Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Fair and Square: A Retention Model of Managerial Compensation

Tore Ellingsen , Eirik Gaard Kristiansen

To cite this article:
Tore Ellingsen , Eirik Gaard Kristiansen (2021) Fair and Square: A Retention Model of Managerial Compensation. Management Science

Published online in Articles in Advance 13 Aug 2021

. https://doi.org/10.1287/mnsc.2021.4062

http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–21
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Fair and Square: A Retention Model of Managerial Compensation

Tore Ellingsen,[a,b] **Eirik Gaard Kristiansen**[b]

[a] Department of Economics, Stockholm School of Economics, S–11383 Stockholm, Sweden; [b] Norwegian School of Economics, N–5045 Bergen, Norway

Contact: gte@hhs.se, https://orcid.org/0000-0001-9501-0137 (TE); eirik.kristiansen@nhh.no, https://orcid.org/0000-0003-3971-0969 (EGK)

Received: July 13, 2020
Revised: November 17, 2020; February 12, 2021
Accepted: February 25, 2021
Published Online in Articles in Advance: August 13, 2021

https://doi.org/10.1287/mnsc.2021.4062

Abstract. We propose a model of how the retention motive shapes managerial compensation contracts. Once employed, a risk-averse manager acquires imperfectly portable skills whose value is stochastic because of industry-wide demand shocks. The manager's actions are uncontractible, and the perceived fairness of the compensation contract affects the manager's motivation. If the volatility of profits is sufficiently large and outside offers are sufficiently likely, the equilibrium contract combines a salary with an own-firm stock option. The model's predictions are consistent with empirical regularities concerning contractual shape, the magnitude of variable pay, the lack of indexation, and the prevalence of discretionary severance pay.

History: Accepted by Axel Ockenfels, behavioral economics and decision analysis.

Funding: Financial support from The Norwegian Finance Market Fund [Grant 197727], The Research Council of Norway [Grant 250506], and the Jan Wallander and Tom Hedelius Foundation [Grant P18-0180] is gratefully acknowledged.

Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2021.4062.

Keywords: compensation contracts • retention • performance pay • reciprocity

1. Introduction

There is a widespread suspicion that managers are overpaid. One reason for the suspicion is that managers are often lavishly rewarded when the firm is lucky and not correspondingly penalized when the firm is unlucky (Bertrand and Mullainathan 2001, Garvey and Milbourn 2006, Bell et al. 2021). Large salaries and long notice periods frequently constitute high lower bounds on pay, and when managers quit or are fired, they tend to receive additional discretionary severance pay (Yermack 2006, Goldman and Huang 2015). For harsh critiques of managerial pay based on these and related observations, see Bebchuk et al. (2002) and Bebchuk and Fried (2004).

Here, we argue that pay for luck, asymmetric rewards, discretionary severance pay, and other controversial compensation practices are not necessarily evidence of corrupt or inept compensation committees. As Oyer and Schaefer (2005) argue in their influential empirical study of firms that offer stock options to broad segments of managers, such practices might instead be understood as natural outcomes of fierce competition to attract and retain scarce talent.[1] Indeed, the contract predicted by our model of competitive compensation closely resembles observed managerial pay practices: the contract specifies a fixed salary combined with a nonindexed stock option package in case the manager stays and the lowest legal payment (typically zero) in case the manager leaves. Although actual pay coincides with contracted pay whenever the manager is retained, separations frequently involve discretionary severance pay. According to the model, this contract is not only an equilibrium contract, but it is also constrained efficient. That is, all the controversial features of this contract are compatible with flawless corporate governance.[2]

The core of the model is imported from Holmström (1983). When the initial match is formed, there is uncertainty about future market conditions. Thus, the optimal compensation contract should let the risk-neutral employer bear as much of the risk as possible. However, because departure penalties are prohibited, the manager cannot commit to stay if the pay is below the outside option.[3] Consequently, full risk sharing is not generally attainable. Rather, the equilibrium contract gives rise to a constant salary in bad states and pay equal to the outside option in good states.

The innovation relative to Holmström's model is to describe the outside option as well as the relationship between the employer and the employee in such a way that the model can be related to observable determinants and features of managerial compensation. For managers, we argue that the outside option in attractive states of the world primarily depends on the portability of the manager's industry-specific human capital. Retention of managers who can bring with them more of their skills to other firms in the same industry requires higher pay in good states. The ease with which managers can bring with them knowledge is affected by the firm's governance. Thus, our model rationalizes the association between weak corporate governance and large variable pay that was documented by Bertrand and Mullainathan (2001) and that they ascribed to managerial skimming. Moreover, we provide reasons for why the variable pay component is agreed in advance and why it is often formulated in terms of own-firm stock options.

In the main case studied by Holmström (1983), in which there are no mobility costs, the employee implicitly pays for insurance against low pay in the future by accepting a low wage in an initial period.[4] Here, instead, pay exceeding productivity in bad states is balanced by pay below productivity in good states. Underpayment in good states is possible because the manager acquires skills that are partly firm-specific. For example, the manager learns about the firm's technology as well as about how to deal with different stakeholders. Because of firm-specificity, the manager's outside value in similar employment is below the inside value, and hence, the manager stays even if underpaid. As in Holmström (1983), overpayment in bad states is possible because employment protection legislation provides asymmetric opportunities for contracting around the legal default. Specifically, although U.S. law stipulates that both sides are equally free to terminate the relationship (employment is "at will"), they are not equally free to relinquish their termination rights. Nothing prevents the employer from committing to a fixed-term or permanent contract, whereas employees typically cannot commit to stay (see Section 2.1).

To capture the notion of inefficient renegotiation outcomes and thereby justify ex ante contracts that condition pay on an index correlated with the outside option, we follow the lead of Halonen-Akatwijuka and Hart (2020). Specifically, we adopt the two main assumptions of the contracts-as-reference-points (CRP) model of Hart and Moore (2008): First, the manager's actions are not (fully) contractible. Second, the manager is reciprocal, supplying the efficient action if and only if the current contract is considered fair.[5] Under these assumptions, a renegotiable flat salary is not an optimal contract, at least not for highly mobile key employees whose wholehearted cooperation is vital to the firm. Renegotiation either entails ex post inefficient outcomes or ex ante inefficient outcomes. In the first case, the employer offers a new contract that the manager considers unfair, in which case the manager is demotivated. In the second case, the employer offers a new contract that the manager considers fair and, therefore, supports efficient effort. But this contract has the feature that pay varies more than it would if the contract were indexed to the outside option ex ante; hence, risk-sharing is inefficient.[6]

More precisely, our model makes the following prediction: all managerial compensation contracts have a salary component. If industry uncertainty is large enough and outside offers are sufficiently likely, the compensation contracts also have a variable component that can be optimally implemented through an own-firm stock-option package. The value of the stock options is proportional to the manager's portable skills. Because this is such an important implication of the model, we devote a separate section (Section 4) to discussing its empirical support.

Intuitively, the contracted variable pay is the smallest compensation that retains the manager whenever the best outside offer comes from within the industry and, hence, the smallest departure from perfect insurance that prevents contract renegotiation. With high enough variability, it is not possible to profitably offer a fixed-salary contract that also satisfies the manager in the best states. This prediction that higher industry variability increases the likelihood of variable pay is exactly opposite the prediction of the standard agency model. The retention motive, thus, provides an account for the mixed empirical evidence on this score.[7]

Another prediction of the model is that it is *not* necessary to contract explicitly on severance pay. To the contrary, discretionary severance pay is never worse than, and under plausible assumptions strictly better than, contracted severance pay. This result is largely a result of the CRP model; it would not hold under most other models of inefficient renegotiation, which attribute inefficiency to disagreement or delayed agreement.

We postpone a detailed discussion of related theoretical literature on compensation contracts until Section 5 with three exceptions.

First, we should further justify our choice of the CRP-model as a foundation for inefficient renegotiation. Although reciprocity furnishes a coherent structural model of contracting frictions, as with other "behavioral" assumptions, it often triggers the question: is fairness really a relevant concern for managerial compensation?[8] Many management scholars certainly think so; see Bosse and Phillips (2016) and Gartenberg and Wulf (2017) for recent contributions and extensive references. Two empirical studies are particularly relevant to our model.

Fong et al. (2010) use cross-section firm data to measure reciprocal responses of CEOs caused by overpayment or underpayment relative to peers. Ockenfels et al. (2014) use differences in the publicness of managerial performance pay within the same firm to document negative responses to perceived underpayment.[9] Thus, we do find the mechanism plausible. Compared with other models of inefficient renegotiation, for example, models based on asymmetric information, the CRP-model is also highly tractable.[10]

Second, despite the huge literature on reciprocity motives in labor markets sparked by Akerlof (1982), there are few theoretical studies of optimal compensation contracts in stochastic environments that take reciprocity motives into account. Englmaier and Leider (2012) is a notable exception, but they focus primarily on the feasibility of motivating locked-in agents without extensive use of variable pay.[11] By contrast, we study the design of optimal contracts for agents that have the opportunity to depart.

Third, we should further clarify the relationship to Holmström (1983) and Oyer (2004). Holmström (1983) primarily seeks to explain downward wage rigidity and layoffs for regular employees. His analysis does not pretend to explain empirically observed contracted variable pay. Nor is it straightforward to reinterpret the analysis in that way for the following reasons. First, the model does not justify contracting in advance what the wage should be in *good* states. Renegotiating pay upward works equally as well as any ex ante variable-pay contract. Indeed, under plausible extensions of Holmström's (1983) model, renegotiation is strictly preferable to an explicit contract.[12] Second, Holmström's (1983) model does not specify how an ex ante optimal contract could be implemented in practice. What are the verifiable indices of a regular employee's outside option, and why, in reality, is variable pay linked to the own firm's stock price rather than some other index? Third, Holmström's (1983) model does not explain why the composition of fixed and variable pay components differs across industries and other observable firm characteristics—a crucial question for the literature on performance pay.

Oyer (2004) constructs a model of contracted variable pay that addresses some of these concerns.[13] Oyer (2004) argues that contracts may link pay to the firm's performance because both the firm's performance and the employees' outside options are likely to be correlated with industry performance. He also argues that renegotiation is likely to be inefficient and that this is the reason why variable pay is agreed in advance. However, despite these virtues, there are several reasons to go beyond Oyer's (2004) analysis.

To start with, Oyer (2004) does not attempt to explain the exact shape of observed contracts. He assumes that contracts are composed of salaries and own-firm stocks rather than deriving this contract from first principles. In fact, the assumed contractual shape is not optimal given the economic environment. Under Oyer's (2004) assumptions, in which the outside option takes one of two values, whereas the price of the own firm's stock is normally distributed around two corresponding conditional expectations, it would be preferable to index pay against a broad stock price index rather than the own-firm stock price. Moreover, when pay is instead linked to the own firm's stock price, stock options would be a better instrument than stocks because they would admit constant pay (only the salary) for a range of low states in which outside options do not bind. But stock options would not be optimal instruments either because they pay too much when the own stock price is in a high range.[14] Although Oyer (2004) is clearly aware of and informally discusses these issues, it is desirable to build the model in such a way that they are reflected in the formal analysis as well.

Also, Oyer's (2004) model of the renegotiation cost is restrictive as it assumes that the employer incurs the same fixed cost whenever the worker leaves or the compensation is renegotiated. But renegotiation for the purpose of encouraging retention is a different event from renegotiation for the purpose of promoting departure, and we, thus, prefer to have a theory of the mechanism behind the renegotiation cost. Finally, and perhaps most importantly, our model complements Oyer's (2004) by offering an additional reason for interindustry variation in variable pay, namely differences in the portability of managers' human capital. Variation in portable human capital is the source of several additional testable predictions as we document in Section 4.

The paper is organized as follows. Section 2 explains the context and sets up the model. Section 3 derives the main results. Section 4 discusses how differences in portability of human capital might explain a variety of empirical regularities. It also briefly compares the explanation and policy implications of our retention theory to those of the managerial entrenchment theory in view of evidence from say-on-pay reforms. Section 5 relates our contribution to prior theoretical literature. Section 6 concludes. An appendix contains most of the proofs.

2. The Model

Before providing the detailed assumptions, it is useful to explain the legal framework.

2.1. Context

In the United States, employment contracts are generally at will. Employees are free to leave, and employers are free to terminate. In both cases, there are

caveats and exceptions, and there are opportunities for writing contracts that supersede the legal default.[15]

But, although there are hardly any restrictions on employers' ability to forgo their right to terminate, there are tight restrictions on employees' ability to do so (e.g., Estlund 2006). No employee can agree to a binding slavery contract, pledging never to quit. At most, one can pledge to not quit in order to join a competitor; this is the content of noncompete clauses, which are legally enforced in many jurisdictions.[16] However, even such limited exceptions from at-will principles are far from universal. There are important jurisdictions, such as California, in which noncompete clauses are held to be illegal.

