



Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **14-06**

Explaining the Association between Monitoring and Controversial CEO Pay Practices: an Optimal Contracting Perspective

Pierre Chaigneau
Nicolas Sahuguet

Février/February 2014

Chaigneau: Corresponding author. HEC Montréal and CIRPÉE. Address: Department of Finance, HEC Montréal, 3000 chemin de la Côte-Sainte-Catherine, Montréal, QC, Canada H3T 2A7. Tel.: 514 340-5661; Fax: 514 340-5632
pierre.chaigneau@hec.ca

Sahuguet: HEC Montréal and CIRPÉE

We thank Marco Becht, Vicente Cunat, Ingolf Dittmann, Alex Edmans, Robert Gibbons, Sergei Kovbasyuk, Patrick Legros, Paul Oyer, Nicola Persico, and Lucy White for interesting comments, as well as participants in seminars at Ecares-Université Libre de Bruxelles, Erasmus University Rotterdam, ESSEC, HEC Montréal, University of North Carolina, the Association Française de Finance December 2011 Conference, the ESSET 2012 Conference, the Petralia economics Workshop 2012, and the 2013 European Economic Association Conference.

Abstract:

Puzzling associations between low levels of ownership concentration and CEO pay practices such as pay-for-luck, a low pay-performance sensitivity, a more asymmetric pay-performance relation, and high salaries, have been documented. They have been interpreted as evidence that CEO pay is not set optimally. We explain these associations in a model in which firms design contracts optimally to attract and retain CEOs. The results are driven by the matching process: firms with greater ownership concentration have a higher monitoring capacity, and can better handle the downside risk of hiring CEOs with more uncertain ability. The outside option of these CEOs is more sensitive to their performance net of luck, which generates a higher pay-performance sensitivity and less pay-for-luck. If managerial skills are sufficiently transferable across firms and the cost of CEO dismissal is sufficiently high, these CEOs are less valuable and therefore receive relatively lower salaries.

Keywords: CEO pay, corporate governance, monitoring, pay-for-luck

JEL Classification: D86, G34, M12

In recent years, CEO pay has attracted considerable attention, both in the popular press and in academic journals. This renewed interest was in part triggered by some puzzling observations such as the evidence that CEOs are paid for luck (Bertrand and Mullainathan (2001)). Based on these observations, Bertrand and Mullainathan (2001) and Bebchuk and Fried (2003), among others, argue that the optimal contracting approach fails. Their argument is based not so much on the *existence* of phenomena such as pay-for-luck – some of which have been explained, see for example the literature review of Edmans and Gabaix (2009) – but on the fact that their extent is *greater* in firms with worse governance. Specifically, Bertrand and Mullainathan (2001) find that pay-for-luck is stronger in firms with worse monitoring or governance, while Bebchuk and Fried (2003) argue that CEO pay will be higher *and* less sensitive to performance in this type of firm, a statement for which there is some empirical evidence (e.g. Core, Holthausen, and Larcker (1999) and Hartzell and Starks (2003)). As these associations would be puzzling in view of an optimal contracting framework with one firm and one CEO, they have contributed to the development of an alternative paradigm of CEO compensation, the “managerial power” or “skimming” approach.

In this paper, we show that these associations are actually consistent with a model of optimal contracting. We complement a standard model of career concerns by introducing firms with different capacities for monitoring, and more or less risky CEOs. In equilibrium, poor monitoring is associated with more pay-for-luck, a sensitivity of CEO pay to firm performance which is lower on average, and more concentrated on the upside than the downside, and higher CEO salaries. This is not because monitoring directly affects any of these dimensions of CEO compensation, but instead because it affects the matching between CEOs and firms: different types of firms hire different types of CEOs, and therefore offer different contracts. The paper thus brings a theoretical response to some as yet unchallenged fundamental arguments of the managerial power theory.

The model is a two-period setting in which firms hire, compensate, and fire CEOs. The ability of CEOs is uncertain and unknown to all parties, and the variance of ability differs across CEOs. For example, the variance of ability tends to be lower for CEOs who are older, who have a specific degree (e.g., MBA), who possess prior industry-specific experience and management experience. Firm performance in the first period provides information about the ability of its CEO, so that it affects the CEO’s second period outside option. As in career concerns models (e.g., Harris and Holmstrom (1982)), a CEO’s outside option depends on the market’s updated belief about his ability. The outside option also depends on the transferability of CEO ability

across firms, and on business conditions or “luck” (as in Oyer (2004)). In this setting, the compensation contracts are designed to match the state-contingent outside options of the CEOs, as determined in a market equilibrium. They are thus driven by market forces, and do not result in rent extraction by the CEOs. We also assume that firms are heterogeneous and differ in their ownership structure. As in Burkart, Gromb, and Panunzi (1997), more concentrated ownership leads to more extensive monitoring of the CEO. Monitoring generates signals on CEO ability, which facilitates CEO dismissal (Cornelli, Kominek, and Ljungqvist (2013)).¹

We determine the optimal matching between CEOs and firms, and derive the optimal contracts resulting from the matching equilibrium without imposing any restriction on the contracting space. In equilibrium, firms with better monitoring are matched with CEOs with more uncertain ability. Intuitively, hiring a CEO with a more uncertain ability increases the upside potential of hiring a very talented CEO whose skills are only partially transferable and are thus partly captured by the firm.² In addition, due to imperfect monitoring, it increases the downside risk of keeping a CEO with low ability in place. Greater monitoring increases the capacity of the firm to handle the downside risk, hence the equilibrium matching.

