

2. $$\frac{dc}{d\delta} > -\frac{dp}{d\delta}; \qquad \frac{dc}{dq} > -\frac{dp}{dq}; \qquad \frac{dc}{d\gamma} < -\frac{dp}{d\gamma}; \qquad \frac{dc}{d\bar{p}} < -\frac{dp}{d\bar{p}}. \tag{15}$$

3. *Let $V'''(x) = 0$. Then for all types μ such that $\phi\left(\frac{-s-\mu}{\sigma}\right) \geq \frac{1}{2+\beta}\phi\left(\frac{-s}{\sigma}\right)$, the effect of a change in δ, q, γ, and $\bar{p}$ on first period salary c will all be of smaller absolute magnitude than their effect on first period perks p.*

Propositions 3 and 4 effectively indicate that changes in δ, q, γ and $\bar{p}$ have a greater impact on perks than on salary. In particular, Proposition 3 indicates that if these variables only change for the manager in question, the effect on consumption c is always smaller and in the opposite direction from the change in perks p. Proposition 4 indicates that if any one of these variables change for both the manager in question and any replacement manager, that 1) the change in c will be in the opposite direction as the change in p if and only if the effect of this change on perks for the manager in question is larger than for a replacement manager; 2) if this change in c is of the opposite sign as the change in p it is always of smaller absolute magnitude; and 3) if $V''' = 0$ then, for "most types" the change in c is of smaller absolute magnitude than the change in p regardless of it's sign.

Intuitively, if these variables did not affect the value of a replacement manager, first period salary would simple reverse the change in anticipated perks through equation (14) (weighted by $\frac{1}{1+\beta}$). So if a given manager was expected to consume more perks, he would have to reduce salary by a similar amount. This is true even though he would prefer to consume salary if perks are inefficient. However, if this effect is mitigated by similar changes in the perks of the replacement manager, the effect on salary is not as large. And if the replacement manager changes perks more than that of a given type of manager in response to a change in an exogenous variable, then the change in both perks and salary will go in the same direction for this type of manager.

While the condition in part 3 in Proposition 4 is restrictive ($V''' = 0$, i.e., V is quadratic), several points are worth noting. First, no similar restrictions are necessary in Proposition 3, where we hold fixed the compensation for the replacement manager; we only need this assumption in this more general setting. While the change in the manager's own perks always leads to an opposite change in salary, the change in the replacement manager's perks goes in the opposite direction. Inequality (21) in the Appendix indicates that taken together, the overall effect on the incumbent's salary will be less than the effect on perks unless this latter effect exceeds the former by a factor of $(2 + \beta)$. We use the assumption of $V''' = 0$, to rule out this possibility for types within the range specified in condition 3.

A second point is that this range of types specified in part 3, given by μ such that $\phi\left(\frac{-s-\mu}{\sigma}\right) \geq \frac{1}{2+\beta}\phi\left(\frac{-s}{\sigma}\right)$, will include "most" types of managers provided that entrenchment s is large. This restriction provides both an upper and lower bound for types, but it is straightforward to show that the lower bound is below $-s$, and that the firing threshold in the first period exceeds that level. (This is shown at the end of the proof to this Proposition.) Thus, all managers below this lower bound (that don't satisfy this restriction on types) will not be retained by the firm anyways. Consequently, the result of part 3 of the Proposition holds for all managers still in the population except the highest types which exceed the upper bound given by this type restriction.

Finally, it is worth remarking that these restrictions are much stronger than necessary for many specific functional forms. Furthermore, by continuity, if V''' is small, we will obtain the same result for a very similar set of types. In summary, under "reasonable" settings, the magnitude of a change in one of the exogenous variables in Propositions 3 and 4 on c will be smaller than the effect on p for most types, even when we allow for this chance in the variable to affect the replacement manager as well as the incumbent.

In contrast with Propositions 2-4, changes in managerial type μ and exogenous entrenchment s both have a direct effect on salary and perks, as summarized in Proposition 5.

Proposition 5 *Provided that γ (uncertainty about a manager's type) is "large enough", if type μ or entrenchment s increases, first period salary increases. First period perks may either increase or decrease, depending on the level of $\mu + s$. In particular, perks increase in μ and s if and only if*

$$s + \mu \geq \frac{\sigma^2}{(1+\beta)(V(\bar{p}) - q)}. \tag{16}$$

Note that if $V(\bar{p}) - q$ is much larger than σ^2, condition (16) is approximately $\mu \geq -s$. The intuition for this is that the marginal increase in the likelihood of being below the firing threshold level if p is increased marginally above its inferred value is greatest for type $\mu = -s$. Up to that point this marginal cost increases, and types will therefore decrease perks. Above that point the marginal cost decreases. Intuitively, types way above $-s$ are unlikely to receive such a low outcome to be fired in the second period even if they undertake a high p, while types way below (if they even survive the first period) are likely to be fired in any event in the second period. The right hand side term is present in this expression because the cost of being fired is mitigated by the lower bound on pay ensured by q. Lower types receive more value from this outside option than higher types, and therefore are marginally more inclined to engage in more p holding the firing effect fixed. As $V(\bar{p}) - q$ grows (the discontinuity from firing becomes more significant, or as σ falls (there is less uncertainty about type), this effect becomes less important.

Note however that shareholders will replace managers in the first period who are below a given threshold (just as in the second period). This threshold will be defined by the type such that even with a salary of $c = 0$, the expected future value of a replacement manager exceeds this type by the level of entrenchment s. While characterizing this threshold is more complicated than doing so for the final period, under a moderate parametric assumption, one can demonstrate that it exceeds $-s + \frac{\sigma^2}{(1+\beta)(V(\bar{p})-q)}$, and consequently, Proposition 5 indicates that perks will increase in μ and s for all types of managers that are retained.

As noted above, if the right hand side of equation (12) is less than $V'(\bar{p})$, then perk consumption will be given by the upper bound $\bar{p}$. This can occur for a number of reasons, the most empirically relevant, perhaps, being a low signal to noise ratio. In particular, if shareholders learn little about a manager's ability from a change in outcome due to increased perks, reputation does not serve to constrain perks. This may be the case for many large firms, or for firms with lots of exogenous noise in outcome.

In such an event, our model would suggest that perks, and overall managerial compensation, are only constrained by the extent to which compensation can be hidden. Note that one critical difference between our model and a standard formulation of managerial pay is that if managerial pay is being determined by a profit maximizing principal, higher perks should be offset with lower salary. To the extent that options, or pension plans, or severance pay is an efficient manner to pay management, when such forms of pay become available, salary should be decreased commensurately. In our setting, if $\bar{p}$ increases and the right hand side of (12) is low, overall pay of all managers increases, as the shareholders cannot constrain even a newly hired manager from consuming a lot of perks, and the pay of other managers follows in relation to this based on ability and entrenchment. We return to this notion in our discussion section, and offer up some thoughts on how this may explain the dramatic rise in CEO pay in the U.S. over the last 50 years.

4 Empirical methodology

When turning to test our model, two methodological issues arise. First, there is the question of what should be included in each of observable and hidden compensation. "Hidden", in the context of our model, means hidden from the mechanism responsible for retaining or firing the manager. While formally most managerial firings come at the board's behest, what may be more relevant is the pressure underlying such decisions. In practice, this mechanism is likely to be complex, and to vary across firms and time.

In the interpretation of our tests, we will often take the perspective that the relevant governance mechanism is rather naive. Managerial replacement often seems to require widespread shareholder dissent or repeated negative media coverage. Such mechanisms are arguably financially "unsophisticated", leading to our interpretation of hidden compensation to be inclusive of all compensation save for salary and bonuses. Thus, in the context of our model, we are effectively arguing that many shareholders (and/or the financial press) do not directly observe or comprehend the value of, say, option grants or perks when given to management. At the same time, however, they are presumed to be capable of observing firm earnings, and forming inferences on managerial ability based on these earnings. Obviously, if one were to view the shareholder governance mechanism as more sophisticated, one would choose to categorize more forms of compensation as observable and fewer forms as hidden in testing our results. Note that even if one takes the governance mechanism to be very sophisticated, our backdating variable gives us a form of compensation that is observable to us today, but was very likely unobservable to shareholders when granted.

Second, there is the issue that only a limited set of proxies of hidden compensation are available to the econometrician. While some forms of compensation might be hidden to the average shareholder but discernable to the financial economist, other forms are likely to be hidden to all but corporate insiders or select board member. CEO perks may include, for instance, the execution of "pet projects" but there is no clear way to identify such endeavors. Similarly, implicit agreements to backdate or reprice options will be hard to observe, but may confer significant value. Thus we are potentially missing a significant fraction of the value of hidden compensation, which biases the estimation procedure against finding any results in accordance with our model.

In light of these issues, we test the implications of the model separately for all available proxies of both observable and hidden pay, so that it is transparent which variables impact each of these components, and whether these effects conform to our predictions.

For all companies tracked by the Execucomp dataset available through Compustat NorthAmerica, we extract information regarding the identity of the CEOs, their observable compensation, proxies for perks, and measures of company performance. Our final panel dataset contains 1724 companies and 2383 CEOs during 1992-2005.[19]

We construct two proxies for inferred CEO ability μ, based on either of two performance measures: the company's return on assets (ROA) and its stock return (RET). Each year we divide the companies in our sample into bins based on their two-digit SIC

[19]The number of CEOs tracked by Execucomp is roughly twice as large as those in our final dataset, but only half of those observations have non-missing entries for all the variables needed in the empirical analysis. The Execucomp variable that is most sparse is *joined_c*, the year when the executive joined the company. We need this variable to calculate a proxy for the executive's firm-specific capital, which we argue is inversely related to his outside option q.

code. Our two proxies for μ, labeled $MeanInferredAbility_t^{ROA}$ and $MeanInferredAbility_t^{RET}$, are constructed for each CEO each year by averaging the yearly performance of the companies managed by the CEO since he entered our dataset up to and including that year. The yearly performance is measured as the ranking of the company's ROA (or RET) relative to the ROA (or RET) of its peers in the corresponding year-industry bin. For instance, if a company ranks 2 out of 20 in year t in its industry group, then its ranking for t is 2/20=0.1. The higher is the ranking, the better is the performance of the company that year, and thus the higher is the inference about the CEO's quality.[20]

As proxy for the level of CEO entrenchment s we used the Governance Index proposed[21] by Gompers, Ishii, and Metrick (2003), as done previously by Fisman, Khurana, and Rhodes-Kropf (2005). By construction, the higher is the governance index, captured by our variable $Entrenchment_t$, the more restricted are the shareholder rights, and thus, the higher is the managerial power[22].