The vast majority of top managers of large companies in the United States have fixed-term contracts (Schwab and Thomas 2006). That is, the employer forgoes the right to unilaterally terminate the relationship before a certain date. Thus, for our purposes of understanding managerial compensation contracts, we think that it is appropriate to view U.S. legislation as a system of one-sided at-will contracting. Similar conditions apply in many other jurisdictions as well.

2.2. Detailed Assumptions

A manager is employable for two periods but only produces in the second period; the first period can be thought of as training. Training is costless for both the firm and the manager. For simplicity, we assume that the first period is so short compared with the second period that it can be neglected for welfare purposes. We revisit this issue at the beginning of Section 3.6.

There are two industries, A and B, and at least three competing firms in each industry.[17] Firms always compete in Bertrand fashion, that is, by making binding contract offers in a noncooperative way. Firms are risk neutral and financially unconstrained.

Both the choice of industry and the compensation contract are endogenously determined by the model. However, the model description is simplified if we take for granted that the manager is employed in a particular industry, say A.

2.2.1. Technology. If the manager is employed by a firm in industry A, the manager's output is

$$y = \begin{cases} e & \text{if stays in firm;} \\ \theta e & \text{if moves within industry;} \\ \alpha e & \text{if moves across industries.} \end{cases} \tag{1}$$

Here, $e \in [0,1]$ denotes the manager's effort, $\alpha \in (0,\theta)$ denotes the manager's lower innate ability in industry B, and $\theta \in (\alpha,1)$ denotes the intraindustry portability of the manager's skills in industry A. The idea is that some of the manager's higher productivity in industry A stems from training, and some of the human capital that stems from training may be impossible to utilize in another firm.[18] In Section 3.6, we relax the assumption that the manager is never worth more in another firm in industry A.

If the manager is unemployed, the manager produces nothing of value. (However, unemployment never occurs in an equilibrium of the model's main specifications.)

2.2.2. Uncertainty. To begin with, we make some strong assumptions to simplify the analysis.

Assumption 1. (i) *The original employer always demands the manager's skill in period 2.* (ii) *There are always at least two external job openings for the manager in each industry.*

Firms take the output market prices p_A and p_B as given. There is uncertainty about these prices, but prices are perfectly negatively correlated.

Assumption 2. *The sum of the two prices is constant,* $p_A + p_B = 1$.

In Section 3.6, we study to what extent the results depend on Assumptions 1 and 2. In short, the insights are robust, but richer versions of the model potentially explain some additional regularities.

Because of Assumption 2, we may replace p_A by p and p_B by $1-p$. As the economy's state is single-dimensional, we can likewise write the ex ante distribution as $h(p)$. Correspondingly, the state space $\mathcal{P}$ is the support of h. Let h be integrable (with cumulative distribution function H) and symmetric around $p = 1/2$. The symmetry assumption is not important, but it simplifies the analysis by ensuring that the manager always starts out in the industry in which the manager's innate ability is highest, which we here take to be industry A.

2.2.3. The Surplus. To characterize the compensation and the contracts, a central concept is the value of the output, henceforth called the surplus, s. By Equation (1), we have that, if the manager remains in the initial industry, A, the surplus is

$$s(e,p) = ep \tag{2}$$

if the manager remains with the original firm and

$$s_A(e,p) = \theta ep \tag{3}$$

if the manager moves to another firm in industry A. If the manager moves to industry B, the surplus is

$$s_B(e,p) = \alpha e(1-p). \tag{4}$$

For future reference, we also define the best outside option

$$\hat{s}(e,p) = \max\{s_A(e,p), s_B(e,p)\}. \tag{5}$$

2.2.4. Contracts. A crucial distinction is that, between a compensation level in state p, which we denote $c(p)$ and the contracted compensation level in that state, which we denote $w(p)$. These can differ because both a contracted compensation can be renegotiated and the manager ultimately receives compensation from more than one employer. In general, we allow a contract w to depend not only on the state p, but also on for whom the manager works. By contrast, we do not allow contracting on the effort, e.[19] We think that this is relatively realistic as a first approximation. For example, stocks and stock options reward managers on the basis of market prices, whereas the actual effort of the manager, and even the output generated, are hard to observe and harder still to verify.

Let f_t denote the firm that employs the manager in period t. A contract, thus, specifies a wage $w(p,f_2) \in \mathbb{R}$. We say that the contract has *limited liability* if there is a restriction $w(p,f_2) \geq 0$. Such a limited liability constraint implies that departure penalties are illegal as they usually are in practice (see Section 2.1).

In addition to the contracts that are agreed upon before the resolution of uncertainty, we allow renegotiation of old contracts as well as new contracts to be signed after p is realized. Because the state is known at this stage, the renegotiation offer only depends on where the manager works. In general, we allow the renegotiation offer to be any real-valued function $w_n(f_2)$.

Throughout, we abstract from side-contracting, for example, in financial markets.

2.2.5. Manager Preferences. The manager cares about compensation, c, and about being fairly treated, but not about the effort level e.[20] When compensation is below the level to which the manager feels entitled, the manager feels a loss l_m that is equal to the difference between the entitlement and the offered compensation. The associated aggrievement can be reduced by imposing similar losses on the employer (typically by shading on performance e). More precisely, the preferences can be expressed by the utility function

$$U = u(c) - \max\{l_m - \tau l_f, 0\}, \tag{6}$$

where $l_m \geq 0$ is the loss that the manager experiences and $l_f \geq 0$ is the loss that the manager imposes in return. The parameter τ is positive, and we typically also assume that it is above one; that is, the manager can get rid of the manager's aggrievement without imposing as large a loss on the employer as the manager has been suffering.[21] The function $u(\cdot)$ is increasing and strictly concave; that is, the manager is risk averse.[22]

When a new contract is signed, either because there is no previous contract or because the previous contract is dominated by an outside option, the manager's entitlement equals the manager's outside option plus a fraction $\beta \in (0,1)$ of the difference between what the manager is worth to the employer, s, and what the manager can earn on the outside, say w^o. We call the difference $s - w^o$ a "relationship rent." If the offered wage w is below the entitlement, the experienced loss is $l_m = w^o + \beta(s - w^o) - w$.

To avoid any misunderstanding, let us emphasize that, in our model, aggrievement is only an issue if and when a contract is renegotiated. Compared with Hart and Moore (2008), our assumption represents a generalization; their assumption is that both parties feel entitled to all of the surplus from renegotiation, which is equivalent to $\beta = 1$ here.

2.2.6. Timing and Information. The timing is as depicted in Figure 1.

We assume that the history is common knowledge among all players. In particular, at stage 2c firms know which offers were made at stage 2b. All offers are verifiable by the proposer and recipient of the offer. Contracts can condition on the future state p but not on future offers (there are no "contracts on contracts").

2.2.7. Key Point. It is useful to illustrate already now why the CRP assumption creates renegotiation costs in some cases but not in all. Because firms are identical ex ante, in equilibrium, there is not any relationship rent associated with contracts offered at date 1. Therefore, there is no aggrievement as long as the original contract does not require renegotiation. But at date 2, because of the relationship-specific skill, the rent is typically positive. A positive relationship rent poses a problem if the original contract must be renegotiated.

For example, suppose at date 2 the original contract specifies $w(p)$, and an outside firm has offered $w^o > w(p)$, so the original contract is no longer viable. Recall that the manager is willing to supply effort $e = 1$ as long as there is no reason to be aggrieved. Thus, the relevant surplus under a fair renegotiation is $s(1,p)$. Assume $s(1,p) > w^o$ and let $w^r > w^o$ denote the new offer by the incumbent employer. Then, the manager is satisfied with the new offer if and only if

$$w^r \geq w^d := w^o + \beta(s(1,p) - w^o). \tag{7}$$

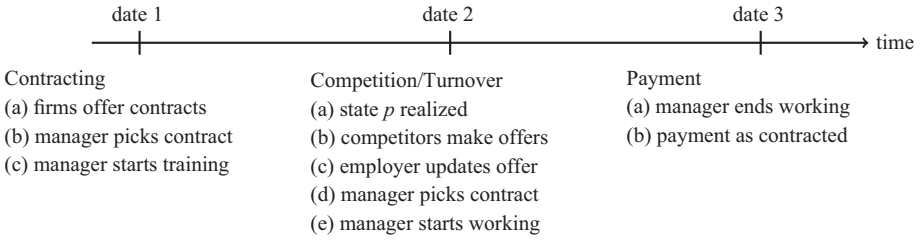
If w^r is smaller than the demanded wage w^d, it follows from maximization of (6) that the manager changes effort by Δe to impose a loss l_f on the employer that corresponds to $1/\tau$ times the own loss, $l_m = (w^d - w^r)$. That is, shading Δe should satisfy the equation

$$\Delta e s(1,p) = -(w^d - w^r)/\tau,$$

implying

$$\Delta e = \frac{(w^d - w^r)}{\tau s(1,p)}. \tag{8}$$

Figure 1. Timing



The shading exactly alleviates the manager's aggrievement, so that the manager's utility is again based on consumption alone, $u(w(p))$. Conveniently, therefore, the manager always accepts the highest offer at date 2.

If an existing contract is not dominated by an outside option, it remains the relevant reference point. The manager does not become aggrieved just because the employer happens to earn a large fraction of the surplus in a particular state. The manager also remains satisfied if the employer proposes replacing the original contract with a new contract as long as the new contract does not reduce the final compensation of the manager.

3. Analysis

We seek to characterize the subgame-perfect equilibrium outcome(s) of this game. We are ultimately interested in explaining the contracts that are observed in practice. However, it is useful to first characterize the manager's final compensation, c, because we may arrive at the same final compensation through different contractual paths.

Most proofs are relegated to the appendix. In particular, for a detailed derivation of constrained optimal contracts we refer to the proof of Lemma 1.

3.1. Welfare

Let $\eta \in (0,1]$ denote the welfare weight on the manager's utility. Because the firms are risk neutral, the (ex ante) welfare W is

$$W = \eta E[u(c(p),e(p))] + (1-\eta)E[s_i(e(p),p) - c(p)], \quad (9)$$

where the surplus s_i in the second term refers to the firm in which the manager ultimately exerts effort (the compensation c can, in principle, come from several employers.)

3.2. Benchmarks: Optimal Outcomes

From now on, we often write $s(p)$ as shorthand for $s(1,p)$. (Clearly, $e = 1$ is a feature of welfare-optimal outcomes.) Let the expected value of the manager's production given optimal assignment be denoted

$$E[s] := E_p[\max\{s(p), s_B(p)\}].$$

As firms are risk neutral and the manager is risk averse, firms ought to carry all risk.

Proposition 1 (First Best). *The compensation $c(p)$ maximizes the welfare function W only if $c(p)$ is constant for all $p \in \mathcal{P}$.*

Because the manager is free to depart without penalty, first-best outcomes are questionable benchmarks. Constrained optimal (or second-best) outcomes take mobility into account.

Definition 1 (Free Mobility). A compensation $c(p)$ satisfies free mobility if and only if, for all p, $c(p) \geq \hat{s}(p)$.

Let $\mathcal{M}$ be the set of all compensation functions that satisfy free mobility.

Proposition 2 (Second Best). *A compensation $c(p)$ maximizes the welfare function W subject to the free-mobility constraint $c(p) \in \mathcal{M}$ only if*

$$c(p) = \begin{cases} \hat{s}(p) & \text{if } \hat{s}(p) > \underline{c}; \\ \underline{c} & \text{otherwise,} \end{cases}$$

where $\underline{c} \geq 0$ is some fixed compensation.

In other words, although fixed pay is desirable for insurance purposes, it is not always sustainable because of the manager's option to depart.

It remains to determine $\underline{c}$. Because of symmetric Bertrand competition, it is reasonable to expect that equilibrium profits are exactly zero (and this is proven as follows), so let us focus attention on welfare-maximizing outcomes with zero profit. Because a welfare-maximizing outcome must allocate the manager efficiently and firms are risk neutral, it follows that the associated expected total compensation must equal the expected value of the manager's production. Thus, the second-best, zero-profit outcome has $\underline{c}$ being pinned down by the equation

$$E[c(p)] = E[s],$$

where $c(p)$ is determined in Proposition 2.