The model generates a series of associations between the intensity of monitoring and CEO pay practices which are consistent with the empirical facts often presented by the managerial power theory as evidence of the failure of the optimal contracting paradigm. First, the model can explain the important finding that “pay for luck diminishes with the presence of a large shareholder” (Bertrand and Mullainathan (2001)), and the positive relation between pay-performance sensitivity (PPS) and institutional ownership concentration (Hartzell and Starks (2003)). Indeed, firms with concentrated ownership and the associated greater monitoring capacity are more willing to hire CEOs whose ability is more uncertain ex-ante. Since firm performance is driven by CEO ability, noise, and luck, firm performance is more informative about CEO ability when this ability is more uncertain ex-ante. It follows that second-period state-contingent pay puts a higher weight on first-period firm performance net of luck for CEOs whose ability is less precisely estimated ex-ante. Thus, the sensitivity of pay to performance will be higher, and there will be less pay-for-luck, in firms with better monitoring. Note that monitoring capacity does not have any direct effect on these two components of

¹In Kuhn and Zwiebel (2009), CEO entrenchment is driven by the cost associated to CEO turnover. In their model, CEOs set their own pay subject to a limited entrenchment constraint. In our model, “entrenchment” is driven by the limited monitoring capacity of the firm’s shareholders, so that the firm optimally sets CEO pay and dismisses the CEO subject to limited monitoring.

²In Hermalin (2005), retention is not an issue and dismissal is costless, so the firm always values uncertainty about the CEO’s ability.

CEO pay – the association follows from the matching between CEOs and firms.³

Second, the model can explain that CEOs in firms with less concentrated ownership are paid higher salaries (see Core, Holthausen, and Larcker (1999), and Hartzell and Starks (2003)). When managerial skills are sufficiently transferable across firms and CEO dismissal is sufficiently costly, CEOs with a more precisely estimated ability are more valuable and therefore receive higher salaries. This is in contrast to the model of Hermalin (2005), in which uncertainty about CEO ability has a positive option value. Intuitively, because of the competition among firms and of the transferability of managerial skills, CEOs capture most of the gains associated with a good first-period performance. In addition, when CEO dismissal is costly, firms can be stuck with bad CEOs – which is all the more likely when CEO ability is more uncertain. The association between high salaries and poor monitoring then follows from the equilibrium matching, given that CEOs with more precisely estimated ability are matched with firms with less concentrated ownership and worse monitoring.

The model can also shed light on the effects of *changes* in monitoring capacity on CEO pay. We show that an improvement in monitoring capacity, whether across the board or confined to the worst firms on this dimension, has spillover effects that increase CEO pay in *all* firms, including those whose monitoring capacity is unchanged.⁴ This is because better monitoring increases the value of a CEO with uncertain ability, which increases CEO pay. In addition, and more importantly, an improvement in the monitoring capacity of the worst firms reduces monitoring heterogeneity, which increases competition for CEOs and reduces the rents earned by firms with superior monitoring capacity, thus increasing CEO pay in all firms. This paper is to our knowledge the first to identify this effect. The facts that CEO pay increased *and* monitoring capacities improved in the past decades (Huson, Parrino, and Starks (2001)), notably via the diffusion of best practices, are hard to reconcile with the managerial power theory, but are consistent with our model. We also find that a firm-specific change in monitoring capacity does not necessarily have an effect on CEO turnover (because of a sorting effect), and it does not affect CEO pay. This is contrary to what the managerial

³The logic of the result is reminiscent of Akerberg and Botticini (2002) who show in the context of sharecropping contracts that endogenous matching can explain apparent discrepancies between theoretical predictions and empirical findings.

⁴In Hermalin (2005), an increase in monitoring capacity reduces CEO utility – which in turn requires an increase in CEO pay – and it increases CEO turnover, but there are no governance spillovers. The spillovers are related to the work of Acharya and Volpin (2010) and Dicks (2012), who show that corporate governance in a firm may generate an externality and influence the compensation of CEOs in other firms. Note that the externality that we identify does not affect Pareto efficiency and therefore does not call for a regulatory intervention.

power theory would predict,⁵ but our results are in line with the evidence in Kaplan and Minton (2012) and Cunat, Gine, and Guadalupe (2012), respectively. This distinction between the effects of economy-wide and firm-specific changes in monitoring or governance should be considered in future empirical studies.

Other predictions of the model are also consistent with the empirical evidence on CEO compensation. As in Harris and Holmstrom (1982), risk neutral firms insure risk averse CEOs against negative updating on their ability. CEO pay is thus downward rigid (consistent with the empirical evidence in Taylor (2013)). This rigidity explains the absence of financial punishment for poor performance and asymmetric pay-for-luck (the fact that CEO pay is less sensitive to bad luck is documented in Bertrand and Mullainathan (2001) and Garvey and Milbourn (2006)). In addition, if firm value is multiplicative in the CEO ability and the luck shock (as in Gabaix and Landier (2008)), then the value of the CEO's outside option is shown to be *linear* in a measure of firm value that does not fully filter out luck. This is in contrast with principal-agent models of effort provision with risk averse CEOs, in which such linearity is elusive (e.g., Dittmann and Maug (2007), Edmans and Gabaix (2011)). In our model, a positive shock to either CEO ability or business conditions raises both firm value and the market value of the CEO – but only to the extent that CEO ability is transferable across firms. Thus, the PPS increases as general managerial skills become relatively more important, which was arguably the case in the 1980s and the 1990s (Murphy and Zabojnik (2004) and Frydman (2007)). The model can therefore contribute to explain the rise of stock-options based compensation in this period.⁶