We use several variables from Execucomp to measure the observed CEO compensation c. $TotalObservedCompensation_t$ is the sum of the CEO's salary ($Salary_t$) and bonus ($Bonus_t$) in fiscal year t (data items *tcc*, *salary* and *bonus* in Execucomp, respectively).

We use three data sources to construct proxies for hidden pay. The first source is Execucomp. Our proxies for hidden CEO compensation p gathered from this source are captured by the variables[23] $Options_t$, $RestrictedStock_t$ and $OtherAnnual_t$ (data items *blk_valu*, $rstkgrnt$ and *othann* in Execucomp, respectively.) We also calculate the total hidden compensation for each CEO each year, $TotalPerks_t$, as the sum of $Options_t$, $RestrictedStock_t$ and $OtherAnnual_t$.

$Options_t$ is the aggregate value of the stock options granted to the executive during

[20]Clearly, variables $MeanInferredAbility_t^{ROA}$ and $MeanInferredAbility_t^{RET}$ are noisy proxies for μ, because ROA and RET are noisy measures of CEO ability. ROA is prone to accounting manipulations, and the stock return RET contains the market's inferences about other firm or CEO characteristics that influence future performance, aside from the CEO's ability.

[21]We thank Andrew Metrick for providing us with the index data.

[22]We currently do not have an appropriate proxy for the maximum perk compensation $\bar{p}$. Empirically, changes in $\bar{p}$ and changes in entrenchment s are difficult to dissentangle, and thus we can not readily test the model's predictions regarding $\bar{p}$.

[23]In the executive compensation literature the norm is to use the log value of compensation as dependent variable in regressions that analyze the determinants of pay, and not the $ value of this quantity. Similarly, log value of firm size is usually included as a right-hand side variable. The predictions of our model dictate that we use the $ value of various types of compensation, as well as $ value of firm size, and not their logs. However, our empirical results are similar when we use log values instead of $ quantities.

year t as valued using the S&P Black Scholes methodology. $RestrictedStock_t$ captures the dollar value of restricted stock granted during the year to the CEO. $OtherAnnual_t$ captures the dollar value of other annual compensation not classified as salary or bonus, and includes items such as perquisites and personal benefits, above market earnings on restricted stock, options or deferred compensation paid during the year but deferred by the CEO. It also includes tax reimbursements, and the value of the difference between the price paid by the CEO for company stock and the actual market price of the stock under a stock purchase plan not available to shareholders or employees of the company.

In addition, we use data from Bizjak, Lemmon, and Whitby (2006) to classify firms as likely to have engaged in options backdating in any given year during 1996-2002.[24] Options backdating fits our model's notion of hidden pay well; the actual grant day is not revealed to shareholders, and hence, neither is the true size of additional compensation received by the CEO from the backdated grant date. Bizjak, Lemmon, and Whitby (2006) classify a firm as likely to have backdated options in a given year if at least 20% of the firm grant days that year exhibit evidence of backdating. An option grant day is identified as being backdated if the market-adjusted stock price declined at least 10% in the 20 days prior to the grant and increased at least 10% in the 20 trading days after the grant. We define the variable $Backdating_t$ to be a dummy equal to 1 for firm-year observarvations for which the firm is likely to have engaged in backdating according to Bizjak, Lemmon, and Whitby (2006).

Finally, we estimate[25] the value of additional compensation obtained by the CEO by backdating stock options granted during year t. For each grant received by the CEO during year t we calculate $Blkvalu\_TranDate_t$, the Black Scholes value of the grant assuming that it was made on the declared transaction date as listed in Thompson Financial Insider Filings Database. We also calculate $Blkvalu\_SecDate_t$, the Black Scholes value of the grant assuming that the CEO actually received the options on the day the company disclosed the grant to the SEC, but used as strike the share price on the declared (and earlier) transaction date. To calculate both $Blkvalu\_TranDate_t$ and $Blkvalu\_SecDate_t$ we use the annual estimates for the firm's volatility and dividend yield as found in Execucomp (data items *bs_volatility* and *bs_yield*), as well as the annual risk free rates used by Execucomp to calculate their own values for executive option grants. The amount of com-

[24] We are grateful to Michael Lemmon and his co-authors for sharing their data with us.

[25] We thank Yaniv Grinstein for suggesting this approach and directing us to the appropriate data sources.

pensation gained by the executive by backdating is defined as: $BackdatingAmount_t = (Blkvalu\_SecDate_t - Blkvalu\_TranDate_t) * 1_{Blkvalu\_SecDate_t > Blkvalu\_TranDate_t}$.

Our proxy for the uncertainty in the manager's ability is the inverse of the length of the officer's tenure at the company as CEO ($InverseCEOTenure_t$). We therefore assume that γ is positively correlated with $InverseCEOTenure_t$.[26]

In the empirical analysis we will assume that the shock η_t in the production process of a company has a firm-specific as well as an industry-specific component, and the latter is observable. The inference about CEO ability in a particular year should therefore be done after filtering out the industry-wide observable shock, which we will assume is the average ROA in the industry that year. If the industry-wide shock is not variable over time, then this component of η_t can be ignored. Our first proxy for the standard deviation δ of η_t is constructed based on this assumption. For each industry, defined by a two-digit SIC code, we calculate the standard deviation of ROA, across all firms in that group, and across all years in our sample. The resulting variable, $\sigma(ROA)^{Industry}$, is our first proxy for the noise in production δ, and will have the same value for all firms with the same two-digit SIC code. Implicitly, we assume that there is no correlation between the industry-specific noise in output δ, and the CEO-level noise in ability, γ, and thus we can identify δ from the industry-specific variation in ROA. If, however, executives with high uncertainty in ability (high γ) tend to work in industries with higher noise in production δ, then using $\sigma(ROA)^{Industry}$ as a proxy for δ would be problematic. However, we have no reason to expect this to be the case.

Our second proxy for δ allows for there to be variation over time in the industry-wide shock to ROA. The inference about a CEO's ability at time t will be made based on the excess ROA of the CEO's company relative to the industry-average ROA that year. We calculate our new proxy for δ, labeled $\sigma(ROA_{i,t} - ROA_{IND,t})$, by taking the standard deviation across all firms in an industry of those firms' excess ROA every year. We find that the two proxies for δ, $\sigma(ROA)^{Industry}$ and $\sigma(ROA_{i,t} - ROA_{IND,t})$, have a correlation of 0.99, indicating that the industry-wide component of the production shock η_t does not

[26]It is worth noting that using $InverseCEOTenure_t$ as a proxy for γ, while quite intuitive, could be problematic, because tenure may be correlated with entrenchment. In our model entrenchment s is exogenous, but in reality it may not be. CEOs who have had their job for longer may have higher values of s, perhaps because they have had more time to become friendly with the board of directors. Thus, the inverse of tenure may be a proxy for a combination of noise in ability γ and entrenchment s, but this problem is alleviated since in our regression models we include the Gompers, Ishii, and Metrick (2003) index as a separate measure of s.

have much time variability during our sample period. Therefore, for brevity, we will only use one of these two variables, namely $\sigma(ROA_{i,t} - ROA_{IND,t})$, as the proxy for δ in the rest of the paper.

We construct several proxies for parameter q, the manager's outside option. The first measure, $Inverse\ of\ Firm - Specific\ Capital$, is the inverse of the number of years the executive worked for the company before becoming its CEO. The length of employment in the company prior to getting the CEO position is a measure of firm-specific capital (Murphy and Zabojnik (2004)), and arguably it is inversely related to the manager's outside option. The second proxy for q is the indicator variable $HasMBA$, which is equal to 1 for observations corresponding to executives who have an MBA degree. These data were obtained by manually searching for the biographical information of all CEOs in our sample in two sources of such data, Marquis' Who's Who and ZoomInfo.com. As suggested by Frydman (2005), having an MBA degree is an indicator general human capital. Therefore, we assume that $HasMBA$ is positively correlated with the manager's outside option q. Finally, our third proxy for q is the dummy variable $Age_t \leq 60$, which is equal to 1 for observations belonging to CEOs who are under 60 years of age in year t. As suggested in the empirical labor literature, older workers have worse employment opportunities (Hutchens (1988)). Thus, $Age_t \leq 60$ should be positively correlated with outside option q. We obtained the CEOs' birth years from the same sources of biographical information as above, Marquis' Who's Who and ZoomInfo.com.

In accordance with prior work (Baker, Jensen, and Murphy (1988), Rosen (1992), Core, Holthausen, and Larcker (1999)) we include as controls in our executive pay regressions proxies for firm complexity and growth opportunities. Firms operating in more complex environments and facing better investment opportunities benefit more from hiring better managers, to whom they have to offer higher wages in equilibrium. This effect is not captured in our model, which deals with one firm only. However, as we test the model using data from a cross-section of firms, it is necessary to include these additional controls. We use the industry-specific[27] mean return on assets during 1993-2005, as well as year fixed-effects, to control for growth opportunities. We use the company size proxied by the value of sales ($Sales_t$) to control for firm complexity.