Our next result shows that, in any constrained optimum consistent with zero profit, there is always a

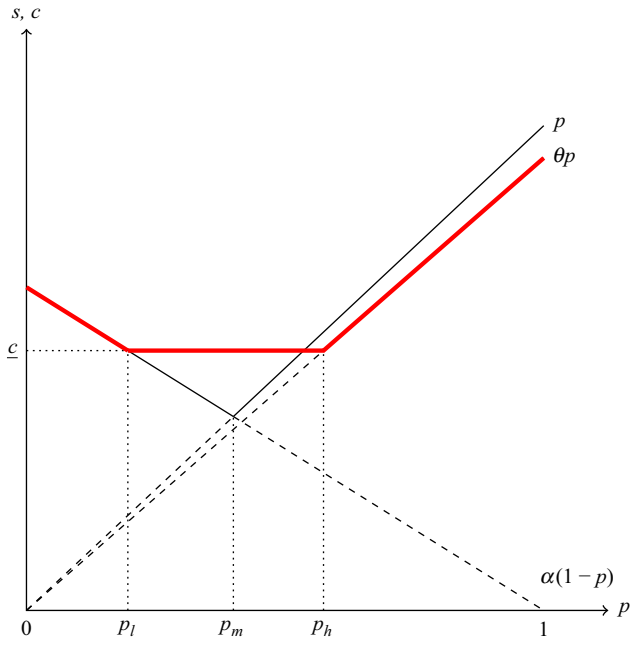
distinct possibility that the manager receives a fixed pay only. No manager receives a contract that yields positive variable pay in all states.

Proposition 3. *There is a nonempty interval of states* $[p_l, p_h]$ *such that any constrained optimal compensation* $c(p)$ *consistent with zero profit is constant for all* $p \in [p_l, p_h]$.

Figure 2 illustrates Propositions 2 and 3, focusing on the case in which the support of the function h is wide enough that pay may deviate from $\underline{c}$ both because of turnover to industry B (when p is low) and because of attractive outside offers from industry A (when p is high). From now on, we stick with this assumption.

Suppose p can take all values between zero and one. Then, the fat line is the second-best compensation as a function of p, taking $\underline{c}$ as given. For low values of p, the manager moves to industry B and is paid the competitive compensation $c(p) = s_B(p) = \alpha(1-p)$ there. As p reaches p_l, s_B decreases below $\underline{c}$. In the states $[p_l, p_h]$, it is optimal to keep compensation fixed. The manager is paid more than the value of the output and is allocated to the industry in which this value is highest, which is a firm in industry B for the states $[p_l, p_m)$ and the initial firm in industry A for the states $(p_m, p_h]$. For the remaining states $(p_h, 1]$, the manager remains with the original firm in industry A, but pay is again higher than $\underline{c}$, because, otherwise, the manager would move to a competing firm in industry A. Thus, free mobility limits risk sharing more when the portable human capital is high (θ is large). Observe that the interval $[p_l, p_h]$ could potentially coincide with $[0,1]$, in which case the compensation $c(p)$ is unconstrained optimal. This happens when α and θ are both sufficiently small.

Figure 2. (Color online) Constrained Optimal Compensation



Note. The fat line indicates the manager's total state-contingent compensation in a constrained optimum, respecting free-mobility and zero profit.

3.3. Manager Effort and Contract Renegotiation

Let us now turn to the analysis of the contracting game. We look for subgame perfect equilibria, so the analysis starts with the last period and moves backward.

The main idea is that compensation is potentially renegotiated under two separate circumstances. One circumstance is that the manager is worth more on the outside and so little on the inside that the firm is making a loss under the original contract. In this case, the manager is offered additional severance pay in order to leave. The second circumstance is that the manager is offered more on the outside than under the original contract but is worth even more on the inside. In this case, the manager is paid more in order to stay.

Lemma 1 describes behavior along any subgame perfect equilibrium path following the signing of an arbitrary stage 1 contract $w(p, f_2)$ with firm f. For completeness, we do not impose any limited liability constraint here, but solve the model for all initial contracts. In order to understand the manager's decision to stay or depart, the key variable is $\Delta w(p) = w(p, f) - w(p, f')$, the difference between the retention pay (associated with staying with the original employer f) and the severance pay associated with departing to another firm. (If there is no contracted severance pay, $\Delta w(p) = w(p, f)$; in our experience, this is the simplest case for gaining intuition.)

Lemma 1 (Renegotiation Outcomes). *Suppose* $\tau > 1$. *Suppose the contract* $w(p, f_2)$ *signed at stage 1b yields zero rent to the employer, f. Then, along any subgame perfect equilibrium path of the game starting at stage 2b:* (i) If $\Delta w > \hat{s} > s$, *the manager departs after receiving discretionary severance pay* $\Delta w - \hat{s}$. (ii) If $s \geq \hat{s} > \Delta w$, *the retention pay is renegotiated to* $w_n = \hat{s}$, *and the manager exerts effort* $e = 1 - \Delta e$. (iii) *In the remaining cases, the manager is compensated according to the original contract, switching to industry B if* $s < \hat{s}$ *and staying with the firm and exerting effort* $e = 1$ *if* $s > \hat{s}$.

Observe the difference between the two renegotiations (i) and (ii). When the manager is leaving, the firm need not worry about shading. When the manager is staying, the firm must weigh the cost of compensation against the cost of shading. Because $\tau > 1$, the optimal renegotiation offer equals the outside offer despite the shading that it entails. (If instead $\tau < 1$, shading would be so costly to the firm that it would be better to offer $c^r(p) = s(1, p)$ in order to prevent shading completely. This would avoid inefficiency resulting from shading but increase inefficiency because of poor risk sharing.)

We are now ready to characterize the equilibrium stage 1 contracts.

3.4. Unconstrained Contracts

Suppose firms can make unconstrained contract proposals. Then, the unique equilibrium outcome is implemented through a fixed salary $E[s]$, together with a severance pay $E[s]-s_B$ that keeps this same level of net compensation upon efficient turnover and a departure penalty that is large enough to prevent inefficient turnover within the industry.

Proposition 4 (Unconstrained Equilibrium Outcomes). *When contracts are unconstrained, the unique equilibrium compensation is $c=E[s]$ in all states. In particular, let $\rho - E[s]>0$ constitute a (prohibitively large) departure penalty. Then there is an equilibrium in which each firm in industry A offers contracts of the form*

$$w_u=\begin{cases} E[s] & \text{if stays in firm;}\\ E[s]-\rho & \text{if moves within industry;}\\ E[s]-s_B & \text{if moves across industries.}\end{cases} \tag{10}$$

The reason that this first-best outcome is not implementable in practice is that the departure penalty that is required to prevent within-industry turnover is either illegal (as in California) or too narrowly permitted. Thus, the model favors broad noncompete clauses in a competitive market for managers. Perfect risk sharing is otherwise unattainable as we now see.

3.5. Main Result: Contracts Under Limited Liability

Let us next characterize the outcome when departure penalties are prohibited. In this case, equilibrium outcomes cannot be fully efficient, but they are constrained efficient. Recall that the total compensation may originate from two separate employers in the case of turnover.

Proposition 5. *Suppose departure penalties are disallowed. Then, the manager's unique equilibrium compensation satisfies*

$$c(p)=\begin{cases}\hat{s}(p) & \text{if } \hat{s}(p)>\underline{c};\\ \underline{c} & \text{otherwise,}\end{cases}$$

where $\underline{c}\geq 0$ solves $E[c(p)]=E[s]$.

In other words, equilibrium compensation satisfies constrained optimality (cf. Proposition 2). In addition, equilibrium pins down the level of pay—the expected profit of the firms must be zero because of Bertrand competition between symmetric firms.

Let us now describe one contract that implements the equilibrium compensation; in the next section, we provide additional conditions under which this contract is the unique contract to do so. The contract combines (i) a salary, (ii) linear variable pay when own industry market conditions are favorable, and (iii) discretionary severance pay.

Proposition 6. *Suppose departure penalties are disallowed. Then, an equilibrium contract is*

$$w_1^*(p,f_2)=\begin{cases}0 & \text{if } f_2\neq f;\\ \underline{w} & \text{if } f_2=f \text{ and } p\leq \underline{w}/\theta;\\ s_A(p) & \text{otherwise.}\end{cases} \tag{11}$$

subsequent to renegotiation, the initial employer's wage bill is

$$w^*(p)=\begin{cases}0 & \text{if } s_B(p)>\underline{w};\\ \underline{w}-s_B(p) & \text{if } \underline{w}>s_B(p)>s(p);\\ w_1^*(p,f) & \text{otherwise.}\end{cases} \tag{12}$$

the wage floor (salary component), $\underline{w}$, is determined by the firm's zero profit condition

$$E[w^*(p)]=E[\mathbb{1}_{s(p)\geq s_B(p)}s(p)]. \tag{13}$$

Note that other equilibrium contracts only differ from $w_1^*(p,f_2)$ in one respect; they admit positive state-contingent severance pay rather than leaving the determination of severance pay entirely to contract renegotiation. As we see in the next section, this indeterminacy vanishes once we make the model more realistic by adding idiosyncratic uncertainty to firm profitability. Then, zero contracted severance pay is uniquely optimal.

Figure 3 illustrates the contracted and actual compensation paid by the original employer.

The following observations are immediate from inspecting Figure 3. First, the variable pay component corresponds exactly to an own-firm stock-option package; the manager has the right to buy stock at some trigger price once the value moves above this trigger.

Second, turnover decisions are always optimal. If p is sufficiently low, the manager switches to industry B. If $(1-p)\alpha>\underline{w}$, there is no severance pay. Otherwise, the severance pay is $\underline{w}-s_B(p)$ (the lower fat line on the interval (p_l,p_m) in Figure 3).

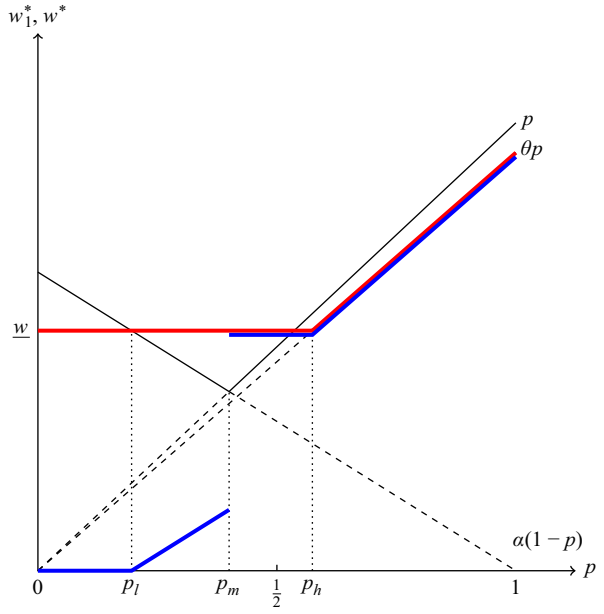
Third, if the uncertainty is sufficiently small (the set of states is clustered sufficiently closely around $p=1/2$), a fixed salary is an optimal contract. The manager is always most productive with the original firm, always stays, and never gets an outside offer that forces renegotiation of pay.

Fourth, when uncertainty is sufficiently large for the optimal contract to have a variable component, the expected variable pay is an increasing function of portability θ (we elaborate on this point in Section 4).[23]

Fifth, if turnover to an alternative industry is always inefficient (specialized manager), variable pay weakly increases as uncertainty increases.

Sixth, if turnover to an alternative industry is efficient in bad states, increased uncertainty increases managers total compensation. This follows from observing that, if the manager departs in the worst states, the firm's

Figure 3. (Color online) An Equilibrium Contract and Actual Compensation



Notes. The upper fat line indicates contracted compensation. The lower fat line indicates actual compensation paid by original employer. Additional pay, when the manager departs to the new employer in industry B, coincides with the line $\alpha(1-p)$ on the interval $p \in [0, p_m]$.

expected wage costs fall when market conditions are poor. Because profits are zero in equilibrium, the salary as well as the total pay must, therefore, be higher. Note also that a more versatile manager—that is, a manager with higher α—is more prone to depart in bad states and, therefore, is offered a higher salary. Let us prove the last results formally.

Corollary 1. *Let $\tilde{h}(p)$ be a mean-preserving spread of $h(p)$ with associated cumulative distribution function $\tilde{H}(p)$.*

i. *Suppose turnover is inefficient even for the lowest states, $\tilde{H}(p^m) = 0$, and the expected variable pay is weakly higher under $\tilde{h}(p)$ than under $h(p)$. The relationships are strict if*

$$\int_0^{p_h} \tilde{H}(p)dp > \int_0^{p_h} H(p)dp,$$

where $p_h = \underline{\tilde{w}}/\theta$.

ii. *Suppose turnover is efficient in some low states, $\tilde{H}(p^m) > 0$; then the expected total pay is weakly higher under $\tilde{h}(p)$ than under $h(p)$. The positive impact of uncertainty on expected pay is larger if the manager is more mobile across industries, that is, α is larger.*

3.6. Extensions

A major simplification relative to Holmström (1983) is that we shorten the initial period so as to be able to neglect the utility impact of consumption in that period. We think that this is without much loss of insight. The main focus of Holmström (1983) is precisely to understand the relationship between first- and second-period compensation when the manager's consumption in both periods is welfare relevant: under Holmström's (1983) assumptions that the manager has no wealth and cannot borrow and utility is time separable, compensation in the first period equals the compensation floor in the second period; this downward wage rigidity of long-term employment contracts is Holmström's (1983) key result and little is gained by repeating it here.[24]

To facilitate the analysis, we have also made strong assumptions concerning the structure of uncertainty. Let us now investigate what happens when we relax some of these assumptions. There are two kinds of extensions. One kind of extension maintains that all firms within an industry are inherently identical but changes what may happen at the industry level. Another kind of extension admits more heterogeneity at the firm level.