The assumption at the core of our analysis that retention is an important determinant of CEO compensation has been tested in the literature. Gabaix and Landier (2008) find strong empirical support for a model in which the level of CEO pay is jointly determined in a competitive market by the distribution of CEO talent and firm size (see also Tervio (2008) and Falato, Li, and Milbourn (2012)). Lazear (2004) and Oyer and Schaefer (2005) emphasize the limitations of the incentives-based explanation for the adoption of variable pay and broad-based stock-options plans, respectively. Rajgopal, Shevlin, and Zamora (2006) present evidence that CEO pay is structured to match the state-contingent outside employment opportunities of managers. Eisfeldt and Kuhnen (2013) find that a competitive assignment model can explain a number of patterns related to CEO turnover. The present paper contributes to this growing lit-

⁵In the managerial power model of Kuhnen and Zwiebel (2009), a CEO who gets more “entrenched” is paid more, *ceteris paribus*.

⁶By contrast, the models of Murphy and Zabojnik (2004) and Frydman (2007) relate the increasing importance of general skills to the increase in the *level* of CEO compensation.

erature which shows that both the level and the form of CEO pay can be explained by retention motives and changes in reservation wages. In particular, our results contrast with the arguments that it is inefficient to use equity-based compensation for retention purposes (Hall and Murphy (2003), Lazear (2004)).⁷

This paper is to our knowledge the first in which the level and structure of CEO pay are designed to retain the CEO given his outside options, which are endogenously determined in a matching equilibrium. Cao and Wang (2013) also study a dynamic market equilibrium in which the outside options are endogenously derived, in order to analyze the relation between the different types of risk faced by the firm and the PPS. In their model, the role of the PPS is to elicit effort, and CEO ability is neither transferable across firms nor persistent over time⁸ – in contrast, (partial) transferability and (partial) persistence of CEO ability are crucial for our results. As in Oyer (2004), we emphasize that state-contingent pay can be used for retention purposes. In addition, we introduce learning on CEO ability, as well as matching between CEOs and firms, which makes the outside option endogenous. Whereas the dimensions of matching used by Gabaix and Landier (2008) enable to explain stylized facts related to the *level* of CEO pay, the dimensions of matching that we introduce enable to explain a number of stylized facts related to the *structure* of CEO pay. Most notably, this is the first model of efficient contracting to explain important associations between firms’ monitoring capacities and CEO pay practices, and to relate the diffusion of best practices to badly governed firms to an across-the-board rise in CEO pay.

1 The model

1.1 Environment

The model builds on the career concerns models of Harris and Holmstrom (1982) and Beaudry and di Nardo (1991). We consider a two-period economy in which firms compete for CEOs. In both periods, the gross profits of a firm (before compensation of the CEO) depend on three factors: the CEO’s ability \tilde{a} , business conditions \tilde{L} , and an idiosyncratic shock $\tilde{\epsilon}_t$. The gross profits in period t , for $t \in \{1, 2\}$, are realized at

⁷According to Lazear (2004), “Worker retention is not a justification for awarding non-vested stock options (...) To the extent that the typical worker is more risk-averse than the outside suppliers of capital, non-vested pay should take the form of bonds rather than equity.” According to Hall and Murphy (2003), “Options clearly provide retention incentives, but do they do so in the most efficient manner? (...) it is not obvious to us that retention incentives should optimally vary with company stock prices.”

⁸More precisely, there is no CEO ability in their model, but a “match quality” between a CEO and a firm, which is specific to the pair and not persistent over time.

the end of the period and take the following form:

$$\pi_t = (\alpha + s_t a + \epsilon_t)L. \tag{1}$$

The multiplicative specification relies on Gabaix and Landier (2008), who show that the dollar effect of CEO “talent” on firm value is increasing in firm value, and that the data is consistent with constant returns to scale. A notable implication is that an exogenous shock to business conditions (L) also affects the value to the firm of CEO ability.⁹

We assume that $\tilde{\epsilon}_t$ is normally distributed with mean zero and variance σ_ϵ^2 , and is independent from other random variables; \tilde{L} is a random variable with positive support and c.d.f. $G(\cdot)$, which is normalized so that $E[\tilde{L}] = 1$. Let \tilde{L} denote the random variable \tilde{L} , and L its realization at the end of period 1. Note that L affects firm profits in periods 1 and 2. We assume that L is observable and contractible. We refer to L as “luck”, since it represents a shock that is not under the control of the CEO but that nevertheless has an effect on firm value.

The variable s_t represents the accumulated experience and firm-specific skills of the CEO. Following Murphy and Zabojnik (2004), we let $s_t = 1$ if the CEO worked for the firm in period $t - 1$, and $s_t = \gamma \in (0, 1)$ otherwise. This means that $s_1 = \gamma$, and $s_2 = 1$ if the CEO stays in the same firm, and $s_2 = \gamma$ in case of CEO turnover. The evidence in Taylor (2013) is consistent with the assumption that managerial skills are not fully transferable across firms, that is, $\gamma < 1$. If general skills predominate, then s approaches one: managerial skills are easily transferable, and CEOs are easier to replace.