[27]We assign companies to industries each year based on two-digit SIC codes extracted from Execucomp.

5 Empirical results

In this section we test the empirical implications of propositions 2 through 5 using pooled OLS regressions.[28] The summary statistics of the variables of interest, namely measures of executive compensation c and p, and our proxies for model parameters s, δ, γ, μ, q as well as control variables, are shown in Table 1. In our empirical tests we only include firm-year observations for which the CEO has been in their current position for at least one year, and who obtained the CEO job after 1980.[29]

Proposition 2 states that perks p are increasing in δ (noise in the production process), q (the manager's outside option) and are decreasing in γ (uncertainty in the manager's ability).

We find strong support for these predictions. We use several measures for the value of hidden pay p, captured by left-hand side variables $Options_t$, $RestrictedStock_t$, $OtherAnnual_t$ and $TotalPerks_t$. The results are shown in Tables 2, 5, 6, and 7 respectively. As explanatory variables we include proxies for CEO inferred ability μ, entrenchment s, outside option q, uncertainty in managerial ability γ and the noise in production process δ. Since executive pay varies across industries, with firm size and over time (Murphy (1999)), we include as control variables the industry-specific return on assets ($\mu(ROA)^{Industry}$) during 1992-2005, the firm size captured by the dollar value of its sales ($Sales_{t-1}$)[30] and year-fixed effects.

Proposition 2 implies that the loadings on $\sigma(ROA_{i,t} - ROA_{IND,t})$ (our proxy for δ), and on $Inverse\ of\ Firm{-}Specific\ Capital$, $HasMBA$ and $Age_t \leq 60$ (our proxies for q) should be positive, while the loading on $InverseCEOTenure_t$ (our proxy for γ) should be negative. As shown in Tables 2, 5, 6 and 7 these predictions are matched by the data, and the effects reach statistical significance in most specifications. The model's predictions receive greatest support when the proxies for hidden pay are $Options_t$ and $TotalPerks_t$.

[28] In all empirical analyses we correct standard errors to allow for heteroskedasticity and correlations among the error terms in observations that belong to the same company-CEO pairing. Our clustering unit is the Execucomp variable *co_per_r* which uniquely defines company-CEO pairs.

[29] Empirical analyses not included here indicate that there are large differences in the composition of compensation offered to executives who became CEOs before and after 1980. In particular, individuals who became CEOs before 1980 receive through time a much smaller fraction of their pay in the form of options and stock, even after controlling for inferred ability, industry, fixed-effects and demographic characteristics of the CEO. We investigate this effect further in another paper.

[30] Using $Size_t$ as a predictor of CEO pay in year t would be problematic, since both CEO pay and firm sales in year t depend on the company's performance that year. Hence we use the lagged value of sales to proxy for firm size.

The least strong results are obtained when using $OtherAnnual_t$ and $RestrictedStock_t$ as proxies for perks.

The economic significance of the effect of the model parameters on the size of hidden pay is considerable. An increase of one standard deviation (see Table 1) in the value of $\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ) corresponds to increases of $499,003 in the value of options granted in a given year to the CEO, and $473,827 in the total value of his hidden compensation. An increase of one standard deviation in the value of $Inverse\ of\ Firm - Specific\ Capital$ (q) corresponds to increases of $302,197 in the value of options granted in a given year to the CEO, and $339,291 in the total value of his hidden compensation (Panel B, tables 2 and 7). CEOs who have an MBA degree (high q) get $395,232 more in option grants per year, and $508,116 in total perks (Panel C, Tables 2 and 7). Similarly, if a CEO is under the age of 60 (high q), he receives $359,841 more in options and $367,854 in total hidden pay per year, compared to a similar CEO who is older than 60 years (Panel D, Tables 2 and 7). Also, an increase of one standard deviation in the value of $InverseCEOTenure_t$ (γ) corresponds to a decrease of $193,766 and $227,876 in option grant and total perks value (Panel B, Tables 2 and and 7). To put these effects into perspective, the mean yearly option grant and value of total perks are $2.22 million and $2.79 million, respectively (table 1).

As additional tests of proposition 2 we examine whether our parameters of interest, δ, q and γ, also correlate in accordance to our model with the likelihood of companies engaging in options backdating, as well as with the estimated size of backdating gains. Options backdating fits the definition of hidden pay, because shareholders are not aware in a timely fashion of the additional value of the compensation received by the CEO as a result of this behavior.

The logit models in Table 3 show[31] that most of the predictions of proposition 2 are also confirmed when we use $Backdating_t$ as a proxy of hidden pay. $\sigma(ROA_{i,t} - ROA_{IND,t})$, our proxy for δ, is a positive and significant predictor of the likelihood of backdating, in all specifications. Two of the proxies for q, $Inverse\ of\ Firm - Specific\ Capital$ and $Age_t \leq 60$, are also positively and significantly correlated with backdating. The third measure of q, $HasMBA$, has a negative impact on backdating, which is opposite to the predicted effect. Finally, we find that γ ($InverseCEOTenure_t$) is negatively correlated

[31] The number of observations in Table 3 is much smaller than that in all other tables, because the backdating data only covers years 1996-2002, half of our main sample (1992-2005).

to backdating likelihood in all specifications in Table 3, as the theory suggests, but the effect does not reach statistical significance.

Using the estimated amount of gains from backdating options ($BackdatingAmount_t$) as a proxy for perks, we also find strong support for proposition 2, as shown in table 4. The gains from backdating increase with our proxies for the outside option q and the noise in the production process δ, and decrease with γ, and all these effects are economically and statistically significant. The magnitude of the effects is similar to those obtained when we use options granted as the proxy for perks (see Table 2).

Proposition 3 and 4 state that changes in the noise in production outcome δ, in the manager's outside option q, and in the uncertainty in the manager's ability γ impact observable compensation c less than they impact perks p, in absolute magnitude.

We test these predictions by comparing the absolute value of the coefficients on parameters δ, q and γ in the regressions in Tables 7 and 8, where the dependent variables are the total value of hidden pay ($TotalPerks_t$) and total value of observed pay ($TotalObservableCompensation_t$), respectively.

Because the right-hand side variables are the same in the two models, the regression results in Tables 7 and 8 are exactly identical to those obtained by estimating the two models simultaneously, in a seemingly-unrelated regressions framework. Estimating the system

$$\begin{cases} c = X_c\beta_c + \tilde{\epsilon_c} \\ p = X_p\beta_p + \tilde{\epsilon_p} \end{cases}$$

using GLS yields best linear unbiased estimates. Estimating each model separately leads to inefficient estimates when the disturbances $\tilde{\epsilon_c}$ and $\tilde{\epsilon_p}$ are correlated. However, when $X_c = X_p$, that is, the two models have the same right-hand side variables, the seemingly-unrelated regressions approach is equivalent to estimating each model separately using OLS, and allows us to do hypothesis-testing on the coefficients β_c and β_p during the joint estimation. We use standard Wald tests to compare the effect sizes of variables in the model on the size of observable pay, c and on perks p.

The coefficient on $\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ) in table 8 is not significantly different from zero in any specification, and therefore δ does not affect observable pay. As noted above, δ is a positive and significant predictor of total perks: the coefficient on $\sigma(ROA_{i,t} - ROA_{IND,t})$ in table 7 is 32.22, with a t-statistic of 3.84 (Panel B, similar values in the other specifications). Hence, we are able to confirm the prediction of proposition 4 that the noise in the production process δ has more impact in absolute terms on perks than on observable pay.

We also match the predictions of proposition 4 regarding the differential impact of the CEO's outside option q on observable compensation and on hidden pay. Our first proxy for q, $Inverse\ of\ Firm-Specific\ Capital$, is not related to observable pay (the coefficients on this variable in Panels A and B of table 8 are not significantly different from zero), but it is a significant and positive predictor of hidden pay (Panels A and B of Table 7). Thus, the prediction of proposition 4 regarding q is met in the data, for this proxy of q. We also confirm the prediction using $HasMBA$ and $Age_t \leq 60$ as proxies for q. The coefficient on $HasMBA$ is 508.116 (t-stat=2.61) in the total perks regression in table 7, and 118.354 (t-stat=1.98) in the total observable pay regression in Table 8. A Wald test performed on these two coefficients shows that we can reject the hypothesis that they are equal ($Chi-squared(1)$ =5.03, p=0.0249). Similarly, the coefficient on $Age_t \leq 60$ is 367.854 (t-stat=2.36) in the total perks regression in table 7, and -171.809 (t-stat=-3.22) in the total observable pay regression in Table 8. The hypothesis that the two coefficients are equal can also be rejected, as indicated by the corresponding Wald test statistic ($Chi-squared(1)$ =15.31, p=0.0001).

Finally, we also match the predictions of proposition 4 regarding the higher absolute value impact of γ on perks compared to observable pay. This holds for all specifications in tables 7 and 8. For instance, using the specification in Panel D of these tables, the coefficient on our proxy for γ, $InverseCEOTenure_t$, is -1328.872 (t-stat=-2.48) in the total perks regression in table 7 and -399.272 (t-stat=-2.49) in the total observable pay regression in Table 8. A Wald test peformed on these two coefficients allows us to reject the hypothesis that they are the same ($Chi2(1)$ =3.79, p=0.051).

Proposition 5 states that managerial observable pay c increases with inferred ability μ and entrenchment s.

We find strong evidence to support this prediction. As shown in the pooled OLS regressions in Tables 8, 9 and 10 total observed CEO compensation (salary plus bonus), and the values of the bonus and salary separately, are positively correlated with both measures of inferred ability, and the effect is statistically significant at conventional levels for the $MeanInferredAbility_{t-1}^{RET}$ proxy for μ, which measures ability based on stock returns. Increasing that measure by one standard deviation increases $TotalObservableCompensation_t$ by \$100,533, $Salary_t$ by \$13,909 and $Bonus_t$ by \$87,067 (Panel B in Tables 8, 9 and 10, similar effects for the specifications in Panels C and D.) When we measure inferred ability μ using performance based on company ROA, the statistical significance of these effects diminishes (Panel A in Tables 8, 9 and 10).