We think it is realistic that industry conditions are verifiable. We think it is unrealistic that the conditions of individual firms are verifiable. Thus, we maintain that contracts signed at date 1 are of the form $w(p, f)$, where p is defined at the industry level only.

A minor change is to generalize the assumption $p_B = 1 - p_A$ to $p_B = 1 - \nu p_A$ with $\nu \in [0,1]$. This generalization has little impact on the results. The nature of the optimal contract is preserved even in the case of no correlation, $\nu = 0$. The only difference is that turnover to industry B is either never associated with severance pay or associated with a constant level of severance pay for a range of p_A values.[25]

3.6.1. Aggregate Shocks. What happens if we relax Assumption 2, that $p_A + p_B = 1$ beyond the case $p_B = 1 - \nu p_A$? Suppose that an aggregate shock may change overall productivity but does not affect the relation between p_A and p_B (i.e., $p_A + p_B = k$ with k stochastic and verifiable). Because firms are risk neutral, they optimally shield managers against these shocks as in Holmström (1983). Thus, the shape of compensation contracts remains the same, but the generalization eliminates the unrealistic feature that owners of firms could perfectly diversify all risk by holding shares in both industries.

3.6.2. Idiosyncratic Shocks and Robustness. In reality, firms face idiosyncratic uncertainty about productivity in addition to the industry-specific uncertainty captured by p.

Let us first relax the assumption of perfect correlation while maintaining the assumption that actual movements within the industry are unprofitable. If it is known at date 1 that the manager's best outside option is employment with a particular competitor, then

the optimal contract should clearly condition on the stock price of that competitor rather than the own firm's stock price. However, in the more realistic case that the most tempting outside offer can come from any number of firms, it is typically better to condition on the own stock price than on the stock price of a particular competitor. If the manager's human capital is most valuable with competitors that have chosen similar technologies and strategies, the price of own firm's stock is better correlated with the most attractive outside offer than is a random firm's stock (or an index of all stocks).[26]

3.6.3. Idiosyncratic Shocks and Uniqueness. Let us next consider idiosyncratic outside offers that might dominate what the original employer is able to pay. For simplicity assume as in Assumption 1(ii) that there is a thick market with "normal" outside offers. But, in addition to these outside offers, suppose that in each state p, there is a probability $g(p)$ that some outside firm values the manager at $s^o(p) > \max\{\underline{w}, s(p)\}$. Let us henceforth call offers by such superior employers "superb."

Proposition 7. *When the market is thick and there is a probability $g(p) \in (0,1)$ of a superb offer, a contract with the shape of $w_1^*(p, f_2)$ (but generally different salary $\underline{w}$) is the unique equilibrium contract.*

Intuitively, the possibility of turnover because of a superb offer affects the salary component but does not affect the slope of variable pay. Most importantly, there is now a unique equilibrium contract, in which the contracted severance pay is exactly zero: because there is always a positive probability that no severance pay needs to be paid in order to induce turnover, any contracted severance pay merely serves to increase the variance of pay compared with the outcome under discretionary severance pay only.[27]

3.6.4. Thin Markets. Let us next relax Assumption 1(ii), that there are always plentiful outside offers (the "thick-market" assumption). Suppose first that there are always plentiful offers from industry B (this makes sense if industry B is shorthand for "other industries") but that an outside offer $w^o \le s_A(1,p)$ from within the industry arrives with probability $\mu \in (0,1]$. Recall that we do not permit initial contracts to condition on such future offers.

Proposition 8. *There is some $\bar{\mu} \in (0,1)$ such that, for all $\mu < \bar{\mu}$, a fixed salary is the unique optimal contract.*

The argument is simple. As μ tends toward zero, the fixed wage contract constitutes a closer and closer approximation to the first best (as the probability of shading tends to zero). By contrast, the loss of risk sharing associated with variable pay components that respond to p (irrespective of the arrival of outside offers) is independent of μ. Thus, the model predicts that contracted variable pay is less common when the market for managers is thinner. This observation might help to explain why there is more variable pay in United States, which has a thick managerial labor market, than in other rich countries despite similar overall levels of pay (Fernandes et al. 2013).[28] To the extent that regulated sectors have fewer firms, Proposition 8 might also explain why there is less variable pay in these sectors (see Murphy 1999, Frydman and Saks 2010).

3.6.5. Contract Length and Renewal. Many managerial contracts are renewed. That is, the original duration of the contract is shorter than the maximum duration of the relationship. Presumably, the reason for limiting the duration of the contract is that the quality of the match might change in ways that cannot easily be contracted on in advance. Thus, a relevant extension of the model would be to add another production period and to allow the manager's productivity to change enough between the two production periods for a replacement to be desirable.

Within the CRP framework, Halonen-Akatwijuka and Hart (2020) analyze the choice between short- and long-term contracts and what they term "continuing contracts"—with which the parties can negotiate a new contract based on the reference point(s) established by the previous contract. The reference point reduces the scope for disagreement compared with a sequence of unrelated contracts. It seems to us that continuing contracts is an adequate description of the situation for many top managers as they near the end of their current term. On the basis of the analysis in Halonen-Akatwijuka and Hart (2020), we offer the following intuition: the manager's firm-specific human capital implies that there is potentially a wide range of outcomes that are consistent with retention. Yet, if pay under the previous contract serves as a reference point, the firm would nonetheless normally avoid forcing the manager to accept pay reductions unless ending the relationship is preferable to continuing under the same terms. Thus, CRP offers a theory for downwardly rigid pay between successive contracts and not only within a contract. We return to this issue when discussing say-on-pay reforms at the end of Section 4.

3.6.6. When Do Results Break Down? Why is the contract piece-wise linear in equilibrium? The immediate reason is plain. The fraction θ of the industry-specific skill that is portable is a constant. Therefore, the value of portable human capital is θp. The linearity arises because the portability of human capital does not depend on the market price p.[29]

There clearly exist generalizations of the model that produce more nonlinearity. One generalization is to admit the distribution of potential outside offers to always have full support on the interval $(0, s(p))$. In this case, it is more difficult to provide clean conditions under which the equilibrium contract is piece-wise linear.[30]

Another challenge to the simplicity of the optimal contract is that there could be uncertainty related directly to portability θ. Suppose, for example, that there is new information about θ at the beginning of date 2. We will not go so far as to completely formulate and solve the resulting problem here, but we note that it invites another common and controversial practice, namely option repricing (for details and another argument for why repricing is benign, see Acharya et al. 2000): if θ increases, the manager's outside option improves. Thus, the strike price of the options needs to be reset downward and the amount of options increased in order to prevent the manager from leaving.[31]

The departures from linearity would be even greater if we introduced a choice of training intensity at date 1, especially if the outcome of training is stochastic. Englmaier et al. (2014) develop a model of knowledge workers in which higher effort stochastically increases both the inside value and the outside option. Although their model only has two outcome states, it already demonstrates that there is a complex interplay between effort incentives and retention incentives when effort affects the outside option.[32]

For all these reasons, own-firm stock options are unlikely to be *exactly* optimal in practice.

4. Portability, Empirical Regularities, and Policy

Preventing employees from leaving with valuable human capital is a central element of good corporate governance; for detailed arguments, see, for example, Rajan and Zingales (2001) and references therein. In our model, we do not endogenize the level of portability. Instead, we focus on how compensation should optimally be designed to match whatever level ensues. Accordingly, the key comparative static implication of our model is that greater portability implies greater variable pay.

In order to evaluate the relevance of our model, the first question is, therefore, which are the factors affecting portability, and how do they show up in empirical work?

4.1. Portability and Empirical Patterns

Portability is affected by two main factors. Portability is higher (i) when the initial firm has weak property rights over the manager's human capital (and it is difficult to transfer such rights from the manager to the firm), and (ii) when there exist other firms that can make good use of the human capital. With these two factors in mind, we see that the positive association between portability and variable pay resonates well with many empirical observations:

1. "Knowledge" firms utilize stocks and especially stock options to a larger degree than do "brick and mortar" firms; see Anderson et al. (2000), Ittner et al. (2003), Murphy (2003), and Oyer and Schaefer (2005). The knowledge firms themselves report that performance-based pay is primarily used for retention purposes; see Ittner et al. (2003). We think of knowledge firms as having high θ relative to brick-and-mortar firms.

2. Option-based compensation is particularly common in "growth firms," both for executives (e.g., Smith and Watts 1992, Gaver and Gaver 1993, Mehran 1995, Himmelberg et al. 1999, Palia 2001) and nonexecutives (e.g., Core and Guay 2001). We think of growth firms as being more recent and having less well-protected technologies and growth options, hence, a higher portability θ.

3. Industries with a higher fraction of outside executives have both a larger fraction of performance-related pay and a smaller degree of indexing, that is, more pay for luck; see Cremers and Grinstein (2014) and also Murphy and Zabojnik (2004, 2006). For reasons provided, we think of high observed mobility as indicating high θ.

4. Firms with high institutional ownership as well as firms that adopt "good governance practices" have fewer stock options and related asymmetric variable pay (Bell et al. 2021). We think of large owners as exercising better control and, hence, reducing θ.

5. There is less performance-based pay in family firms (e.g., Kole 1997, Anderson and Reeb 2003, Bandiera et al. 2015), especially when the manager is a family member (Gomez-Mejia et al. 2003). We think that family members are more reluctant to move; hence, θ is low.

6. There is less mobility of managers and managers have weaker performance incentives in jurisdictions with stronger enforcement of noncompete employment clauses (Garmaise 2011). This is perhaps the most direct interpretation of our theory; θ is smaller when noncompete clauses are stronger.[33]

7. There is a weaker link between pay and industry-specific shocks ("luck") in companies with large owners, especially when these large owners sit on the company's board of directors (Bertrand and Mullainathan 2001; see also Fahlenbrach 2009, and Bell et al. 2021).

8. Performance hurdles for option contracts are increasing in the quality of corporate governance; see Bettis et al. (2010). This point is similar to the previous one. If good corporate governance entails better safeguards against managers departing with valuable assets, good corporate governance is associated with a higher hurdle price p_h (see Figure 3).

In addition to this cross-section evidence, Murphy and Zabojnik (2004, 2006) argue that the relative importance of transferable talent has increased over time as evidenced by the executives' education as well as the increasing frequency of externally hired executives. If this view is accepted, our model can account for the increase in variable pay over the last few decades that is documented by Frydman and Saks (2010).

4.2. Retention or Skimming?

A common interpretation of more asymmetric pay and greater pay for luck in "weakly governed" firms (items 4 and 7) is that powerful managers take advantage of passive owners to raise their pay. Specifically, as argued by Bertrand and Mullainathan (2001), managers may influence weak or friendly boards to raise managerial pay in states of the world in which owners will not notice it too much. They maximize their own pay—engage in "skimming"—subject to owners' "outrage constraints." As noted in the introduction, this line of argument is often advanced in favor of corporate governance reforms. Because our theory explains the same findings, the questions arises: can we empirically disentangle our version of the managerial retention hypothesis from the managerial power hypothesis?[34]

In their pioneering work, Bertrand and Mullainathan (2001) consider the following three observations to favor the skimming hypothesis over the retention hypothesis: (i) asymmetric response of compensation to industry conditions,[35] (ii) lack of procyclical turnover,[36] and (iii) lack of positive correlation between turnover and pay for luck.[37] But unlike Bertrand and Mullainathan's (2001) depiction of the retention hypothesis, our retention model is consistent with all these three regularities. The asymmetry is an optimal response to the demand for risk-sharing, the countercyclical turnover is an optimal response to smaller (and eventually negative) gains from keeping the relationship together when industry condition are poor, and the lack of correlation between pay for luck and turnover follows from the fact that turnover is driven by the quality of the match, which is unrelated to pay for luck (differences in pay for luck are driven exclusively by the portability of human capital in our model).

4.3. Evidence from Say-on-Pay Reforms

If correlational evidence does not discriminate between the skimming hypothesis and our rentention hypothesis, perhaps data from governance reforms can do so? The most natural reforms to study are those that target managerial power. If overpayment is a big problem, then such reforms ought to substantially reduce the pay of influential managers. If, instead, the pay practices are nearly optimal, these reforms should not have a lasting impact on the level of pay.[38]

What is the evidence that governance reforms have altered managerial pay in a sustained fashion? Rather than offering a comprehensive survey, let us discuss the reforms that target pay most directly, namely say-on-pay regulations.[39] The pioneering say-on-pay reform was adopted by the United Kingdom in 2002 and granted all shareholders the right to participate in a nonbinding vote on executive pay.[40] Subsequently, several countries have enacted similar reforms and sometimes even imposed binding votes; see Ferri and Göx (2018) for a comprehensive survey of the literature about say-on-pay.