The ability \tilde{a} of a CEO is normally distributed with mean $\bar{a} > 0$ and variance σ_a^2 . It is initially unknown to the firm and to the CEO. CEOs are risk averse with utility function $u(\cdot)$, with $u' > 0$ and $u'' < 0$. We assume a limited supply of CEOs. Firms without CEOs can be run by managers, whose ability is normalized to zero.¹⁰ There is an infinite supply of such managers.

Firm net profits are the gross profits (henceforth “profits”) net of compensation costs. Both gross profits and net profits are observable and contractible. For simplicity, we assume a zero interest rate and no time discounting. Firms pay out their net profits realized over both periods to shareholders at the end of the second period.

⁹The “luck” shocks considered in Bertrand and Mullainathan (2001), are observable shocks “that a CEO does not influence” through his actions, for example mean industry performance. Our specification is consistent with this definition.

¹⁰The ability of managers could also be random, but this would not affect the main results. This formulation simplifies the algebra. It also implies that managers do not accumulate firm-specific skills.

We assume that a firm can commit to a long-term contract, but a CEO cannot. While firms can propose enforceable long-term contracts to their employees, constraints on involuntary servitude prevent employees from forgoing the option to quit a job. This one-sided-commitment assumption was introduced first in Harris and Holmstrom (1982). We also assume that CEOs can neither save nor borrow; they cannot transfer income from one period to another.¹¹ Two types of contracts are feasible: spot contracts and long-term contracts. In case of a spot contract, the employment of the first period CEO terminates at the end of the first period. A long-term contract specifies the wage that the firm commits itself to pay the CEO in both periods.

1.2 Monitoring

As in Burkart, Gromb, and Panunzi (1997), a fraction $\delta \in (0, 1]$ of shares is held by a single shareholder, while a fraction $1 - \delta$ is dispersed among a continuum of small shareholders.¹² Shareholders are risk neutral, and their objective is to maximize expected profits. Each shareholder can exert a nonverifiable monitoring effort $e \in [0, 1]$ at a cost $c\frac{e^2}{2}$ during the first period. As in Hermalin and Weisbach (1998), monitoring effort increases the probability that the shareholder receives a private signal on CEO ability. Specifically, with probability e , the shareholder receives the following signal:

$$y = \alpha + a + \epsilon_t \tag{2}$$

and with probability $1 - e$, he receives no signal. The signal can be shared with other shareholders who then decide whether or not to dismiss the CEO before the end of the first period.¹³ The cost of dismissal is denoted by $K \geq 0$; the parameter K represents the cost of involuntary CEO turnover. After dismissing its CEO, a firm hires a new CEO (or manager) on the spot market at the beginning of period 2.

The timeline of the model is depicted in Figure 1. The analysis proceeds in two steps. In section 2, we solve for optimal contracts for a given match of firms and

¹¹As in Harris and Holmstrom (1982) and Holmstrom and Ricart i Costa (1986), the optimal contract is such that the saving restriction is inconsequential.

¹²In the words of Bertrand and Mullainathan (2001), “Shleifer and Vishny (1986), among others, argue that large shareholders improve governance in a firm. A single investor who holds a large block of shares in a firm will have greater incentives to watch over the firm than a dispersed group of small shareholders.”

¹³These assumptions parsimoniously capture the notion that monitoring facilitates the assessment of CEO ability and CEO dismissal – before publicly observable measures of performances are realized. They also allow to avoid asymmetric information on the market for CEOs at the beginning of the second period, which is an interesting but separate issue that would markedly complicate the model without qualitatively affecting the main results.

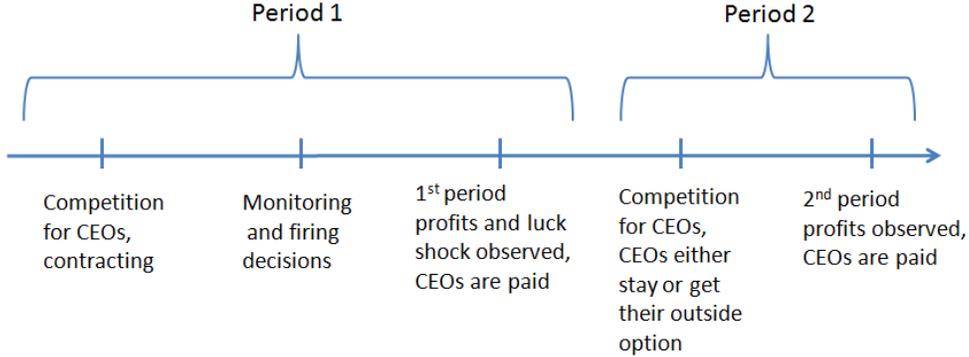


Figure 1: Timeline of the model.

CEOs. In section 3, we derive the equilibrium matching of firms and CEOs with the reservation utilities at the beginning of the first period determined in equilibrium.

2 Optimal contracts

Consider a firm with a shareholder owning a fraction δ of the shares and a CEO with variance of ability σ_a^2 and reservation utility over both periods denoted by \bar{U} . We first derive the CEO's outside option in the second period, after beliefs about his ability have been updated, and then solve backwards for the optimal contract in the first period.