The predicted positive link between entrenchment s and observable pay is also strongly supported by the data. In tables 8, 9 and 10 our proxy for s, $Entrenchment_t$, is a positive and significant predictor of total observable compensation, and of each of its components, salary and bonus. Increasing $Entrenchment_s$ by one standard deviation corresponds to an increase of \$151,145 in $TotalObservableCompensation_t$, \$60,176 in $Salary_t$ and of \$92,069 in $Bonus_t$ (Panel B of Tables 8, 9 and 10, similar effects in all other specifications.) These effects are large, on the order of 10% of the average values of total observable pay, salary and bonus (\$1.55 million, \$0.70 million, and \$0.85 million, respectively, as shown in table 1)

We find that our control variable that should proxy for industry growth opportunities ($\mu(ROA)^{Industry}$) is positively correlated with the value of observed compensation, but not related to most types of perks. Moreover, in all empirical specifications we match the findings in the extant literature regarding the link between firm complexity (as proxied by $Sales_{t-1}$) and CEO pay. Thus, executives of larger firms, and of firms in industries that have experienced higher returns on assets during our sample period (1993-2005) have enjoyed significantly higher compensation.

To summarize, we find empirical evidence in support of the predictions of our model. Certain variables (δ, q and γ) affect perks much more than observable compensation, and others (μ, s) impact observable pay. The effects we find are of significant magnitude in light of the average value of compensation for the CEOs in our sample.

6 Discussion

6.1 The Magnitude of Executive Compensation

A central question addressed by much recent research on executive compensation has been why has this compensation risen so sharply in recent years.[32] For example, explanations include higher pay as a manner of better incentives(Holmstrom and Kaplan (2001) and

[32] Frydman and Saks (2005) find an increase in CEO pay that is about 8-fold over the last 25 years. Jensen, Murphy, and Wruck (2004) and Murphy (1999) document similar increases in pay.

Inderest and Mueller (2005)), the increased importance of general skills (Frydman (2005) and Murphy and Zabojnik (2004)), changes in communication technology (Garciano and Rossi-Hansberg (2006)), increased entrenchment (Bebchuk and Fried (2003) and Bebchuk and Fried (2004)) and increased value due to rising firm size (Gabaix and Landier (2006)). While this question is not the primary motivation of our work (i.e., we believe our approach would be interesting even if the level of CEO pay had been stagnant), given the importance of this question, it is worth discussing how this rise in pay can be explained by our model, and how this relates to some of the existing literature.

Our model offers two different possible explanations for such a rise in pay, notably an increase in entrenchment or an increase in the maximum allowable hidden pay $\bar{p}$. The first explanation, an increase in entrenchment, is familiar to other stories where managers have some influence over their own pay (Bebchuk and Fried (2004)). However, there are a few specific points from our setting that are worth emphasizing, so we discuss this here briefly. The second explanation, an increase in the maximum allowable hidden pay $\bar{p}$, in new to our model and we will discuss this in more detail below.

One objection to the explanation of increased entrenchment is that managerial firings have increased concurrent with the rise in executive pay. This is taken to be inconsistent with the hypothesis of greater entrenchment. However, in a setting where managers are choosing their own pay subject to partial entrenchment, the relationship between entrenchment and firings is ambiguous. In our model, a manager extracts rents from greater entrenchment by paying himself more, which in turn increases the likelihood of firing. In effect, the extra leeway that the manager has from an exogenous increase in entrenchment is spent on more salary, rather than on less firing. Consequently, while there is a clear relationship between entrenchment and salary, entrenchment and firings need not be related.[33]

Another possibility is that entrenchment increases proportionally with firm size. For example, entrenchment may derive from a cost of adjustment for a new manager, or from a loss in firm reputation – both of which are plausibly proportional to size. If, however, at the same time the benefits from managerial ability are proportionate to size as well, it

[33] To be precise, in the model above in Sections 2 and 3, all managers above a threshold pay themselves as much as they can without triggering a firing, so the only prediction regarding firing is that only the worst performing managers get fired. However, further interesting implications for firings follow with the extension of adding noise to the level of entrenchment: that is, by making E_t noisy. This extension is discussed immediately below in subsection 6.2.

would follow from a model like ours that firings remain constant, while managerial salary increases in proportion with size. Intuitively, managers with low inferences would still be fired, whereas good managers would both be worth proportionately more relative to a new draw and would have proportionately more entrenchment. This would translate into compensation that increases with firm size.

Indeed, an interesting question in this regard is whether our framework can generate a prediction for the cross-sectional or cross-time variation in executive pay similar to that of (Gabaix and Landier (2006)). In their paper, they posit: 1) that the value of a CEO is related to his talent times firm size; 2) a general talent distribution for managers derived from extreme value theory; and 3) a value-maximizing managerial labor market where managers are matched efficiently with firms (assortative by talent-size) and are paid an equilibrium wage. From this they find a distribution for executive wages that strikingly corresponds with cross-sectional and cross-time patterns in the data. It is plausible, however, that a similar distribution would follow from a model with their assumptions on the value and distribution of managerial talent, with a managerial-optimal mechanism determining compensation in place of their value-maximizing mechanism. This is because in our setting, managers still do get compensated for their inferred ability (as in a value-maximizing mechanism), but also receive further compensation for their entrenchment.

The second explanation offered by our model for the rise executive pay is quite different from the familiar explanation of increased entrenchment. In particular, if total allowable hidden compensation $\bar{p}$ increases in our model, so does overall compensation. This result in our model is not simply a consequence of relaxing a binding constraint on pay; in fact, first period compensation increase in $\bar{p}$ for a given manager even if this constraint is not binding for him. Rather, executive compensation rises for all types in $\bar{p}$ because when $\bar{p}$ rises, the replacement manager of type 0 obtains higher compensation in equilibrium. This is true both in period 1 (when the replacement manager can be given 0 salary and obtains all compensation through perks), and in period 2 (when perk consumption cannot be constrained.[34] The compensation that other managers can obtain is always relative to such a potential replacement; managers' set their own pay to be as high as possible subject to not being replaced. Hence, all managers benefit from an increase in $\bar{p}$ because this gives the replacement manager greater scope for compensation, and the

[34]Note that for this result, it is critical that the IR constraint for such a manager is not binding, as discussed above. This is of course one of key features that distinguish our model from others on executive compensation.

retention decision for all other managers is made with respect to how they compare with such a replacement.

It is also worth remarking that in a value-maximizing framework, one would not expect any relationship between $\bar{p}$ and total managerial compensation. First, one would need some explanation why hidden pay is used at all in such a setting. But even given such an explanation (say, for tax purposes), the ability to increase hidden compensation should correspond directly with a decrease in observable compensation if compensation is being set to maximize firm value. That is, even if for some reason a hidden form of pay is more efficient than direct salary, a value-maximizing board should still give a total package that holds a manager to his IR constraint. In contrast, we have argued above that it seems implausible to us that most top executives are operating under a binding IR constraint. (That is, that most CEOs are paid with a package that leaves them indifferent to working for their firm or not.) Since the IR constraint is not binding in our setting, total compensation can change more readily with changes in $\bar{p}$.

While only anecdotal in nature, we would argue that the scope for hidden managerial pay has increased significantly over the last couple of decades. Our view here derives from a conception of boards that differs a bit from the notion of complete managerial capture. Suppose instead that boards are willing to rubber-stamp the CEO's pay request, unless they deem the request is something that could lead to a legal challenge. As such, boards might be reluctant to give compensation in a form that is not well accepted or well-justified, under the legal doctrine of the prudent man rule. Under this view, the widespread acceptance and adoption of various forms of incentive pay over the last 20 years have made certain forms of hidden pay more acceptable, and left managers with much greater scope to propose hidden pay that is acceptable to the board.[35]

Consistent with this interpretation, Murphy (1999) documents that while executive pay has increased dramatically, the largest share of this increase is not an increase in salary, but in options and other incentive compensation. Our model predicts that an increase in $\bar{p}$ can translate both into an increase in hidden pay and in salary. However, it suggests that a greater availability of the ability to pay compensation in a indirect form has played a critical role in the increase in overall compensation.

[35]For example, a number of board members have defended their firms' practices of backdating options by noting that such activity was commonplace and widely accepted.

6.2 Managerial Firings

In the model where managerial entrenchment is deterministic, unexpected firings do not occur. Managers below a threshold of inferred ability cannot prevent a firing, and all other managers make sure to keep their compensation to a level where they will not be fired. However, if some noise is added to the retention decision, for example, through the level of entrenchment, our model also yields interesting implications on firings.

Let entrenchment now be given by

$$E_t = s_t - \beta c_t + \tilde{\epsilon}_t. \tag{17}$$

The new term $\tilde{\epsilon}_t$ is a random noise distributed as $\tilde{\epsilon}_t \sim N(0, \xi^2)$, independent across time and with other noise terms in the model. The realization of this noise term is unknown to both the manager and shareholders when the manager sets c_t, and is observed by both prior to the shareholders' replacement decision. This shock in entrenchment allows for the possibility that managers will be surprised by shareholders' replacement decisions, and consequently, unexpected firings will occur at times.

The implications for such added noise on executive compensation and on firings is straightforward. Consider first the second period. Without any noise, managers set salary according to equation (3), which was the highest salary they could pay themselves without triggering a firing. Adding the noise term $\tilde{\epsilon}_t$ to entrenchment makes managers more cautious; now any increase in pay increases the probability they will be fired. In the final period, a manager loses his salary and perks from this period alone if fired. This cost is increasing in the manager's type. Furthermore, the benefit of raising salary c_2 more is the same for all types, and the increased probability of falling below the threshold is the same for all managers for an equivalent increase above the salary given by equation (3). Consequently, since higher types have more to lose, they will be more cautious with salary s_2, and will therefore be fired less in the second period.