Studies of firms in the United Kingdom and the United States, where the concerns about excessive pay have been greatest, have found little evidence that these reforms markedly reduce average CEO pay. In fact, there is even some evidence that average pay has gone up as a consequence of the reform; see Iliev and Vitanova (2019). This is not to argue that the legislation has not affected governance. For example, Cuñat et al. (2016) find that say-on-pay raises productivity, suggesting that the threat to intervene over pay helps to discipline managers. Nor does the lack of impact on average pay imply that the composition of pay is unaffected. For example, Ferri and Maber (2013) find that managers are rewarded less for poor performance. Specifically, votes against compensation plans were associated with movement toward shorter notice periods (and, thus, smaller severance pay) in case of dismissal as well as less lenient retesting provisions for stock options. These three studies suggest that governance improves because the reform but that overall pay levels were already quite competitive.

Correa and Lel (2016) use a much broader sample of firms from 38 countries, of which 11 adopted some version of say-on-pay during the period 2001 to 2012. They find that the growth of average CEO pay slows markedly in a country after the reform and that the reduction is concentrated in firms with weak governance and in which managers with long tenures are overpaid relative to observed indices of performance. They interpret these findings as indicating skimming.[41]

Our model suggests the following alternative interpretation: the introduction of say-on-pay legislation encourages shareholders to renegotiate implicit contracts with managers with relatively poor outside options. Specifically, consider managers who reach the end of their current contracts and whose pay is, therefore, up for renewal (see our discussion of contract renewal in Section 3.6). Suppose the directors are inclined to apply the logic of CRP theory and offer a continuing contract with conditions that are no worse than existing conditions. However, as the manager's human capital by now is partly firm-specific,

shareholder activists might point out that this level of pay is well above the manager's outside option and insist that pay is reduced. Whether the pay reduction benefits the shareholders depends on the manager's response. If the manager is not aggrieved (perhaps because the directors in fact had proposed a "fair" continuing contract and the manager remains loyal to them), this extraction of managers' quasi-rents might benefit the shareholders in the short run. However, because the prospect of similar hold-up harms risk sharing in the future, there could still be long-run damage entailing either higher average pay for new managers or the best managers preferring to work for companies where say-on-pay does not apply.[42] Thus, even in this apparently clear-cut case for the skimming hypothesis, the retention hypothesis offers a complementary perspective that might affect the final judgment of say-on-pay reforms.

5. Related Literature

The literature on optimal compensation contracts is huge. Yet, besides Harris and Holmström (1982), Holmström (1983), and Oyer (2004), which we have already discussed, there are only a handful of closely affiliated contributions.

Most of the compensation contracting literature focuses on effort incentives within a relationship, neglecting the issue of mobility.[43] That is, it primarily studies the bilateral relationship between a single employer and a single employee with market forces determining the average level of pay but otherwise playing a subordinate role.[44] Moreover, most models consider the case in which the manager is entirely selfish, and contractual payments can be made contingent on some noisy measure of output. Because the corresponding general effort inducement problem tends to produce fairly complicated contracts, the previous literature frequently considers a restricted set of contractual shapes, in particular, linear ones. For example, as we document, very few models attempt to explain why contracts simultaneously comprise both total pay floors and linear performance pay components. Within the class of general effort-inducement models, perhaps de Meza and Webb (2007) come closest to producing this shape. They show that loss-aversion can generate optimal contracts with a flat segment, just like a base salary. However, the strictly increasing part of the compensation is not generally linear, and it often starts with a jump; see also Herweg et al. (2010) and Dittmann et al. (2010).[45]

A notable precursor to our work is Holmström and Ricart i Costa (1986). There too, optimal compensation takes the form of an option contract with the fixed salary being a result of the employee's risk aversion and the variable pay being a result of the employee's inability to commit to staying with the current employer when outside opportunities become attractive. However, where Holmström and Ricart i Costa (1986) emphasize uncertainty about employee characteristics, we emphasize uncertainty about future market conditions. Therefore, we are able to address many empirical regularities regarding which their model is silent. For example, we can explain why regular own-firm stock options are used to reward employees whose talents are well known and whose effort does not greatly affect the value of the firm; in their model, the option is instead tied to what is revealed about the specific skills of individual employees, for which the stock price is typically a less precise indicator.[46] Another difference is that Holmström and Ricart i Costa (1986) assume that mobility barriers are absent. Without any benefit from retention, the magnitude of their fixed wage component is bounded by the principal's ability to extract surplus from the employee through low pay in an initial period. In our model, the magnitude of the fixed wage is instead largely driven by the size of the mobility barrier and can, thus, attain a more realistic size. Apart from Holmström and Ricart i Costa (1986), we are not aware of any previous model that explains why employees are paid a combination of fixed salary and nonindexed stock options.[47]

The role of portable human capital has also recently been studied in the literature on relational contracting. For example, Kvaløy and Olsen (2012) argue that increased portability tends to favor individual incentives over group incentives and sometimes yields individual pay that is driven by outside options rather than inside contributions.

Our work is complementary to the literature on matching and compensation of executives, such as Gabaix and Landier (2008), Terviö (2008), and Edmans et al. (2009). That literature also studies the compensation of managers with heterogeneous skills. However, the emphasis is on the equilibrium level of pay and on the matching of managers within an industry composed of heterogeneous firms, whereas we instead focus on the shape of the compensation contract and neglect intraindustry firm heterogeneity.[48]

The model might also be compared with the various other theories that seek to explain the same empirical regularities. We have already noted that *effort-inducement models* typically struggle to rationalize the findings except one regularity at a time; this is what motivated our analysis in the first place. From the point of view of predictions, perhaps the most closely related effort-inducement models are those that focus jointly on effort inducement and optimal turnover. Inderst and Mueller (2010) may be closest.[49] There, the manager receives private information about match quality at an interim stage, and the optimal contract needs to induce managers to work hard, to induce poorly matched managers to quit, and to

induce well-matched managers to stay. The solution is to let pay be sharply increasing in performance and more so if it is easier for the manager to shirk or if the firm value is more variable. Inderst and Mueller (2010) also provide a reason why the magnitude of severance pay need not be agreed in advance (severance pay has less problematic incentive effects at the interim stage than before the manager exerts effort). Because Inderst and Mueller (2010) confine attention to the case of two possible outcomes, it is not clear under which additional assumptions this approach could also account for the more detailed shape of actual contracts.

Another important class of models focuses on *risk-inducement*. By this logic, risk-averse managers are inclined to take too little risk from the perspective of the better diversified owners, and stock options can better align their interests. Although we would not deny the plausibility of this mechanism for some CEOs, we note that risk-inducement does not offer a convincing explanation in the common case of employees who receive stock options despite hardly affecting the stock price at all (Oyer and Schaefer 2005). And, even for CEOs, the risk-inducement models fail to explain the absence of indexing.[50]

Finally, from the perspective of contract theory more broadly, our analysis contributes to the recent literature on CRP. The core question in this literature is how best to write contracts that are inevitably incomplete. In their seminal paper, Hart and Moore (2008) study the costs and benefits of flexible and rigid contracts. Flexible contracts allow adaptation to the economic environment within the boundaries of the contract, but flexibility invites shading whenever this adaptation is considered unfair. By contrast, a rigid contract does not invite shading even if the realized allocation of surplus is unbalanced as long as the contract was considered fair when it was agreed. Our contribution abstracts from the need for flexibility within the contract; the desired action is independent of the state. Instead, we focus on the renegotiation of contracted compensation in the presence of outside offers. Because of limited liability, outside offers by inferior employers always constitute a nuisance. In the absence of any fairness concerns, employers would deal with this nuisance through renegotiation. But, because employers foresee the grievance and shading that would be associated with renegotiation, they instead offer explicit variable pay contracts. This insight is not unique to us; it is also prominent in Halonen-Akatwijuka and Hart (2020).[51] However, by considering risk aversion, our analysis does add a twist: even if employers would find it desirable to keep employees happy and motivated in case of contract renegotiation, such renegotiated contracts involve more variability and, hence, worse risk sharing than indexed contracts do. Thus, it is not necessary to believe that renegotiation would cause bad feelings in order for indexed contracts to be strictly preferable.

Observe that the CRP theory is key to explaining the existence of performance-pay contracts that are renegotiated in case of departure but not in case of retention. This outcome could not be rationalized by standard bargaining theories, such as the alternating-offer bargaining (AOB) model associated with Rubinstein (1982) and Binmore et al. (1989).[52] According to AOB, renegotiation under symmetric information is always ex post efficient, and a party with a relatively attractive outside option gets no more than this outside option in equilibrium. Thus, there is no reason to specify retention terms in advance. By contrast, according to AOB, there is typically a reason to contract about *severance pay* in advance. If, instead, the firm were to rely on renegotiation in order to induce the manager to leave, the manager would receive a fixed fraction of the benefit from separation; the payment would, thus, exceed the contracted salary in these states (except if the manager is very impatient or risk averse, in which case the employer has all the bargaining power). To compensate the firm for the expected renegotiation payments, the salary would need to be lower to begin with, implying worse risk sharing.

6. Conclusion

Critics of generous compensation practices often take for granted that optimal contracts should encourage high effort while providing adequate insurance. Against this backdrop, they argue that non-indexed performance pay, high pay floors, and discretionary severance pay cannot be efficient and are, thus, prima facie evidence of managerial rent extraction. Our analysis shows that another interpretation is possible: when contracts are designed to balance insurance and retention, optimal contracts have all three features. Our version of the retention hypothesis can also explain several other regularities that were previously considered to be evidence against the retention hypothesis, such as countercyclical turnover.

The regularity that is most difficult to square with the retention hypothesis is the cross-country evidence that say-on-pay reforms are associated with subsequent reductions in the growth of pay for top managers. But, as we see, even this evidence is potentially consistent with our retention model as the reductions might well be due to extraction of quasi-rent from locked-in managers. Only if there is a negative impact of say-on-pay on the net present value of new managerial contracts can we be sure that the evidence supports the entrenchment hypothesis over the retention hypothesis. As time passes, that research question should soon be answerable.

Acknowledgments

This paper supersedes our earlier efforts "Paying for Staying: Compensation Contracts and the Retention Motive" and "Competitive Compensation Contracts," which were vulnerable to the critique that we artificially restricted the space of feasible contracts. The authors are deeply grateful to Bengt Holmström for insightful, patient, and exceedingly polite guidance and to Axel Ockenfels (editor) and his team of anonymous referees for many productive comments and suggestions. Thanks also to Carsten Bienz, Mike Burkart, James Dow, Mariassunta Giannetti, Robert Gibbons, Oliver Hart, Hans Hvide, Daniel Metzger, Espen Moen, Trond Olsen, Marco Ottaviani, Patrick Rey, Per Strömberg, Steve Tadelis, Marko Terviö, Joel Watson, and previous anonymous referees for helpful comments and discussions. Remaining shortcomings are the authors'.

Appendix. Proofs

Proof of Proposition 1

Supposed to the contrary that $c(p)$ is not constant on the support $\mathcal{P}$. Then, by strict concavity of u, there is some constant compensation $\underline{c} < E[c(p)]$ such that $u(\underline{c}) > E[u(c(p))]$. Because $E[s(p) - \underline{c}] > E[s(p) - c(p)]$, it follows from (9) that $W(\underline{c}) > W(c(p))$.

Proof of Proposition 2

(i) Consider any $c(p)$ that is not constant when $c(p) > \hat{s}(p)$. By the same argument as in the proof of Proposition 1, this compensation is dominated. (ii) Consider any $c(p)$ such that $c(p) > \hat{s}(p) > \underline{c}$ on a positive measure of states $\mathcal{P}_m$. This compensation is welfare dominated by a compensation $\tilde{c}(p)$ generated in the following way: $\tilde{c}(p) = c(p) - \epsilon$ for $p \in \mathcal{P}_m$, where $\epsilon > 0$ is small enough that $\tilde{c}(p) > \max\{\underline{c}, \hat{s}(p)\}$ for all $p \in \mathcal{P}_m$; in remaining states, $\tilde{c}(p) = \underline{c} + \delta$, where $\delta > 0$ is the unique solution to $E[\tilde{c}(p)] = E[c(p)]$. The welfare of the firms is unchanged. Because $c(p)$ is a mean-preserving spread of $\tilde{c}(p)$ and u is concave, it follows that $E[u(\tilde{c}(p))] > E[u(c(p))]$.