2.1 Spot market in period 2

After observing first-period profits π_1 and the luck shock L , firms use Bayes' rule to update their belief about the CEO's expected ability \hat{a} :

$$\hat{a} = \frac{1}{\gamma} \frac{\gamma^2 \sigma_a^2 (\pi_1/L - \alpha) + \sigma_\epsilon^2 \gamma \bar{a}}{\gamma^2 \sigma_a^2 + \sigma_\epsilon^2}. \quad (3)$$

In the second period, a firm can hire a manager with zero ability on a spot contract for a zero wage, in which case its expected profits are αL . Firms can also compete for a CEO with updated expected ability \hat{a} . All other firms with vacant positions are willing to pay up to $\gamma \hat{a} L$ to hire this CEO, which corresponds to the additional expected profits generated by a given CEO relative to a zero-ability manager. Competition between firms drives the second period reservation wage to $W_2(\hat{a}, L) = \gamma \hat{a} L$. Because of one-sided commitment, a CEO with expected ability \hat{a} can earn this wage in the second

period, whether he entered a spot contract or a long-term contract in the first period. It follows that a firm that employed a CEO in the first period needs to match this reservation wage to retain him. Because of the imperfect transferability of managerial skills, the market value of CEOs does not fully adjust to ability or luck shocks: any given firm would be willing to pay up to $\hat{a}L$ to retain its CEO in the second period.

2.2 Monitoring and dismissals

Consider a firm that hired a CEO under a long-term contract. The firm optimally dismisses its CEO when it receives a signal on his ability *and* this signal is lower than a threshold.

First, consider the case in which shareholders do not receive a signal. The expected second period profits of a given firm that does not dismiss its CEO are:

$$(\alpha + E[\tilde{a}])E[\tilde{L}] - w_2, \quad (4)$$

where w_2 denotes the (as yet undetermined) compensation promised to the CEO in the second period. The expected second period profits of the firm if it dismisses its CEO and hires either a new CEO or a manager on the spot market at the beginning of the second period are:

$$\alpha E[\tilde{L}] - K - w_2, \quad (5)$$

where w_2 is again the contractual second period compensation of the initial CEO. Given that $E[\tilde{a}] = \bar{a}$, $E[\tilde{L}] = 1$, comparing the expressions in (4) and (5) yields the optimal firing rule: a firm will dismiss its CEO before the end of the first period if and only if

$$\bar{a} < -K. \quad (6)$$

This condition is never satisfied given that $\bar{a} > 0$.

Second, consider a firm that receives a signal on CEO ability in the first period. As above, its expected second period profits if it does not dismiss its CEO are:

$$(\alpha + E[\tilde{a}|y])E[\tilde{L}|y] - w_2. \quad (7)$$

Likewise, the expected second period profits of a firm which dismisses its CEO are:

$$\alpha E[\tilde{L}|y] - K - w_2. \quad (8)$$

Given that L and y are independent and $E[\tilde{L}] = 1$, and comparing the expressions in

(7) and (8) yields the optimal firing rule: a firm will dismiss its CEO before the end of the first period if and only if

$$E[\tilde{a}|y] < -K. \quad (9)$$

In summary:

Lemma 1. A firm that does not receive a signal on its CEO's ability does not dismiss its CEO. A firm that receives a signal y dismisses its CEO if and only if $E[\tilde{a}|y] < -K$.

Dismissing a CEO under a long-term contract is optimal if the updated expected ability of the CEO in place is lower than a threshold. This threshold is decreasing in the cost of dismissal K . Given that a firm that receives no signal retains its CEO, some CEOs with low ability remain in place. Monitoring effort increases the probability to receive a signal, so that firms that monitor more tend to dismiss their CEOs more often.

Shareholders use a cost-benefit analysis to choose the monitoring intensity. Due to free riding by small shareholders,¹⁴ only the large shareholder incurs the cost of monitoring. Since he is risk neutral, the benefit of monitoring is increasing linearly in δ_i , while the cost of monitoring is independent of δ_i .

Denoting by φ the p.d.f. of \tilde{a} , and by ϕ the p.d.f. of the signal \tilde{y} , the problem of the shareholder is

$$\max_e \delta \left[e \int_{-\infty}^{\infty} \left(\int_{-\infty}^{-K} -K \phi(y|x) dy + \int_{-K}^{\infty} x \phi(y|x) dy \right) \varphi(x) dx + (1-e) E[\tilde{a}] \right] - c \frac{e^2}{2} \quad (10)$$

The expected benefit of successful monitoring is

$$M(K, \sigma_a^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{-K} (-K - x) \phi(y|x) \varphi(x) dy dx = \int_{-\infty}^{-K} (-K - x) \hat{\varphi}(x) dx.$$

where $\hat{\varphi}$ denotes the density of the distribution of the updated CEO ability before y is observed. Simple algebra shows that this is the density of a normal distribution with mean \bar{a} and variance $\frac{\sigma_a^2}{\sigma_a^2 + \sigma_\varepsilon^2} \sigma_a^2$. Thus, the problem in (10) reduces to:

$$\max_e \delta_i e M(K, \sigma_a^2) - c \frac{e^2}{2}$$

¹⁴Note that the nonverifiability of the monitoring effort prevents shareholders from sharing the cost of monitoring. With a continuum $1 - \delta_i$ of other shareholders, this could alternatively be microfounded by assuming an arbitrarily small transaction cost.