Similarly, the same effect will be present in period 1. Now even more so than in period 2, higher type managers have more to lose from being fired. Hence when choosing the optimal level s_1, they will be more cautious about the possibility of getting a bad draw of $\tilde{\epsilon}_t$, and will therefore be more cautious in setting first period salary. Once again, this will lead to less variation is salary across types than when entrenchment is deterministic, and firings that decrease in managerial type. The overall pattern of firings predicted by the model would then be that all managers whose inferred types are below a threshold that is well below the average replacement manager are fired, and above this level, there is some probability that any manager will be fired, but this probability decreases in type. Translated to recent performance, we would then expect to see a strong negative relationship between firing and performance at the very bottom tail of the performance distribution, and a weak negative relationship above this bottom tail.[36] This seems to match well prior evidence on CEO firings (Warner, Watts, and Wruck (1988)).

7 Conclusion

We develop a new framework for considering executive compensation, based on the notion of CEOs setting their own pay subject to shareholder constraints, rather than shareholders setting CEO pay subject to informational constraints. This "managerial optimal" perspective can immediately rationalize the importance of hidden pay and yields a number of new empirical implications regarding hidden pay, observable pay and managerial retentions. The model matches a number of anecdotal facts well and addresses initial questions about standard models of executive pay. Our empirical analysis, conducted in a large sample of companies during 1993-2005, finds support for the most of the new predictions yielded by the model.

[36] It is worth emphasizing that this discussion relies heavily on the assumption of a fixed outside option for all managers. If higher ability managers also had equivalently higher outside options, then they would no longer have more to lose from firings, and these results would not follow.

8 Appendix

Proof to Proposition 2: If first period perk consumption has an interior solution, then p satisfies equation (12). It is immediate to see that the right hand side of (12) decreases in β and q and increasing in $\bar{p}$, and consequently, equilibrium p increases in β and q and decreases in $\bar{p}$ since V is concave.

Showing the result for δ and γ takes more work, given the dependence of σ on δ and γ. In showing these results, we make use of the fact that types $\mu < -s$ will be fired at the beginning of period 1.[37]

We must show that the right hand side of equation (12) is decreasing in δ. Note first that

$$\frac{d\sigma}{d\delta} = \frac{\frac{1}{2}(\delta^2+\gamma^2)^{-\frac{1}{2}}2\delta\gamma^2}{\gamma^2+\delta^2} = -\frac{\delta\gamma^2}{(\delta^2+\gamma^2)^{\frac{3}{2}}} < 0. \tag{18}$$

Differentiating the right hand side of (12) with respect to δ yields:

$$\begin{aligned} &\frac{d\sigma}{d\delta}\left[\frac{1}{\gamma^2}(V(\bar{p})-q)\phi\left(\frac{-s-\mu}{\sigma}\right)+\frac{1}{\gamma^2}\left(\frac{-s-\mu}{\sigma}\right)^2\right] \\ -\quad &\frac{1}{1+\beta}\left[2\delta\frac{\gamma^2}{\delta^2+\gamma^2}\left(1-\Phi\left(\frac{-s-\mu}{\sigma}\right)\right)-\frac{\gamma^2}{\delta^2+\gamma^2}\phi\left(\frac{-s-\mu}{\sigma}\right)\frac{s+u}{\sigma^2}\frac{d\sigma}{d\delta}\right]; \end{aligned}$$

which equals

$$\begin{aligned} =\quad &\frac{d\sigma}{d\delta}\left[\frac{1}{\gamma^2}(V(\bar{p})-q)\phi\left(\frac{-s-\mu}{\sigma}\right)\right]\left[1+\frac{(s+\mu)^2}{\sigma^2}\right] \\ -\quad &\frac{\delta\gamma^2}{(1+\beta)(\delta^2+\gamma^2)^2}\left[2\left(1-\Phi\left(\frac{-s-\mu}{\sigma}\right)\right)+\left(\frac{-s-\mu}{\sigma}\right)\phi\left(\frac{-s-\mu}{\sigma}\right)\right]. \end{aligned} \tag{19}$$

The first line of (19) is negative, since both bracketed terms are positive (recall that $V(\bar{p}) > q$), and (18) indicates that $\frac{d\sigma}{d\delta}$ is negative. The second line of (19) is also negative, since the function $r(x) \equiv (1-\Phi(x)) + x\phi(x)$ is greater than 0 for all x.[38] Consequently,

[37] This follows from straightforward, albeit tedious, calculations not included here. The intuition for this result is straightforward. Recall that this result holds (and is trivial to show) in period 2. Given this, it is not surprising that this also holds in period 1: the incentive to fire are greater in period 1 given there are two periods of performance that remain, and firing costs are the same as before. In fact, the firing threshold in period 1 is greater than that in period 2, but we will only need the fact that it exceeds $-s$.

[38] This can be seen immediately by noting that $\lim_{x\to-\infty} r(x) = 1$, $\lim_{x\to\infty} r(x) = 0$, and $r'(x) =$

the right hand side of (12) is decreasing in δ, and therefore, perks are increasing in δ.

A similar argument shows the desired relationship for γ.

If instead perks are given by a corner solution, it follows trivially that perks may remain unchanged (at 0 or $\bar{p}$) with changes in any of these variable. And if $p = \bar{p}$ is binding, it also follows trivially that relaxing this constraint will lead to an increase in p. $\square$

Proof to Proposition 4: Using equation (14) and letting x represent and of the variables δ, γ, q, and $\bar{p}$, it follows that: all affect equilibrium salary of type μ only through their effects on the first period perks that this type will be anticipated to pay relative to those of a replacement manager; that is, through their effects on $P_1(0) - P_1(\mu)$. Hence for a change in any of these variables, the corresponding change in first period salary will be given by:

$$\frac{dc}{dx} = \frac{1}{1+\beta}\left(\frac{dP_1(0)}{dx} - \frac{dP_1(\mu)}{dx}\right). \tag{20}$$

Using Proposition 2, and also noting that this proposition indicates that P_1 is monotonic in all four of the variables δ, γ, q, and $\bar{p}$, part 1 and 2 of the proposition follow immediately.

We show part 3 of the proposition for q and $\bar{p}$. Results follow in a similar manner for δ and γ but only with tedious calculations that are omitted here. For q, equation (20) indicates that we must show that

$$\frac{1}{1+\beta}\left(\frac{dP_1(0)}{dq} - \frac{dP_1(\mu)}{dq}\right) \leq \frac{dP_1(\mu)}{dq};$$

or

$$\frac{dP_1(0)}{dq} \leq (2+\beta)\frac{dP_1(\mu)}{dq}. \tag{21}$$

Using (12), it follows that

$$\frac{dP_1(\mu)}{dq} = -\frac{1}{V''(p(\mu))}\frac{\gamma^2}{\delta^2+\gamma^2}\frac{1}{\sigma}\phi\left(\frac{-s-\mu}{\sigma}\right).$$

Since $V''(p) < 0$, and by assumption is constant in the proposition, equation (21) will be satisfied provided that

$$\phi\left(\frac{-s}{\sigma}\right) \leq (2+\beta)\phi\left(\frac{-s-\mu}{\sigma}\right); \tag{22}$$

$-x\phi^2(x) < 0$.

as stated in the Proposition.

For $\bar{p}$, using (12), it follows that

$$\frac{dP_1(\mu)}{d\bar{p}} = \frac{1}{V''(p(\mu))} \frac{\gamma^2}{\delta^2+\gamma^2} \frac{1}{\sigma} \phi\left(\frac{-s-\mu}{\sigma}\right) V'(\bar{p}). \tag{23}$$

Since p is decreasing in $\bar{p}$, (20) indicates that we must show that

$$\frac{dP_1(0)}{d\bar{p}} \geq (2+\beta)\frac{dP_1(\mu)}{d\bar{p}},$$

which follows from (23) if inequality (22) is satisfied.

Finally it is worth remarking that while the restriction on types in inequality (22) implies both a lower and an upper bound for μ, the lower bound will not be relevant as types below this bound will be fired at the beginning of the first period. To see this, note that a manager will be fired in the first period if salary c as specified by equation (14) is negative. And since Proposition 5 indicates that $P_1(0) - P_1(\mu) < 0$, for all $\mu < -s$, it follows that all types $\mu < -s$ will be below the first period firing threshold. Since the lower bound for inequality (22) is below $-s$, the desired result follows.

□

Proof to Proposition 5: Differentiating the first order condition (12) with respect to type μ yields:

$$\begin{aligned}
\frac{dp}{d\mu} &= \frac{1}{V''(p)} \frac{\gamma^2}{\delta^2+\gamma^2}\left[\frac{1}{\sigma^2}\phi\left(\frac{-s-\mu}{\sigma}\right)\left(\frac{-s-\mu}{\sigma}\right)(V(\bar{p})-q) + \frac{1}{\sigma(1+\beta)}\phi\left(\frac{-s-\mu}{\sigma}\right)\right] \\
&= -\frac{\phi\left(\frac{-s-\mu}{\sigma}\right)}{V''(p)\gamma^2}\left[\left(\frac{s+\mu}{\sigma}\right)\frac{1}{\sigma}(V(\bar{p})-q) - \frac{1}{1+\beta}\right]. \qquad (24)
\end{aligned}$$

It follows immediately from (24) that first period perks are increasing in type if and only if

$$s+\mu \geq \frac{\sigma^2}{(1+\beta)(V(\bar{p})-q)};$$

as stated in the proposition. Note that the same is true for $\frac{dp}{ds}$, as s enters into equation (12) in the identical manner as μ.