Proof of Proposition 3

Consider a constrained optimal compensation $c(p)$ as defined in Proposition 2. Suppose $c(p) = \hat{s}(p)$ in all states p. Then, the manager is paid the full value of the production in states p such that $\hat{s}(p) = s_B$ but less than the full value of the production in states p such that $\hat{s}(p) = s_A(p)$. The latter states include all $p \geq 1/2$. Hence, if the manager receives only constrained optimal variable pay, firms earn positive profit, a contradiction. Finally, because the outside options are greatest for extreme values of p, the states associated with fixed pay cover a single interval $[p_l, p_h]$.

Proof of Lemma 1

Because we seek subgame-perfect equilibria, the analysis starts at the last stage.

Stage 3: If the relevant contract in place, $w(p)$, grants the manager at least a fraction β of the available surplus at the time the contract was signed (the available surplus at this stage is the difference between the expected pay associated with $w(p)$ and the expected pay associated with the best competing offer w^d), the manager exerts effort $e = 1$; otherwise, the manager's effort is

$$e = 1 - \left(\frac{w^d - w(p)}{\tau s(1,p)} \right), \tag{A.1}$$

as computed in (8).

Stage 2d: Because the manager ultimately only obtains utility from consumption, the manager accepts the offer (one of the offers) that yields the highest total pay (there is no uncertainty at this stage).

Stage 2c: There is Bertrand competition for the manager's service. Suppose the manager either does not hold an offer or holds an offer that is below $\hat{s} = \max\{s_A, s_B\}$. Then, by the standard Bertrand logic, in any subgame-perfect equilibrium, at least two firms make offers of exactly $\hat{s}$.

Stage 2b: Suppose the manager is employed in firm f. Recall that $\Delta w(p) = w(p,f) - w(p,f')$ denotes the net pay that is specified by f's contract if the manager stays rather than leaves for firm f'. By the analysis of stage 2c, we know that the manager is only retained if $s \geq \hat{s}$. The two cases require analysis because they potentially involve renegotiation.

Part (i) (paying to induce departure): $\Delta w > \hat{s} > s$. Note that this implies $s - w(p,f) < 0$, so the firm makes a loss and is willing to pay the manager for leaving. The manager is worth more on the outside, but net pay from staying is higher than the best outside offer. There is a range of mutually profitable contracts, in which f increases the pay in case of departure by at least $\Delta w - \hat{s}$. Thus, the equilibrium (increase in) severance pay is exactly this amount, establishing part (i). (Note: shading is not a concern in this case because the manager will not remain in the firm.)

Part (ii) (paying more than Δw to induce retention): $s > \hat{s} > \Delta w$. In this case, the employed manager is worth more at the incumbent firm than outside but will not stay in the firm under the current contract. Thus, the incumbent makes a renegotiation offer. How large should the new wage be? From (A.1) we see that $de(w(p))/dw = 1/\tau s(1,p)$. Recalling that $s(e,p) = ep$, it follows that $d(s-w)/dw = 1/\tau - 1$, which is negative when $\tau > 1$ and positive if $\tau < 1$. Thus, the optimal renegotiated wage equals $\hat{s}(p)$ when $\tau > 1$ and $s(1,p)$ when $\tau < 1$.

Part (iii): There are four subcases. (a) $s > \Delta w > \hat{s}$, (b) $\Delta w > s > \hat{s}$, (c) $\hat{s} > \Delta w > s$, (d) $\hat{s} > s > \Delta w$. In cases (a) and (b), the outside options are the worst, and the manager prefers to stay. There is no reason to renegotiate, so the manager's pay is $w(p,f)$. Because the firm got none of the rent when the contract was signed, the manager's effort is $e = 1$. In cases (c) and (d), the outside option dominates both the manager's current contract and the maximum amount that the current employer could pay. Therefore, the manager takes up the option to leave and collects the contracted severance pay $w(p,f')$.

Proof of Proposition 4

Note that the manager pays a penalty equal to s_B if departing to industry B and a prohibitively large penalty if departing to a competing firm in industry A.

First, we show that this contracting outcome is part of a subgame-perfect equilibrium. From Lemma 1, we know that the manager earns $\hat{s}$ by departing. If $s \geq \hat{s}$, so the

incumbent employer is the optimal employer, it does not pay to depart. If, instead, $s < \hat{s}$, then the final net pay is $E[s]$ regardless of whether the manager stays or leaves. Let the manager resolve the indifference according to the surplus maximizing assignment. Thus, turnover is efficient. Note also that the incumbent employer cannot gain by making another offer. If the offer is more generous, it creates a loss. If it is less generous, the manager signs another contract.

Second, we show that no subgame-perfect equilibrium supports any other final pay. (i) Contracts that yield an expected pay above $E[s]$ are clearly loss making for a risk-neutral employer. (ii) Suppose there were an equilibrium in which the manager is employed at some contract $\hat{w}$ yielding an expected total pay strictly below $E[s]$. Suppose first that a departing worker is underpaid by a new employer. This is inconsistent with equilibrium because the new employer earns a positive rent; it would be profitable for another employer in the same industry to bid higher. Underpayment must, thus, be associated with states in which the worker is retained. But, if so, the worker is underpaid on average by the initial employer. But this is also inconsistent with equilibrium. If the incumbent employer earns a rent, it would have been profitable for another employer in industry A to bid higher. (iii) Similarly, there cannot be an equilibrium in which the manager obtains the same expected pay but receives nonconstant pay across states as the manager prefers w_u.

Proof of Proposition 5

(i) $E[c(p)] \leq E[s]$ follows from the fact that firms can offer $w(p) = 0$ for all p (such a frivolous contract guarantees a nonnegative profit). (ii) $E[c(p)] \geq E[s]$ follows from competition at stages 1a and 2c entailing a nonpositive profit (the standard Bertrand argument). (iii) $c(p) = \hat{s}(p)$ follows from competition at stage 2c (if the condition were not true, a firm could make a profitable offer at this stage). (iv) The constancy of pay in remaining states follows from competition at stage 1a and the manager's demand for insurance; cf proofs of Propositions 1 and 2.

Proof of Proposition 6

To check that w_1^* is an equilibrium contract, let us begin by characterizing behavior at stage 2. (i) In any state p such that $s_B(p) > \underline{w}$, competition from industry B yields departure and total pay $s_B(p)$ (cf. Lemma 1(v)). (ii) In states satisfying $\underline{w} > s_B(p) > s(p)$, there is departure with discretionary severance pay $\underline{w} > s_B(p)$ (cf. Lemma 1(i)); hence, total pay is $\underline{w}$. (iii) In remaining states, pay is w_1^*. By implication, whenever $w_1(p) \neq \underline{w}$, it coincides with the constrained optimal pay $\hat{s}(p)$. Because all firms earn zero profit, we must have $\underline{w} = \underline{c}$. Thus, the manager indeed obtains the equilibrium compensation described in Proposition 5.

Proof of Corollary 1

Part i. First consider the effect on expected performance pay. To show that expected performance pay is increasing in a mean-preserving spread (MPS), we need to show that

$$\int_{p_h}^{\infty} (p\theta - \underline{w})\tilde{h}(p)dp \geq \int_{p_h}^{\infty} (p\theta - \underline{w})h(p)dp. \tag{A.2}$$

Observe that

$$\begin{aligned}\int_{p_h}^{\infty} (p\theta - \underline{w})h(p)dp &= \int_{0}^{\infty} (p\theta - \underline{w})h(p)dp \\ &\quad - \int_{0}^{p_h} (p\theta - \underline{w})h(p)dp \\ &= \theta/2 - \underline{w} \\ &\quad - \int_{0}^{p_h} (p\theta - \underline{w})h(p)dp \\ &= \theta/2 - \underline{w} \\ &\quad + \theta \int_{0}^{p_h} H(p)dp.\end{aligned}$$

Note that $E[p] = 1/2$ and that turnover is inefficient, so $p \geq p_m$. The last equality follows from integration by parts. By deriving the analogous expression for $\tilde{h}(p)$, it follows that Inequality (A.2) holds if $\int_0^{p_h} H(p)dp \leq \int_0^{p_h} \tilde{H}(p)dp$, which follows from the definition of a MPS. Inequality (A.2) is strict if $\int_0^{p_h} H(p)dp < \int_0^{p_h} \tilde{H}(p)dp$. Second, observe that, if the expected performance pay increases because of an MPS, the wage floor has to decrease to satisfy the firm's break-even constraint.

Part (ii). In a competitive equilibrium, the manager's expected compensation equals the productivity. Hence, we need to show that expected productivity increases with an MPS:

$$\int_0^{\infty} \max\left[(1-p)\alpha, p\right]\tilde{h}(p)dp \geq \int_0^{\infty} \max\left[(1-p)\alpha, p\right]h(p)dp. \tag{A.3}$$

These expressions can be rewritten as follows:

$$\begin{aligned}\int_0^{\infty} \max\left[(1-p)\alpha, p\right]h(p)dp &= \int_0^{\frac{\alpha}{1+\alpha}} \alpha(1-p)\,h(p)dp \\ &\quad + \int_{\frac{\alpha}{1+\alpha}}^{\infty} p\;h(p)dp \\ &= E[p] - \int_0^{\frac{\alpha}{1+\alpha}} p\,h(p)dp + \int_0^{\frac{\alpha}{1+\alpha}} \alpha(1-p)\;h(p)dp \\ &= E[p] + \int_0^{\frac{\alpha}{1+\alpha}} \left[\alpha - p(1+\alpha)\right]h(p)dp \\ &= E[p] + (1+\alpha)\int_0^{\frac{\alpha}{1+\alpha}} H(p)dp.\end{aligned}$$

The last equality follows from integration by parts. Inequality (A.3) holds if $\int_0^{\frac{\alpha}{1+\alpha}} H(p)dp \leq \int_0^{\frac{\alpha}{1+\alpha}} \tilde{H}(p)dp$, which follows from the definition of an MPS. Inequality (A.3) is strict if $\int_0^{\frac{\alpha}{1+\alpha}} H(p)dp < \int_0^{\frac{\alpha}{1+\alpha}} \tilde{H}(p)dp$. Finally, note that an MPS has a larger effect on expected productivity if α is larger.

Proof of Proposition 7

Because the incumbent employer beats any outside offer $w^o < s(p)$, in any equilibrium, a superior outside employer offers $w^o \in [s(p), s^o(p))$, and the manager departs. (We need not consider whether it is optimal for the outside employer to elicit $e = 1$ or not.)

Suppose $g(p) > 0$ for all p. Anticipating this, $\underline{w}$ must be adjusted to preserve zero profit, moving the threshold p_h, but otherwise, there is no gain from adjusting $w(p)$. The only new feature is that all contracts with severance pay $w(p,f') > 0$ are now strictly dominated: when there is a superb outside offer, the manager earns $s(p) + \beta(s^o(p) - s(p)) + w(p,f')$, and because $s^o(p) > \underline{w}$, the contract with positive contracted severance pay has more variable compensation than the contract with zero contracted severance pay.

Endnotes

[1] Related theoretical suggestions have been made by Rosen (1992), Holmström and Kaplan (2003), Hubbard (2005), and especially Holmström and Ricart i Costa (1986) and Oyer (2004). Oyer (2004) is discussed in the introduction; other contributions are discussed in Section 5.

[2] Of course, we do not mean to imply that flawless corporate governance generally implies socially desirable outcomes. See Bénabou and Tirole (2016) for a recent theory of how optimal managerial contracts in a fiercely competitive market can entail severely inefficient outcomes.

[3] For a more nuanced description of the legal context, see Section 2.1.

[4] The second case that Holmström discusses in some detail has infinite mobility costs. Our case of intermediate mobility costs is only mentioned cursorily. Had it been pursued, Holmström's assumption that mobility costs are state-independent would still make the analysis quite different from ours. In particular, his mobility cost assumption would not justify own-firm stock options.

[5] Other structural models of ex post inefficient renegotiation, such as models of incomplete information bargaining, are more complicated. As becomes clear, they are also unlikely to account for the full set of empirical regularities that we aim to explain.

[6] This second case cannot arise in the analysis of Halonen-Akatwijuka and Hart (2020) because they assume risk neutrality.

[7] See Rajgopal et al. (2006) for a previous argument along these lines as well as evidence about lack of indexing of CEO compensation; see Prendergast (2002) for an overview of other relevant evidence. We discuss more of the evidence in Section 4. Needless to say, there are also competing theories that can fit some of this evidence. For example, Raith (2003) shows that changes in competition can introduce a positive relationship between uncertainty and the strength of incentives in an otherwise conventional effort-inducement model. Others argue that observed pay for luck to some extent reflects firms' lack of commitment power; they only keep promises to pay managers more than the salary when they think managers are likely to be somewhat persistently lucky; see DeVaro et al. (2018) and Fahn and Klein (2019).

[8] The notion that fairness is a key concern for lower level employees is more rarely disputed and also empirically rather well established. See, for example, Breza et al. (2018) and the references therein.