To ensure an interior solution, we assume that $c > M(K, \sigma_a^2)$. Given that the problem is globally concave, the first-order condition is necessary and sufficient.

Lemma 2. The monitoring effort is $e^ = \delta_i M(K, \sigma_a^2)/c$. It is strictly increasing in the stake δ_i of the shareholder and in the variance of CEO ability, σ_a^2 .*

2.3 The optimal long-term contract

A long-term contract consists in a first-period wage w_1 and a second period wage $w_2(\hat{a}, L)$ contingent on the observed variables L and \hat{a} (through π_1). The optimal contract minimizes total expected compensation subject to two types of participation constraints. The state-contingent participation constraints guarantee that the second period contractual wage is at least as high as the reservation wage of the CEO. The first-period participation constraint guarantees that the expected utility associated with the two-period contract is as high as the CEO's reservation utility over two periods, \bar{U} .

As seen before, given that it only depends on the information on CEO ability received at the end of the first period, the dismissal decision is independent of the contract. The dismissal cost in turn does not affect the optimal long-term contract.¹⁵ In the remainder of the paper, we assume that the expected dismissal cost is not too large, so that the optimal long-term contract dominates a sequence of spot contracts.¹⁶

The optimal long-term contract solves:

$$\min_{\langle w_1, w_2(\hat{a}, L) \rangle} w_1 + E[w_2(\hat{a}, L)] \quad (11)$$

$$\text{subject to } w_2(\hat{a}, L) \geq W_2(\hat{a}, L) = \gamma \hat{a} L \quad \text{for all } \hat{a}, L \quad (12)$$

$$u(w_1) + E[u(w_2(\hat{a}, L))] \geq \bar{U}. \quad (13)$$

Lemma 3. The optimal long-term contract is characterized by a first period wage of w_1^ and a second period wage of:*

$$w_2^*(\hat{a}, L) = \max\{w_1^*, \gamma \hat{a} L\}. \quad (14)$$

The value of w_1^ is determined by the first period participation constraint, and depends*

¹⁵A firm dismisses its CEO if and only if $\hat{a} < -K$. The participation constraint (12) is binding only when $\hat{a} \geq \frac{w_1^*}{\gamma L} \geq 0$. But in cases when the firm dismisses its CEO, \hat{a} is negative so that (12) cannot be binding. It follows that the dismissal decision does not affect the contract.

¹⁶This assumption can be microfounded by assuming that σ_a^2 and K are sufficiently low.

on the value of the outside option \bar{U} .

A long term contract is fully characterized by the first-period wage w_1^* . The second period wage is either equal to w_1^* or adjusts to match the reservation wage $W_2(\hat{a}, L)$, when $W_2(\hat{a}, L)$ is larger than w_1^* . As in Harris and Holmstrom (1982), the risk averse CEO gets partial insurance: should his second period reservation wage fall below w_1^* , the firm nevertheless pays w_1^* in the second period. In this case, this payment either takes the form of a fixed wage, or of a severance payment, in case the CEO is dismissed. Thus, because of the CEO's risk aversion and of the ability of the firm to commit, it is inefficient to punish the CEO for "failure". To summarize, the optimal contract features downside protection for insurance purposes, and upside participation for retention purposes.

Even though firing a CEO is costly, there is no rent extraction in equilibrium. Indeed, the first period wage w_1^* adjusts so that the CEO is at his reservation level of utility. Intuitively, the CEO pays an insurance premium in the first period, which brings w_1^* to a lower level than what the CEO would get on the spot market. Our models thus differs on that dimension from Kuhnen and Zwiebel (2009), where firing costs make rent extraction possible. In addition, the CEO is less exposed to risk with this long-term contract than he would be with a sequence of spot contract because of the embedded insurance. Here our results differ from Oyer's (2004). In Oyer, there is a trade-off between exposing the CEO to risk (by indexing his pay on some variable which is imperfectly correlated with his reservation wage) and incurring renegotiation or transactions costs with interim re-contracting. We also differ from standard models of moral hazard, where the optimal contract trades-off incentives for effort and risk-sharing.

The optimal contract can be implemented in two different ways – this does not matter for our subsequent results.¹⁷ First, the firm can simply commit to paying the CEO the fixed wage w_1^* in the first and second periods, and adjust upward CEO pay ex-post (at the beginning of the second period) in cases when $W_2(\hat{a}, L) > w_1^*$, as specified in (14). Indeed, this will be ex-post optimal for retention purposes. Second, at the beginning of the first period, the firm can offer the CEO an explicit long-term contract based on indexed firm value, as described in the next paragraph.

¹⁷In particular, the results in Proposition 1 hold in any case, as the factors at play in (18) are by construction the same as in (14). Note that the explicit contract described in the following paragraphs would strictly dominate the ex-post adjustment described in this paragraph with a renegotiation cost, no matter how small. By contrast, in Oyer (2004), with an arbitrarily small renegotiation cost, it would be optimal to use spot contracts and renegotiation instead of explicit long-term contracts.