From equation (14) it follows that

$$\frac{dc}{ds} = 1 + \frac{dP_1(0)}{ds} - \frac{dP_1(\mu)}{ds}, \tag{25}$$

and

$$\frac{dc}{d\mu} = \frac{1+2\beta}{1+\beta} + \frac{dP_1(0)}{d\mu} - \frac{dP_1(\mu)}{d\mu}. \tag{26}$$

Now provided that γ is large enough, it follows from equation (24) and an equivalent equation for $\frac{dp}{ds}$ that the terms $\frac{dP_1(\mu)}{d\mu}$, $\frac{dP_1(0)}{d\mu}$, $\frac{dP_1(\mu)}{d\mu}$, and $\frac{dP_1(0)}{d\mu}$ can all be bounded below 1. (That is, when total managerial slack increases by 1, perks for any type of manager increase by no more than 1.) And then, the first terms on the right hand side of equations (25) and (26) dominate, ensuring that $\frac{dc}{d\mu}$ and $\frac{dc}{ds}$ are both positive. $\square$

References

Baker, G., M. Jensen, and K. J. Murphy (1988). Compensation and incentives: Practice vs. theory. *Journal of Finance 43*(3), 593–616.

Bebchuk, L. and J. Fried (2003). Executive Compensation as an Agency Problem. *Journal of Economic Perspectives* **17(3)**, 71–92.

Bebchuk, L. and J. Fried (2004). *Pay without Performance*. Cambridge, MA and London, England: Harvard University Press.

Bebchuk, L. A. and R. J. Jackson (2005). Executive pensions. *Journal of Corporation Law*.

Bizjak, J., M. Lemmon, and R. Whitby (2006, November). Options backdating and board interlocks. *Working paper*.

Carter, M. and L. Lynch (2001). An examination of executive stock option repricing. *Journal of Financial Economics* (61), 207–225.

Core, J., R. Holthausen, and D. Larcker (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics*.

DeGroot, M. H. (1970). *Optimal Statistical Decisions*. New York: McGraw Hill.

Fisman, R., R. Khurana, and M. Rhodes-Kropf (2005). Governance and CEO turnover: Do something or do the right thing? *Working paper. Columbia University*.

Frydman, C. (2005). Rising Through the Ranks. The Evolution of the Market for Corporate Executives, 1936-2003. *Harvard University, Working Paper*.

Frydman, C. and R. Saks (2005). Historical Trends in Executive Compensation, 1936-2003. *Harvard University, Working Paper*.

Gabaix, X. and A. Landier (2006). Why Has CEO Pay Increased So Much? *Working Paper*.

Garciano, L. and E. Rossi-Hansberg (2006). Organization and Inequality in a Knowledge Economy. *Quarterly Journal of Economics*.

Garvey, G. and G. Hanka (1999). Capital Structure and Corporate Control: The Effect of State Antitakeover Laws on Firm Leverage. *Journal of Finance* **54**, 519–546.

Gompers, P., J. Ishii, and A. Metrick (2003). Corporate governance and equity prices. *Quarterly Journal of Economics*.

Grossman, S. and O. Hart (1982). *The Economics of Information and Uncertainty*, Chapter "Corporate Financial Structure and Managerial Incentives". University of Chicago Press.

Grossman, S. and O. Hart (1983). An Analysis of the Principal-Agent Problem. *Econometrica* **51**, 7–45.

Hallock, K. F. (1997, September). Reciprocally interlocking boards of directors and executive compensation. *Journal of Financial and Quantitative Analysis 32*(3), 331–344.

Hart, O. (2001). Financial Contracting. *Journal of Economic Literature* **34**, 1079–1100.

Hermalin, B. E. and M. S. Weisbach (1998). Endogenously chosen boards of directors and their monitoring of the CEO. *The American Economic Review*, 96–118.

Hermalin, B. E. and M. S. Weisbach (2003, April). Boards of directors as an endogenously determined institution: A survey of the economic literature. *FRBNY Economic Policy Review*, 7–26.

Heron, R. A. and E. Lie (2006). Does backdating explain the stock price pattern around executive stock option grants? *Journal of Financial Economics, forthcoming*.

Holmstrom, B. (1979). Moral Hazard and Observability. *The Bell Journal of Economics* **13**, 324–40.

Holmstrom, B. (1982). *in Essays in Economics and Management in Honor of Lars Wahlbeck*, Chapter Mangerial incentive problems: A dynamic perspective. Helsinki, Finland. Republished in the Review of Economic Studies, 1999, 66, 169-182.

Holmstrom, B. and S. Kaplan (2001). Corporate Governance and Merger Activity in the U.S. *Journal of Economic Perspectives* **15(2)**, 121–144.

Holmstrom, B. and P. Milgrom (1987). Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica* **7**, 303–328.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization* **7**, 24–52.

Holmstrom, B. and J. Ricart i Costa (1986). Managerial Incentives and Capital Management. *Quarterly Journal of Economics* **101**, 835–860.

Hutchens, R. M. (1988, October). Do job opportunities decline with age? *Industrial and Labor Relations Review 42*(1), 89–99.

Inderest, R. and H. Mueller (2005). Keeping the Board in the Dark: CEO Compensation and Entrenchment. *Working Paper*.

Jensen, M. (1986). Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review* **76(2)**, 323–329.

Jensen, M. and W. Meckling (1976). Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure. *Journal of Financial Economics* **3**, 360–395.

Jensen, M., K. J. Murphy, and E. Wruck (2004). Remuneration: Where we've been, how we got to here, what are the problems, and how to fix them. *Unpublished Paper*.

Lazear, E. and S. Rosen (1981). Rank-Order Tourmanments as Optimum Labor Contracts. *The Journal of Political Economy* **89**, 841–862.

Main, B., C. O'Reilly, and J. Wade (1995). The CEO, the Board of Directors , and Executive Compensation: Economic and Psychological Perspectives. *Industrial and Corporate Change* **95**, 292–332.

Mirrlees, J. (1974). *in Notes on Welfare Economics, Information, and Uncertainty*, Chapter Essays on Economic Behavior Under Uncertainty. North Holland.

Mirrlees, J. (1976). The Optimal Structure of Incentives and Authority Within an Organization. *Bell Journal of Economics* **7**, 105–131.

Murphy, K. J. (1999). Executive Compensation. *Handbook of Labor Economics* **3**.

Murphy, K. J. and J. Zabojnik (2004). CEO Pay and Appointments: A Market-based explanation for recent trends. *American Economic Review Papers and Proceedings* **94(2)**, 192–196.

Novais, W. (2003). Capital Structure Choice When Managers Are in Control: Entrenchment versus Efficiency. *Journal of Business* **76**, 49–82.

Prendergast, C. (1999, March). The Provision of Incentives in Firms. *Journal of Economic Literature* **37**, 7–62.

Rajan, R. G. and J. Wulf (2006). Are perks purely managerial excess? *Journal of Financial Economics* **79**, 1–33.

Rosen, S. (1992). *Contract Economics*, Chapter "Contracts and the market for executives". Backwell Publishers.

The Journal of Corporation Law (2005). Symposium on Bebchuk & Fried's Pay Without Performance. *The Journal of Corporation Law* **30**, 647–900.

Warner, J., R. Watts, and K. Wruck (1988). Stock prices and top management changes. *Journal of Financial Economics* **20**, 461–492.

Yermack, D. (1997). Good timing: CEO stock option awards and company news announcements. *Journal of Finance* **52**, 449–476.

Yermack, D. (2006a). Flights of fancy: Corporate jets, CEO perquisites, and inferior shareholder returns. *Journal of Financial Economics* **80**, 211–242.

Yermack, D. (2006b). Golden handshakes: Separation pay for retired and dismissed CEOs. *Working paper, NYU*.

Zwiebel, J. (1996). Dynamic Capital Structure under Managerial Entrenchment. *The American Economic Review* **86**, 1197–1215.

Table 1: Summary statistics and variable definitions

$Salary_t$ and $Bonus_t$ are the salary and bonus paid to the CEO during year t (data items *salary* and *bonus* in Execucomp). $TotalObservedCompensation_t$ is the sum of $Salary_t$ and $Bonus_t$. $TotalPerks_t$ is the sum of $Options_t$, $RestrictedStock_t$ and $OtherAnnual_t$. $Options_t$ is the aggregate value of the stock options granted to the executive during year t (data item *blk_valu*). $RestrictedStock_t$ is the value of restricted stock granted during the year to the CEO (data item *rstkgrnt*). $OtherAnnual_t$ is the value of other annual compensation not classified as salary or bonus, and includes perquisites and personal benefits, as well as other gains such as above market earnings on stock purchase plans (data item *othann*). $Backdating_t$ is an indicator equal to 1 for firm-year observations (for 1996-2002 only) for which options backdating is likely to have occured, according to Bizjak, Lemmon, and Whitby (2006). $BackdatingAmount_t$ is the estimated value of additional compensation obtained by the CEO by backdating stock options granted during year t. Compensation variables are expressed in \$ thousands and are windsorized at 99.5% at the upper tail. $MeanInferredAbility_t^{ROA}$ and $MeanInferredAbility_t^{RET}$ are the historical averages up to t of the yearly ranking of a CEO's firm relative to its peers, in terms of ROA, and stock return, respectively. $Entrenchment_t$ is the value of the Governance Index in Gompers, Ishii, and Metrick (2003). $Inverse\ of\ Firm-Specific\ Capital$ is the inverse of the # of years the executive worked for the company before becoming its CEO. $HasMBA$ and $Age_t \leq 60$ are indicators equal to 1 for firm-year observations where the CEO has an MBA degree, and where the CEO is younger than 60, respectively. $InverseCEOTenure_t$ is the inverse of the # of years the executive has been the firm's CEO. $\sigma(ROA_{i,t} - ROA_{Ind,t})$ is the standard deviation of the firm's excess ROA relative to the industry mean. $Sales_t$ is the firm's net annual sales (\$ millions). $\mu(ROA)^{Industry}$ is the industry-wide mean ROA.