[9] Fehr et al. (1993) initiate a large experimental literature on reciprocity; for references, see Fehr et al. (2009a). It is noteworthy that, at least under laboratory circumstances, reciprocity tends to be stronger among business people than among students (e.g., Fehr and List 2004). The more recent wave of CRP experiments extends this earlier literature by involving a renegotiation stage; see Fehr et al. (2009b, 2011, 2015), Kessler and Leider (2012), Bartling and Schmidt (2015), and Brandts et al. (2016). Under such laboratory circumstances, the basic tenet of CRP receives solid support.

[10] For applications of reference point theory to other aspects of the theory of the firm, see Hart (2008, 2009) and Hart and Holmström (2010).

[11] For a complementary perspective that emphasizes private information about parties' characteristics, see Non (2012).

[12] Specifically, renegotiation is better if the managerial labor market is thin, that is, if the manager is not certain to get an outside offer that is as attractive as the compensation specified in the contract; see Proposition 8. (For rank-and-file workers with little industry-specific human capital, there is an additional reason for not contracting on variable pay: there may be no financial security or other verifiable index whose value correlates sufficiently well with their outside option. In particular, there is weak correlation between the own firm's stock value and these workers' attraction to other employers.)

[13] Oyer does not cite Holmström (1983). Instead, he relates his analysis to Harris and Holmström (1982), which makes some of the same assumptions.

[14] These restrictions on the space of contracts also affect the model's relevance for policy discussions. By assuming that variable pay is composed of stocks rather than stock options, Oyer assumes away the asymmetric pay—downward inflexibility and upward flexibility—which is at the center of many critiques about managerial compensation (see Section 4.2).

[15] For a treatment of employers' rights to terminate, see Verkerke (2009).

[16] Noncompete arrangements are not confined to top managers. For a recent analysis of the causes and consequences of noncompete arrangements in a representative sample of U.S. employees, see Starr et al. (2020).

[17] For simplicity, we refrain from introducing other managers, implicitly assuming that we can neglect issues of scale and scope of firms. Likewise, we disregard the question of who or what it is that the managers manage.

[18] Portability of human capital is affected both by technology and institutions; for empirical illustrations and relevant references, see, for example, Groysberg et al. (2008) and Marx et al. (2009).

[19] We interpret effort broadly as any action that affects the outcome. It is straightforward to extend the model to allow the firm to force some nonzero minimum effort level.

[20] The lack of direct concern for e is not to say that managers necessarily find each action equally onerous; it could just mean that managers' actions are guided more by professionalism and decency than by desire for an easy life.

[21] If $\tau < 1$, it would be optimal to avoid aggrievement by offering the manager a large enough fraction of the relationship rent. In this case, contract renegotiation would still entail ex ante inefficiency, but now, all the inefficiency is due to a shortage of risk sharing.

[22] The only difference from Hart and Moore (2008) is that our formulation admits risk aversion.

[23] Because total pay is larger in each state that variable pay is positive, fixed pay must be lower in order to keep total expected pay constant.

[24] Perhaps repeating it might serve to defend against the suggestion, which we have sometimes heard from experts who ought to know better, that employers can easily and legally handle the retention problem by offering low initial salaries coupled with large retention bonuses. That argument rests on the unrealistic assumption that credit markets are perfect.

[25] This is easily seen by positioning the relevant horizontal line for αp_B in Figure 3.

[26] For example, the following model can be easily articulated and solved, but we merely sketch it here: Suppose firms are horizontally differentiated on a (Salop-) circle. Customers have heterogeneous valuations so that a customer values firm i's product at v_i minus a

linear transportation cost, but let heterogeneity be small enough that each firm only effectively competes with the two closest neighbors in the product space. Similarly, let these two closest neighbors be the only candidates for recruiting the manager, whose human capital is also more portable to similar firms. Then, if the identities of the closest firms are known at date 1, the optimal contract conditions on the stock price of that neighboring firm that has the highest willingness to pay for the manager's services. Thus, under this scenario, it is no longer optimal to condition pay on the original employer's stock price. But, if we instead suppose that locations on the circle and, thus, the identity of neighbors, are not known at date 1, it is unwise at date 1 to tie pay to the performance of a particular competitor. Instead, the own stock price again becomes the best measure at the contracting date for capturing what close competitors (neighbors) can come to offer in the future. At least for industries with much innovation and unstable firm rankings, we consider this scenario to be the most plausible. Thus, own-firm stock options are a robust feature of the model.

[27] We might also consider negative idiosyncratic shocks, for example, that the manager becomes disaffected. Let there be a probability $\xi > 0$ of a large negative shock to the manager's nonpecuniary benefit from working at the original firm, making turnover desirable regardless of the state p. With this additional probability of separation, the salary component clearly goes down to reflect the lower expected surplus, but in all other respects, the simplest optimal contract remains the same as before. Note also that this new turnover is most likely to occur within the industry (because $h(p)$ is symmetric and $p_m < 1/2$; see Figure 2). As with the possibility of superb offers, this extension, thus, allows the model to better fit the evidence that much managerial turnover occurs within industries.

[28] We are not aware of any empirical work explicitly investigating the relationship between managerial market thickness and contracted performance pay. It could be an interesting avenue for future research.

[29] Ellingsen and Kristiansen (2011) advance a similar argument for the optimality of simple securities.

[30] The manager having constant absolute risk aversion is presumably necessary but not sufficient.

[31] As a result of higher expected costs of compensation, the firm's expected profit goes down as does the stock price.

[32] An analysis of this interplay offers a new argument against the use of noncompete clauses.

[33] A related finding is that mobility is smaller when firms defend their patents more aggressively; see Ganco et al. (2015).

[34] The skimming hypothesis focuses on the downside of managerial power. Of course, in a world of incomplete contracts, giving managers influence can, in principle, be optimal; see Almazan and Suarez (2003) for an analysis of optimal managerial entrenchment and contracted severance pay.

[35] "First, our suggestive evidence of asymmetry in pay for luck may be hard to reconcile with [the retention] view. Average CEO compensation in the oil industry always goes up when the price of crude oil goes up but does not always go down when the price of crude oil goes down … "

[36] "We found no statistically significant relationship between a CEO's turnover and industry returns (after controlling for the firm's returns) and a point estimate that was negative. This suggests that, if anything, turnover is countercyclical."

[37] "If pay for luck were caused by market competition for CEOs, then industries with higher turnover should exhibit the greatest pay for luck. For accounting measures, we found that industries with the highest turnover in fact showed the least pay for luck. For market measures of performance, we found no relationship between industry turnover and pay for luck."

[38] Of course, better governance in a particular firm would allow that firm to reduce its average pay because the riskiness of its pay is lower, a feature that is attractive to risk-averse managers. However, compared with alleged levels of overpayment, the potential for reducing average pay because of risk reduction would seem modest.

[39] Among other governance reforms, transparency reforms also have as one of their purposes to curb excess compensation. Perhaps surprisingly, transparency reforms seem to be associated with higher pay rather than lower pay; see, for example, Perry and Zenner (2001) and Gipper (2020). Although these findings go against the skimming view, our retention view does not obviously offer a particularly helpful perspective on them (except that reference-dependent preferences might play a similar role).

[40] In the 1860s, the default was that the general assembly determined managerial pay, so shareholder say on pay is not a modern invention.

[41] Because their identifying assumptions are very different, it remains to fully understand how the findings from this international study reconciles with the evidence from the single-country studies.

[42] As noted by Ferri and Göx (2018), market reactions to say-on-pay legislation have tended to be positive. Thus, these long-term negative effects must be rather small if they are to be outweighed by one-time rent extraction. However, to the extent that say-on-pay has large disciplining effects of the kind discussed by Cuñat et al. (2016), there is more scope for negative effects from hold-up and inefficient risk-sharing to be large as well.

[43] When competition among employers is considered, there is typically more emphasis on how it influences job assignments and internal career patterns than on the optimal the contractual shapes; see Waldman (1984) for an early theoretical contribution and Waldman (2012) for a recent survey of the internal labor market literature.

[44] See, for example, the surveys of Murphy (1999) and Lazear and Oyer (2012).

[45] For a recent summary of what effort-inducement models can and cannot explain in the domain of executive compensation, see Edmans and Gabaix (2016).

[46] In Holmström and Ricart i Costa (1986), the option value is linear in the stock price when there is a strong impact of worker's ability on the firm's value, which is only true for exceptionally important employees.

[47] Models that attempt to explain how option packages vary with firm and market conditions, such as Johnson and Tian (2000), exogenously impose a combination of salary and options. Among previous theoretical models of compensation contracts that consider the retention motive, Hashimoto (1979) and Blakemore et al. (1987) assume that contracts are piece-wise linear. Oyer (2004) and Giannetti (2011) assume linear contracts. Dutta (2003) derives a linear contract from first principles. All three, thus, fail to account for the lower bound to payments. Finally, Pakes and Nitzan (1983) examine how contracts can be designed to retain research personnel. Their focus is similar to ours, but the contract that they derive is generally not linear in performance and depends on the nature of output market competition.

[48] As noted by Rosen (1981), already Adam Smith understood that small differences in talent could translate into massive differences in compensation levels in a competitive market.

[49] Bénabou and Tirole (2016) is another important contribution to this literature but with a different focus. They emphasize inefficiencies that arise when competition for talented managers distort effort in a multitask setting. As mentioned, Englmaier et al. (2014) consider the issues that arise when current effort affects the value of the manager both inside the firm and outside of it.

[50] This is not to claim that the retention-based model is quantitatively superior. Comparisons between the various approaches are likely to depend on the domain and to require empirical investigations that are designed with the explicit purpose of evaluating the relative importance of the different mechanisms.

[51] See also Halonen-Akatwijuka and Hart (2013) and Herweg and Schmidt (2015) for previous analyses of inefficient renegotiation.

[52] For a particularly clear exposition, see Osborne and Rubinstein (1990, chapter 3).

References

Acharya VV, John K, Sundaram RK (2000) On the optimality of resetting executive stock options. *J. Financial Econom.* 57(1):65–101.

Akerlof G (1982) Labor contracts as partial gift exchange. *Quart. J. Econom.* 97(4):543–569.

Almazan A, Suarez J (2003) Entrenchment and severance pay in optimal governance structures. *J. Finance* 58(2):519–547.

Anderson RC, Reeb D (2003) Founding-family ownership and firm performance: Evidence from S & P 500. *J. Finance* 58(3): 1301–1328.

Anderson MC, Banker RD, Ravindran S (2000) Executive compensation in the information technology industry. *Management Sci.* 46(4):530–547.

Bandiera O, Guiso L, Prat A, Sadun R (2015) Matching firms, managers and incentives. *J. Labor Econom.* 33(3):623–681.

Bartling B, Schmidt KM (2015) Reference points, social norms, and fairness in contract renegotiations. *J. Eur. Econom. Assoc.* 13(1):98–129.

Bebchuk LA, Fried JM (2004) *Pay Without Performance: The Unfulfilled Promise of Executive Compensation* (Harvard University Press, Cambridge, MA).

Bebchuk LA, Fried JM, Walker DI (2002) Managerial power and rent extraction in the design of executive compensation. *Univ. Chicago Law Rev.* 69(3):751–846.

Bell B, Pedemonte S, van Reenen J (2021) CEO pay and the rise of relative performance contracts: A question of governance. *J. Eur. Econom. Assoc.* Forthcoming.

Bénabou R, Tirole J (2016) Bonus culture: Competitive pay, screening, and multitasking. *J. Political Econom.* 124(2):305–370.

Bertrand M, Mullainathan S (2001) Are CEOs rewarded for luck? The ones without principals are. *Quart. J. Econom.* 116(3): 901–932.

Bettis C, Bizjak J, Coles J, Kalpathy S (2010) Stock and option grants with performance-based vesting provisions. *Rev. Financial Stud.* 23(10):3849–3888.

Binmore KG, Shaked A, Sutton J (1989) An outside option experiment. *Quart. J. Econom.* 104(4):753–770.

Blakemore AE, Low SA, Ormiston MB (1987) Employment bonuses and labor turnover. *J. Labor Econom.* 5S:S114–S135.

Bosse DA, Phillips RA (2016) Agency theory and bounded self-interest. *Acad. Management Rev.* 41(2):276–297.

Brandts J, Charness G, Ellman M (2016) Let's talk: How communication affects contract design. *J. Eur. Econom. Assoc.* 14(4):943–974.

Breza E, Kaur S, Shamdazani Y (2018) The morale effects of pay inequality. *Quart. J. Econom.* 133(2):611–663.

Core JE, Guay WR (2001) Stock option plans for non-executive employees. *J. Financial Econom.* 61(2):253–287.

Correa R, Lel U (2016) Say on pay laws, executive compensation, pay slice, and firm valuation around the world. *J. Financial Econom.* 122(3):500–520.