We now show that the optimal second-period compensation of the CEO can be expressed as a function of firm value and the luck shock. In a competitive market for firm shares with risk neutral shareholders, firm value at the beginning of the second period is:

$$V = \pi_1 - w_1 + E[\pi_2|\pi_1, L] - w_2(\hat{a}, L) = \pi_1 - w_1^* + (\alpha + \hat{a})L - \max\{w_1^*, \gamma\hat{a}L\}, \quad (15)$$

if a signal y is received by the shareholder and $E[\tilde{a}|y] > -K$. Otherwise, the initial CEO is dismissed at the beginning of the second period, with a compensating payment of w_1^* , while the new CEO receives a fixed wage, so that firm value does not matter for compensation purposes. Substituting for $\pi_1 = V + w_1^* - (\alpha + \hat{a})L + \max\{w_1^*, \gamma\hat{a}L\}$ in (3) and isolating \hat{a} , we get:

$$\begin{aligned} w_2^*(\hat{a}, L) &= w_1^* \text{ if } \gamma\hat{a}L \leq w_1^*, \\ w_2^*(\hat{a}, L) &= \gamma\hat{a}L = \left(\frac{\gamma^2\sigma_a^2 + \sigma_\epsilon^2}{\gamma^2\sigma_a^2} + \frac{1-\gamma}{\gamma} \right)^{-1} \left(V + w_1^* - 2\alpha L + \gamma\bar{a}L \frac{\sigma_\epsilon^2}{\gamma^2\sigma_a^2} \right) \text{ otherwise.} \end{aligned}$$

That is,

$$w_2^*(\hat{a}, L) = \max\{w_1^*, \psi(w_1^* + V + \eta L)\}, \quad (16)$$

$$\text{where } \psi \equiv \left(\frac{\gamma^2\sigma_a^2 + \sigma_\epsilon^2}{\gamma^2\sigma_a^2} + \frac{1-\gamma}{\gamma} \right)^{-1} \quad \text{and} \quad \eta \equiv \gamma\bar{a} \frac{\sigma_\epsilon^2}{\gamma^2\sigma_a^2} - 2\alpha. \quad (17)$$

The optimal contract described in Lemma 3 can be implemented by making CEO pay in the second period contingent upon the measure $P(V, L) = V + \eta L$. Indeed, the state-contingent payment $w_2^*(\hat{a}, L)$ in (14) can be expressed as

$$w_2^*(\hat{a}, L) = \max\{w_1^*, \psi w_1^* + \psi P(V, L)\}. \quad (18)$$

We conclude that the optimal state-contingent payment may be implemented with a fixed wage of w_1^* in the first and second periods, and indexed stock-options based on P with exercise price $\kappa = \frac{w_1^*(1-\psi)}{\psi}$ which vest at the beginning of the second period.

We now summarize notable features of the optimal contract:

Proposition 1. The compensation contract has the following characteristics:

- (i) *The sensitivity of the performance measure to luck is increasing in the variance of the idiosyncratic shock $\tilde{\epsilon}$: $\frac{d\eta}{d\sigma_\epsilon^2} > 0$.*
- (ii) *The sensitivity of the performance measure to luck is decreasing in the variance of the CEO's ability: $\frac{d\eta}{d\sigma_a^2} < 0$.*

(iii) *The contract displays asymmetric pay-for-luck.*

(iv) *The sensitivity of pay to performance is increasing in the relative importance of general managerial skills: $\frac{d\psi}{d\gamma} > 0$.*

As in Oyer (2004), the compensation of the CEO in the second period depends on business conditions, or “luck” (L): $\frac{d}{dL}P(V, L) \neq 0$. Compensation adjusts to the level required to retain the CEO, and this level depends on business conditions. The degree of pay-for-luck relative to pay-for-performance (which is essentially pay-for-ability), as measured by η , is increasing in σ_ϵ^2 and decreasing in σ_a^2 : there is more pay-for-luck relative to pay-for-performance when firm value V is a noisy measure of CEO ability a , and when the initial uncertainty on the ability of the CEO is low. This suggests that pay-for-luck is stronger for old CEOs or CEOs with a long tenure (with low σ_a^2), and that young CEOs or CEOs with a short tenure (with high σ_a^2) should be less paid for luck. In the limit, as the ratio $\frac{\sigma_\epsilon^2}{\sigma_a^2}$ tends to infinity, η also tends to infinity, and state-contingent remuneration only depends on luck. In addition, the CEO is “rewarded for good luck”, but he is not symmetrically “penalized for bad luck”: pay-for-luck is asymmetric. This is due to the insurance provided to risk averse CEOs.

The sensitivity of CEO pay to performance, as measured by ψ , is increasing in the relative importance of general managerial skills (γ). An increase in γ means that CEOs are more transferable across firms. The sensitivity of CEO compensation to firm value and luck must then increase in order to match the reservation wages of CEOs state-by-state. To the extent that general skills became progressively more important in the 1980s and the 1990s, as argued by Murphy and Zabochnik (2004) and Frydman (2007), then the model explains why CEOs received increasing amounts of stock-options over this period (Frydman and Jenter (2010), figure 2). This prediction is also consistent with the evidence in Cunat and Guadalupe (2009) that stronger international competition – which in their words “could be an additional reason why general skills are more important” – is associated with more performance-related pay. Finally, this prediction may explain the finding in Murphy (2012) that non-U.S. CEOs do not receive as much equity-based pay as U.S. CEOs. Indeed, skills might be relatively less transferable in countries which are comparatively small and not Anglophone.