Variable	Proxy for parameter	Execucomp data item	Obs.	Mean	Standard Deviation	Min	Max
$Salary_t$	c	*salary*	10048	\$697.00	\$322.85	\$0	\$2110.42
$Bonus_t$	c	*bonus*	10048	\$847.92	\$1227.71	\$0	\$8500.00
$TotalObservableCompensation_t$	c	*tcc*	10048	\$1550.92	\$1430.04	\$0	\$10000.00
$TotalPerks_t$	p	-	10004	\$2791.51	\$4855.37	\$0	\$47284.57
$OtherAnnual_t$	p	*othann*	10048	\$54.47	\$162.03	\$0	\$1437.53
$RestrictedStock_t$	p	*rstkgrnt*	10048	\$506.74	\$1561.28	\$0	\$11695.10
$Options_t$	p	*blk_valu*	10004	\$2229.73	\$4345.51	\$0	\$35589.46
$Backdating_t$	p	-	5307	0.07	0.25	0	1.00
$BackdatingAmount_t$	p	-	4087	\$435.00	\$1636.60	0	\$16198.18
$MeanInferredAbility_{t-1}^{ROA}$	μ	-	10007	0.54	0.24	.01	1.00
$MeanInferredAbility_{t-1}^{RET}$	μ	-	10048	0.52	0.18	.01	1.00
$Entrenchment_t$	s	-	10048	9.52	2.62	2.00	17.00
$Inverse\ of\ Firm-Specific\ Capital\ (q)$	q	-	5848	0.40	0.42	0.02	1.00
$HasMBA$	q	-	10048	0.27	0.44	0	1.00
$Age_t \leq 60$	q	-	10048	0.65	0.48	0	1.00
$InverseCEOTenure_t$	γ	-	10048	0.20	0.13	0.04	0.50
$\sigma(ROA_{i,t} - ROA_{Ind,t})$	δ	-	10048	15.73	14.71	0	52.99
$Sales_{t-1}$	control	*sales*	10048	\$5503.99	\$13780.08	\$0.10	\$263989.00
$\mu(ROA)^{Industry}$	control	-	10048	2.69	2.71	-5.41	19.10

Table 2: Predictors of Options Granted

The dependent variable, $Options_t$, is the aggregate value of the stock options granted to the executive during year t calculated using using the S&P Black Scholes method (data item *blk_valu* in Execucomp). All independent variables are defined in table 1. Years covered are 1992-2005. T-statistics in parentheses.

Dependent variable	$Options_t$			
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility^{RET}_{t-1}$ (μ)		1984.086	1976.873	2008.673
		(4.12)***	(5.36)***	(5.40)***
$MeanInferredAbility^{ROA}_{t-1}$ (μ)	1353.298			
	(3.56)***			
$Entrenchment_t$ (s)	8.826	–5.268	–23.410	–16.329
	(.27)	(–.16)	(–.87)	(–.61)
$Inverse\ of\ Firm-Specific\ Capital$ (q)	769.850	719.518		
	(3.14)***	(2.95)***		
$HasMBA$ (q)			395.232	
			(2.26)**	
$Age_t \leq 60$ (q)				359.841
				(2.54)**
$InverseCEOTenure_t$ (γ)	–1519.240	–1537.831	–972.580	–1200.148
	(–2.73)***	(–2.71)***	(–2.07)**	(–2.44)**
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	32.393	33.932	35.462	35.279
	(3.73)***	(4.55)***	(5.48)***	(5.46)***
$\mu(ROA)^{Industry}$	–42.060	–40.905	–66.041	–65.653
	(–1.06)	(–1.07)	(–2.09)**	(–2.08)**
$Sales_{t-1}$	.064	.063	.065	.066
	(4.68)***	(4.82)***	(5.53)***	(5.57)***
Adj. R^2	.100	.100	.100	.100
No. of obs	6856	6760	10004	10004

Included: year FEs

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Pooled OLS, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level

Table 3: Predictors of Options Backdating

Logit model for the likelihood of options backdating. Dependent variable is a dummy equal to 1 for firm-year observations for which the firm is likely to have engaged in backdating, according to the algorithm in Bizjak, Lemmon, and Whitby (2006). Years covered by backdating data are 1996-2002. All independent variables are defined in Table 1. T-statistics in parentheses.

Dependent variable		$Backdating_t$		
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility_{t-1}^{RET}$ (μ)		.0129	–.1117	–.1068
		(.03)	(–.31)	(–.30)
$MeanInferredAbility_{t-1}^{ROA}$ (μ)	–.5674			
	(–1.73)*			
$Entrenchment_t$ (s)	–.0660	–.0623	–.0923	–.0933
	(–2.25)**	(–2.10)**	(–3.92)***	(–4.04)***
$Inverse\ of\ Firm-Specific\ Capital$ (q)	.5023	.5259		
	(2.83)***	(3.00)***		
$HasMBA$ (q)			–.4408	
			(–2.93)***	
$Age_t \leq 60$ (q)				.3329
				(2.38)**
$InverseCEOTenure_t$ (γ)	–.4053	–.2901	–.0367	–.2867
	(–.66)	(–.47)	(–.08)	(–.59)
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	.0123	.0128	.0114	.0101
	(2.04)**	(2.10)**	(2.33)**	(2.10)**
$\mu(ROA)^{Industry}$	–.0001	.0048	–.0264	–.0343
	(–.00)	(.16)	(–1.04)	(–1.35)
$Sales_{t-1}$ X 10^{-3}	–.0176	–.0179	–.0160	–.0155
	(–2.24)**	(–2.24)**	(–2.17)**	(–2.15)**
$Options_t$ X 10^{-3}	.0341	.0350	.0399	.0375
	(3.38)***	(3.46)***	(4.80)***	(4.36)***
Pseudo R^2	.088	.086	.087	.085
No. of obs	2820	2790	4928	4928

Included: year FEs

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Logit model, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level

Table 4: Predictors of Amount of Backdating

The dependent variable, $BackdatingAmount_t$, is the estimated value of additional compensation obtained by the CEO by backdating stock options granted during year t. For each grant received by the CEO during year t we calculate $Blkvalu\_TranDate_t$, the Black Scholes value of the grant assuming that it was made on the declared transaction date as listed in Thompson Financial Insider Filings Database. We also calculate $Blkvalu\_SecDate_t$, the Black Scholes value of the grant assuming that the CEO actually received the options on the day the company disclosed the grant to the SEC, but used as strike the share price on the declared (and earlier) transaction date. To calculate both $Blkvalu\_TranDate_t$ and $Blkvalu\_SecDate_t$ we use the annual estimates for the firm's volatility and dividend yield as found in Execucomp (data items *bs_volatility* and *bs_yield*), as well as the annual risk free rates used by Execucomp to calculate their own values for executive option grants. The amount of compensation gained by the executive by backdating is defined as: $BackdatingAmount_t = (Blkvalu\_SecDate_t - Blkvalu\_TranDate_t) * 1_{Blkvalu\_SecDate_t > Blkvalu\_TranDate_t}$, where we winsorize the top 0.05% of values to eliminate the effect of outliers. All independent variables are defined in table 1. Years covered are 1996-2005. T-statistics in parentheses.

Dependent variable	$BackdatingAmount_t$			
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility^{RET}_{t-1}$ (μ)		706.782	519.346	527.128
		(2.80)***	(2.60)***	(2.63)***
$MeanInferredAbility^{ROA}_{t-1}$ (μ)	375.322			
	(2.19)**			
$Entrenchment_t$ (s)	–7.434	–12.722	–18.140	–15.715
	(–.47)	(–.80)	(–1.27)	(–1.10)
$Inverse\ of\ Firm-Specific\ Capital$ (q)	269.283	240.401		
	(2.56)**	(2.29)**		
$HasMBA$ (q)			92.945	
			(1.25)	
$Age_t \leq 60$ (q)				133.276
				(2.07)**
$InverseCEOTenure_t$ (γ)	–427.190	–443.765	–394.798	–488.501
	(–1.43)	(–1.48)	(–1.68)*	(–1.98)**
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	14.339	14.686	11.389	11.447
	(2.78)***	(3.06)***	(3.11)***	(3.11)***
$\mu(ROA)^{Industry}$	18.691	15.703	–5.719	–3.849
	(.81)	(.74)	(–.38)	(–.25)
$Sales_{t-1}$	.008	.008	.007	.007
	(1.92)*	(1.87)*	(1.99)**	(2.25)**
Adj. R^2	.058	.059	.048	.048
No. of obs	2676	2638	4087	4087

Included: year FEs * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Pooled OLS, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level

Table 5: Predictors of Restricted Stock Grant Compensation

The dependent variable, $RestrictedStock_t$, captures the dollar value of restricted stock granted during the year to the CEO (data item *rstkgrnt* in Execucomp). All independent variables are defined in Table 1. Years covered are 1992-2005. T-statistics in parentheses.