Cremers M, Grinstein Y (2014) Does the market for CEO talent explain controversial CEO pay practices? *Rev. Finance* 18(3): 921–960.

Cuñat V, Giné M, Guadalupe M (2016) Say pays! Shareholder voice and firm performance. *Rev. Finance* 20(5):1799–1834.

de Meza D, Webb DC (2007) Incentive design under loss aversion. *J. Eur. Econom. Assoc.* 5(1):66–92.

DeVaro J, Kim J-H, Vikander N (2018) Non performance pay and relational contracting: Evidence from CEO compensation. *Econom. J. (London)* 128(613):1923–1951.

Dittmann I, Maug E, Spalt O (2010) Sticks or carrots? Optimal CEO compensation when managers are loss-averse. *J. Finance* 65(6):2015–2050.

Dutta S (2003) Capital budgeting and managerial compensation: Incentive and retention effects. *Accounting Rev.* 78(1):71–93.

Edmans A, Gabaix X (2016) Executive compensation: A modern primer. *J. Econom. Literature* 54(4):1232–1287.

Edmans A, Gabaix X, Landier A (2009) A multiplicative model of optimal CEO incentives in market equilibrium. *Rev. Financial Stud.* 22(12):4881–4917.

Ellingsen T, Kristiansen EG (2011) Financial contracting under imperfect enforcement. *Quart. J. Econom.* 126(1):323–371.

Englmaier F, Leider S (2012) Contractual and organizational structure with reciprocal agents. *Amer. Econom. J. Microeconom.* 4(2):146–183.

Englmaier F, Muehlheusser G, Roider A (2014) Optimal incentive contracts for knowledge workers. *Eur. Econom. Rev.* 67(1): 82–106.

Estlund CL (2006) Between rights and contract: Arbitration agreements and non-compete covenants as a hybrid form of employment law. *Univ. Pennsylvania Law Rev.* 155:379–445.

Fahlenbrach R (2009) Shareholder rights, boards, and executive compensation. *Rev. Finance* 13(1):81–113.

Fahn M, Klein N (2019) Relational contracts with private information on the future value of the relationship: The upside of implicit downsizing costs. *Amer. Econom. J. Microeconom.* 11(4):33–58.

Fehr E, List J (2004) The hidden costs and returns of incentives—Trust and trustworthiness among CEOs. *J. Eur. Econom. Assoc.* 2(5):743–771.

Fehr E, Goette L, Zehnder C (2009a) A behavioral account of the labor market: The role of fairness concerns. *Annual Rev. Econom.* 1:355–384.

Fehr E, Hart O, Zehnder C (2009b) Contracts, reference points, and competition–behavioral effects of the fundamental transformation. *J. Eur. Econom. Assoc.* 7(2–3):561–572.

Fehr E, Hart O, Zehnder C (2011) Contracts as reference points—Experimental evidence. *Amer. Econom. Rev.* 101(2):493–525.

Fehr E, Hart O, Zehnder C (2015) How do informal agreements and revision shape contractual reference points? *J. Eur. Econom. Assoc.* 13(1):1–28.

Fehr E, Kirchsteiger G, Riedl A (1993) Does fairness prevent market clearing? An experimental investigation. *Quart. J. Econom.* 108(2):437–459.

Fernandes N, Ferreira MA, Matos P, Murphy KJ (2013) Are U.S. CEOs paid more? New international evidence. *Rev. Financial Stud.* 26(2):323–367.

Ferri F, Göx RF (2018) Executive compensation, corporate governance, and say on pay. *Foundations Trends Accounting* 12(1): 1–103.

Ferri F, Maber DA (2013) Say on pay votes and CEO compensation: Evidence from the UK. *Rev. Finance* 17(2):527–563.

Fong EA, Misangyi VF, Tosi HL (2010) The effect of CEO pay deviations on CEO withdrawal, firm size, and firm profits. *Strategic Management J.* 31(6):629–651.

Frydman C, Saks R (2010) Executive compensation: A new view from a long-term perspective, 1936–2005. *Rev. Financial Stud.* 23(5):2099–2138.

Gabaix X, Landier A (2008) Why has CEO pay increased so much? *Quart. J. Econom.* 123(1):49–100.

Ganco M, Ziedonis R, Agarwal R (2015) More stars stay, but the brightest ones still leave: Job hopping in the shadow of patent enforcement. *Strategic Management J.* 36(5):659–685.

Garmaise MJ (2011) Ties that truly bind: Non-competition agreements, executive compensation and firm investment. *J. Law Econom. Organ.* 27(2):376–425.

Gartenberg C, Wulf J (2017) Pay harmony: Social comparison and performance compensation in multibusiness firms. *Organ. Sci.* 28(1):39–55.

Garvey G, Milbourn TT (2006) Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad. *J. Financial Econom.* 82(1):197–225.

Gaver JJ, Gaver KM (1993) Additional evidence on the association between the investment opportunity set and corporate financing, dividend, and compensation policies. *J. Accounting Econom.* 16(1–3):125–160.

Giannetti M (2011) Serial CEO incentives and the structure of managerial contracts. *J. Financial Intermediation* 20(4):633–662.

Gipper B (2021) The economic effects of expanded compensation disclosure. *J. Accounting Econom.* Forthcoming.

Goldman EM, Huang PP (2015) Contractual vs actual separation pay following CEO turnover. *Management Sci.* 61(5):1108–1120.

Gomez-Mejia LR, Larraza-Kintana M, Makir M (2003) The determinants of executive compensation in family-controlled public corporations. *Acad. Management J.* 46(2):226–237.

Groysberg B, Lee L-E, Nanda A (2008) Can they take it with them? The portability of star knowledge workers' performance. *Management Sci.* 54(7):1213–1230.

Halonen-Akatwijuka M, Hart O (2013) *More Is Less: Why Parties May Deliberately Write Incomplete Contracts* (Harvard University).

Halonen-Akatwijuka M, Hart O (2020) Continuing contracts. *J. Law Econom. Organ.* 36(2):284–313.

Harris M, Holmström B (1982) A theory of wage dynamics. *Rev. Econom. Stud.* 49(3):315–333.

Hart O (2008) Reference points and the theory of the firm. *Economica* 75(299):404–411.

Hart O (2009) Hold-up, asset ownership, and reference points. *Quart. J. Econom.* 124(1):267–300.

Hart O, Holmström B (2010) A theory of firm scope. *Quart. J. Econom.* 125(2):483–513.

Hart O, Moore J (2008) Contracts as reference points. *Quart. J. Econom.* 123(1):1–48.

Hashimoto M (1979) Payments, on-the-job training, and lifetime employment in Japan. *J. Political Econom.* 87(5):1086–1104.

Herweg F, Schmidt KM (2015) Loss aversion and inefficient renegotiation. *Rev. Econom. Stud.* 82(1):297–332.

Herweg F, Müller D, Weinschenk P (2010) Binary payment schemes: Moral hazard and loss aversion. *Amer. Econom. Rev.* 100(5):2451–2477.

Himmelberg CP, Hubbard RG, Palia D (1999) Understanding the determinants of managerial ownership and the link between ownership and performance. *J. Financial Econom.* 53(3):353–384.

Holmström B (1983) Equilibrium long-term labor contracts. *Quart. J. Econom.* 98(Supplement):23–54.

Holmström B, Kaplan S (2003) The state of US corporate governance. *J. Appl. Corporate Finance* 15(3):8–20.

Holmström B, Ricart i Costa J (1986) Managerial incentives and capital management. *Quart. J. Econom.* 101(4):835–860.

Hubbard RG (2005) Pay without performance: A market equilibrium critique. *J. Corporate Law.* 30(Summer):717–720.

Iliev P, Vitanova S (2019) The effects of say on pay in the U.S. *Management Sci.* 65(10):4505–4521.

Inderst R, Mueller HM (2010) CEO replacement under private information. *Rev. Financial Stud.* 23(8):2935–2969.

Ittner CD, Lambert RA, Larcker DF (2003) The structure and performance consequences of equity grants to employees of new economy firms. *J. Accounting Econom.* 34(1–3):89–127.

Johnson SA, Tian YS (2000) The value and incentive effects of nontraditional executive stock option plans. *J. Financial Econom.* 57(1):3–34.

Kessler J, Leider S (2012) Norms and contracting. *Management Sci.* 58(1):62–77.

Kole SR (1997) The complexity of compensation contracts. *J. Financial Econom.* 43(1):79–104.

Kvaløy O, Olsen T (2012) The rise of individual performance pay. *J. Econom. Management Strategy* 21(2):493–518.

Lazear EP, Oyer P (2012) Personnel economics. Gibbons R, Roberts DJ, eds. *Handbook of Organizational Economics*, vol. 3 (Elsevier, New York), 479–519.

Marx M, Strumsky D, Fleming L (2009) Mobility, skill, and the Michigan non-compete experiment. *Management Sci.* 55(6):875–889.

Mehran H (1995) Executive compensation structure, ownership, and firm performance. *J. Financial Econom.* 38(2):163–184.

Murphy KJ (1999) Executive compensation. Aschenfelter O, Card D, eds. *Handbook of Labor Economics*, vol. 3B (Elsevier, New York), 2485–2563.

Murphy KJ (2003) Stock-based pay in new economy firms. *J. Accounting Econom.* 34(1–3):129–147.

Murphy K, Zabojnik J (2004) CEO pay and appointments: A market-based explanation for recent trends. *Amer. Econom. Rev.* 94:192–196.

Murphy K, Zabojnik J (2006) Managerial capital and the market for CEOs. Queen's Economics Department Working Paper No. 1110.

Non A (2012) Gift-exchange, incentives, and heterogeneous workers. *Games Econom. Behav.* 75(1):319–336.

Ockenfels A, Sliwka D, Werner P (2014) Bonus payments and reference point violations. *Management Sci.* 61(7):1496–1513.

Osborne MJ, Rubinstein A (1990) *Bargaining and Markets* (Academic Press, San Diego).

Oyer P (2004) Why do firms use incentives that have no incentive effects? *J. Finance* 59(4):1619–1649.

Oyer P, Schaefer S (2005) Why do some firms give stock options to all employees? An empirical examination of alternative theories. *J. Financial Econom.* 76(1):99–133.

Pakes A, Nitzan S (1983) Optimal contracts for research personnel, research employment, and the establishment of" rival" enterprises. *J. Labor Econom.* 1(4):345–365.

Palia D (2001) The endogeneity of managerial compensation in firm valuation: A solution. *Rev. Financial Stud.* 14(3):735–764.

Perry T, Zenner M (2001) Pay for performance? Government regulation and the structure of compensation contracts. *J. Financial Econom.* 62(3):453–488.

Prendergast C (2002) The tenuous trade-off between risk and incentives. *J. Political Econom.* 110(5):1071–1102.

Raith M (2003) Competition, risk and managerial incentives. *Amer. Econom. Rev.* 93(4):1425–1436.

Rajan R, Zingales L (2001) The firm as a dedicated hierarchy: A theory of the origins and growth of firms. *Quart. J. Econom.* 116(3):805–851.

Rajgopal S, Shevlin T, Zamora V (2006) CEOs' outside employment opportunities and the lack of relative performance evaluation in compensation contracts. *J. Finance* 61(4):1813–1844.

Rosen S (1981) The economics of superstars. *Amer. Econom. Rev.* 71(5):845–858.

Rosen S (1992) Contracts and the market for executives. Werin L, Wijkander H, eds. *Contract Economics* (Blackwell, Cambridge, MA and Oxford), 181–211.

Rubinstein A (1982) Perfect equilibrium in a bargaining model. *Econometrica* 50(1):97–109.

Schwab SJ, Thomas RS (2006) An empirical analysis of CEO employment contracts: What do top executives bargain for? *Washington Lee Law Rev.* 65(1):231–270.

Smith CJ, Watts RL (1992) The investment opportunity set and corporate financing, dividend, and compensation policies. *J. Financial Econom.* 32(3):263–292.

Starr E, Prescott JJ, Bishara ND (2021) Noncompete agreements in the U.S. labor force. *J. Law Econom.* Forthcoming.

Terviö M (2008) The difference that CEOs make: An assignment model approach. *Amer. Econom. Rev.* 98(3):642–668.

Verkerke JH (2009) Discharge. Dau-Schmidt KG, Harris S, Lobel O, eds. *Labor and Employment Law and Economics*, vol. 2, Encyclopedia of Law and Economics, 2nd ed. (Edward Elgar Publishing, Northampton, UK), 447–479.

Waldman M (1984) Job assignments, signalling, and efficiency. *RAND J. Econom.* 15(2):255–267.

Waldman M (2012) Theory and evidence on internal labor markets. Gibbons R, Roberts J, eds. *Handbook of Organizational Economics* (Princeton University Press, Princeton, NJ), 520–571.

Yermack D (2006) Golden handshakes: Separation pay for retired and dismissed CEOs. *J. Accounting Econom.* 41(3):237–256.