Dependent variable		$RestrictedStock_t$		
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility^{RET}_{t-1}$ (μ)		89.851	95.972	103.132
		(.80)	(.85)	(.91)
$MeanInferredAbility^{ROA}_{t-1}$ (μ)	3.575			
	(.03)			
$Entrenchment_t$ (s)	31.864	30.098	10.868	12.110
	(3.64)***	(3.52)***	(1.38)	(1.53)
$Inverse\ of\ Firm-Specific\ Capital$ (q)	82.180	80.425		
	(1.03)	(1.02)		
$HasMBA$ (q)			107.226	
			(1.62)	
$Age_t \leq 60$ (q)				24.574
				(.48)
$InverseCEOTenure_t$ (γ)	–206.263	–242.500	–111.733	–122.568
	(–.94)	(–1.13)	(–.63)	(–.71)
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	–2.828	–2.229	–.147	–.082
	(–1.25)	(–.96)	(–.07)	(–.04)
$\mu(ROA)^{Industry}$	6.705	8.067	15.362	15.723
	(.52)	(.58)	(1.07)	(1.10)
$Sales_{t-1}$	.026	.026	.028	.028
	(4.53)***	(4.49)***	(5.63)***	(5.71)***
Adj. R^2	.096	.095	.097	.096
No. of obs	6889	6793	10048	10048

Included: year FEs

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Pooled OLS, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level

Table 6: Predictors of Other Annual Compensation

The dependent variable, $OtherAnnual_t$, captures the dollar value of other annual compensation not classified as salary or bonus, and includes items such as perquisites and personal benefits, above market earnings on restricted stock, options or deferred compensation paid during the year but deferred by the CEO. It also includes tax reimbursements, and the value of the difference between the price paid by the CEO for company stock and the actual market price of the stock under a stock purchase plan not available to shareholders or employees of the company (data item *othann* in Execucomp). All independent variables are defined in Table 1. Years covered are 1992-2005. T-statistics in parentheses.

Dependent variable		$OtherAnnual_t$		
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility^{RET}_{t-1}$ (μ)		–13.558	–16.548	–16.427
		(–.93)	(–1.47)	(–1.47)
$MeanInferredAbility^{ROA}_{t-1}$ (μ)	–2.635			
	(–.17)			
$Entrenchment_t$ (s)	3.355	3.327	1.522	1.428
	(2.20)**	(2.16)**	(1.37)	(1.29)
$Inverse\ of\ Firm-Specific\ Capital$ (q)	7.864	7.376		
	(.81)	(.73)		
$HasMBA$ (q)			6.142	
			(.86)	
$Age_t \leq 60$ (q)				–15.662
				(–2.85)***
$InverseCEOTenure_t$ (γ)	–22.759	–23.027	–20.716	–9.424
	(–.93)	(–.94)	(–1.12)	(–.52)
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	.357	.481	.437	.466
	(.84)	(1.10)	(1.41)	(1.51)
$\mu(ROA)^{Industry}$	2.867	3.387	3.634	3.716
	(1.49)	(1.65)*	(2.46)**	(2.54)**
$Sales_{t-1}$	.002	.002	.002	.002
	(3.56)***	(3.55)***	(4.21)***	(4.18)***
Adj. R^2	.037	.037	.043	.045
No. of obs	6889	6793	10048	10048

Included: year FEs

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Pooled OLS, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level

Table 7: Predictors of Total Hidden Compensation

The dependent variable, $TotalPerks_t$, is the sum of $Options_t$, $RestrictedStock_t$ and $OtherAnnual_t$. All independent variables are defined in Table 1. Years covered are 1992-2005. T-statistics in parentheses.

Dependent variable	$TotalPerks_t$			
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility^{RET}_{t-1}$ (μ)		2073.243	2061.876	2101.010
		(4.04)***	(5.17)***	(5.23)***
$MeanInferredAbility^{ROA}_{t-1}$ (μ)	1363.513			
	(3.20)***			
$Entrenchment_t$ (s)	43.198	27.194	–11.561	–3.364
	(1.25)	(.76)	(–.40)	(–.12)
$Inverse\ of\ Firm-Specific\ Capital$ (q)	860.932	807.838		
	(3.17)***	(3.00)***		
$HasMBA$ (q)			508.116	
			(2.61)***	
$Age_t \leq 60$ (q)				367.854
				(2.36)**
$InverseCEOTenure_t$ (γ)	–1751.646	–1808.542	–1102.315	–1328.872
	(–2.77)***	(–2.84)***	(–2.13)**	(–2.48)**
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	29.942	32.220	35.745	35.656
	(3.17)***	(3.84)***	(5.02)***	(5.03)***
$\mu(ROA)^{Industry}$	–32.225	–29.194	–47.511	–46.682
	(–.71)	(–.64)	(–1.29)	(–1.26)
$Sales_{t-1}$	.092	.090	.095	.097
	(6.78)***	(6.90)***	(7.68)***	(7.68)***
Adj. R^2	.122	.123	.125	.124
No. of obs	6856	6760	10004	10004

Included: year FEs

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Pooled OLS, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level

Table 8: Predictors of Total Observable Compensation

The dependent variable, $TotalObservedCompensation_t$, is the sum of the CEO's salary ($Salary_t$) and bonus ($Bonus_t$) in fiscal year t (data items *tcc*, *salary* and *bonus* in Execucomp, respectively). All independent variables are defined in table 1. Years covered are 1992-2005. T-statistics in parentheses.

Dependent variable	$TotalObservableCompensation_t$			
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility^{RET}_{t-1}$ (μ)		549.361	501.354	505.890
		(4.95)***	(4.82)***	(4.87)***
$MeanInferredAbility^{ROA}_{t-1}$ (μ)	183.594			
	(1.59)			
$Entrenchment_t$ (s)	61.273	57.623	38.028	37.470
	(5.15)***	(4.79)***	(3.89)***	(3.81)***
$Inverse\ of\ Firm-Specific\ Capital$ (q)	60.913	43.714		
	(.70)	(.51)		
$HasMBA$ (q)			118.354	
			(1.98)**	
$Age_t \leq 60$ (q)				–171.809
				(–3.22)***
$InverseCEOTenure_t$ (γ)	–887.919	–924.272	–526.149	–399.272
	(–4.32)***	(–4.47)***	(–3.33)***	(–2.49)**
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	.617	1.416	2.241	2.610
	(.17)	(.36)	(.75)	(.88)
$\mu(ROA)^{Industry}$	35.789	35.181	37.716	38.825
	(1.86)*	(1.71)*	(2.36)**	(2.45)**
$Sales_{t-1}$	.042	.041	.043	.043
	(6.74)***	(6.80)***	(7.22)***	(7.18)***
Adj. R^2	.238	.240	.220	.221
No. of obs	6889	6793	10048	10048

Included: year FEs

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Pooled OLS, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level

Table 9: Predictors of Salary

The dependent variable, $Salary_t$, is the salary paid to the CEO during year t (data item *salary* in Execucomp). All independent variables are defined in Table 1. Years covered are 1992-2005. T-statistics in parentheses.

Dependent variable	$Salary_t$			
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility_{t-1}^{RET}$ (μ)		76.007	75.505	76.096
		$(2.36)^{**}$	$(2.82)^{***}$	$(2.86)^{***}$
$MeanInferredAbility_{t-1}^{ROA}$ (μ)	3.901			
	(.12)			
$Entrenchment_t$ (s)	23.822	22.942	18.423	18.129
	$(8.77)^{***}$	$(8.41)^{***}$	$(8.46)^{***}$	$(8.36)^{***}$
$Inverse\ of\ Firm-Specific\ Capital$ (q)	–32.752	–34.223		
	$(-1.67)^{*}$	$(-1.75)^{*}$		
$HasMBA$ (q)			23.724	
			$(1.73)^{*}$	
$Age_t \leq 60$ (q)				–53.230
				$(-4.36)^{***}$
$InverseCEOTenure_t$ (γ)	–228.328	–231.718	–119.203	–80.656
	$(-5.03)^{***}$	$(-5.02)^{***}$	$(-3.30)^{***}$	$(-2.15)^{**}$
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	–.219	–.132	.424	.526
	(–.32)	(–.18)	(.78)	(.98)
$\mu(ROA)^{Industry}$	8.240	7.837	6.539	6.828
	$(2.11)^{**}$	$(1.83)^{*}$	$(2.10)^{**}$	$(2.21)^{**}$
$Sales_{t-1}$	.011	.011	.011	.011
	$(7.90)^{***}$	$(7.91)^{***}$	$(8.33)^{***}$	$(8.30)^{***}$
Adj. R^2	.308	.306	.268	.273
No. of obs	6889	6793	10048	10048

Included: year FEs

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Pooled OLS, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level

Table 10: Predictors of Bonus

The dependent variable, $Bonus_t$, is the bonus paid to the CEO during year t (data item *bonus* in Execucomp). All independent variables are defined in Table 1. Years covered are 1992-2005. T-statistics in parentheses.

Dependent variable	$Bonus_t$			
	Panel A	Panel B	Panel C	Panel D
$MeanInferredAbility^{RET}_{t-1}$ (μ)		475.778	430.270	434.445
		(5.23)***	(4.88)***	(4.93)***
$MeanInferredAbility^{ROA}_{t-1}$ (μ)	182.631			
	(1.86)*			
$Entrenchment_t$ (s)	37.837	35.101	20.196	19.978
	(3.57)***	(3.28)***	(2.31)**	(2.27)**
$Inverse\ of\ Firm-Specific\ Capital$ (q)	87.907	71.883		
	(1.15)	(.94)		
$HasMBA$ (q)			97.838	
			(1.86)*	
$Age_t \leq 60$ (q)				–116.915
				(–2.51)**
$InverseCEOTenure_t$ (γ)	–644.168	–676.446	–386.162	–298.804
	(–3.54)***	(–3.69)***	(–2.76)***	(–2.13)**
$\sigma(ROA_{i,t} - ROA_{IND,t})$ (δ)	.778	1.467	1.727	1.994
	(.23)	(.42)	(.64)	(.75)
$\mu(ROA)^{Industry}$	26.363	25.923	29.616	30.443
	(1.53)	(1.41)	(2.06)**	(2.13)**
$Sales_{t-1}$	.030	.029	.031	.031
	(5.64)***	(5.69)***	(6.19)***	(6.15)***
Adj. R^2	.167	.170	.160	.161
No. of obs	6889	6793	10048	10048

Included: year FEs

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$

Pooled OLS, st. errors corrected for heteroskedasticity, clustered at the firm-CEO level