3 Competition for CEOs and matching equilibrium

In section 2, we derived the optimal contract with exogenous reservation utilities \bar{U} . We now introduce competition between firms to endogenize the CEOs’ first-period

reservation utilities. We also extend the model to incorporate matching between CEOs and firms, in order to study the link between monitoring and CEO pay practices. We do not model the competitive process explicitly, but we identify the stable matching between firms and CEOs.

3.1 A matching model of CEOs and firms

We now assume that firms differ in their monitoring capacity, because of differences in ownership structures. A firm i with a large outside shareholder (high δ_i) has more incentives to monitor the CEO to get information about his ability and dismiss him if necessary. We also assume that the variance of ex-ante ability σ_a^2 differs across CEOs.

As in the baseline model, there are no information asymmetries: for any given CEO, the value of σ_a^2 is common knowledge, but neither the firms nor the CEO observe a . We denote by $A_f = \{\delta_1, \delta_2, \dots, \delta_n\}$ the set of firms with $\delta_1 > \delta_2 > \dots > \delta_n$. Likewise, we denote by $A_c = \{\sigma_1^2, \sigma_2^2, \dots, \sigma_l^2, \sigma_{l+1}^2, \dots, \sigma_p^2\}$ the set of l CEOs, with $E[\tilde{a}] = \bar{a}$ and $\sigma_1^2 > \sigma_2^2 > \dots > \sigma_l^2$, and $p - l$ managers, with $a = 0$ and $\sigma_k^2 = 0$ for $k \geq l + 1$, where $l < n < p - l$. We thus assume that CEOs are on the short side of the market.

Following Roth and Sotomayor (1989), the matching process can be defined as a matching function $\mu : A_f \cup A_c \rightarrow A_f \cup A_c$ such that $\mu(\delta_i) \in A_c \cup \{\delta_i\}$ for all δ_i , $\mu(\sigma_i^2) \in A_f \cup \{\sigma_i^2\}$, for all $\sigma_i^2 \in A_c$, and $\mu(\delta_i) = \sigma_j^2$ if and only if $\mu(\sigma_j^2) = \delta_i$ for all $(\delta_i, \sigma_j^2) \in A_f \times A_c$. An equilibrium is defined by a matching function that specifies which type of firm employs which type of CEO/manager, and the associated contracts. A firm is unmatched if $\mu(\delta_i) = \delta_i$. Similarly, a CEO or manager is unmatched if $\mu(\sigma_i^2) = \sigma_i^2$.

The first condition for equilibrium is that the matching function is consistent; each manager or CEO is matched with one firm at most. The second condition is that no firm can break its match and increase its expected profit by proposing a contract to an already matched manager or CEO who would prefer that contract. The model corresponds to a matching model with nontransferabilities, as studied in Legros and Newman (2007). They derive sufficient conditions on the Pareto frontiers generated by a match that ensure positive or negative assortative matching.

Proposition 2. Consider a matching μ associated with optimal contracts. This matching is stable if and only if the $n - l$ firms with the lowest δ 's are matched with managers, and there is positive assortative matching between the l CEOs and the l firms with the highest δ 's: $\mu(\delta_1) = \sigma_1^2, \mu(\delta_2) = \sigma_2^2, \dots, \mu(\delta_l) = \sigma_l^2$.

For any fixed set of reservation utilities, a match between risky CEOs and firms with better monitoring capacity generates more surplus. Indeed, it is more likely that the estimated ability of a more risky CEO (with a higher σ_a^2) will fall below the firing threshold of any given firm. To minimize the costs of inefficient continuation of CEOs with low ability, it is more efficient to match a risky CEO with a firm with a good monitoring capacity. The firm will exert more monitoring effort and thus will be in a better position to dismiss a CEO with low ability.

The reservation utilities of CEOs and the wages associated with their optimal contracts are determined in the matching equilibrium. The wage of CEO σ_l matched with firm δ_l is such that if firm δ_{l+1} were to attract this CEO, it would make zero expected profits, which is what it gets by hiring a manager. This condition pins down the wage that firm δ_l needs to give its CEO, and also determines the expected profits of that firm. The wage of CEO $l - 1$ and the expected profits of firm $l - 1$ are similarly determined by the condition that firm l cannot hire CEO $l - 1$ and make higher expected profits than it would by employing CEO $l - 1$ with the contract previously determined.

The existing evidence, while not conclusive, is consistent with the prediction that more risky CEOs are hired by firms with a greater monitoring capacity: Shen and Cannella (2002) and Agrawal, Knoeber, and Tsoulouhas (2006) find a statistically significant association between hiring an industry outsider as CEO and the proportion of independent directors.

3.2 CEO salaries and monitoring

The model has implications for the distribution of CEO salaries across firms with different monitoring capacities. Since a firm matches the outside option of the CEO in the second period, the fixed wage w_1 fully characterizes the contract. Denoting by w_1^i the first-period wage of CEO i , the following Proposition characterizes the equilibrium distribution of salaries across firms.

Proposition 3. The CEO wage w_1 is higher in a firm with relatively worse monitoring if and only if the cost of dismissal K or the transferability of skills γ are sufficiently large:

- (i) For any $K \in (0, \infty)$, there exists $\underline{\gamma} \in (0, 1)$ with $w_1^{i+1} > w_1^i$ for $\gamma > \underline{\gamma}$, and $w_1^{i+1} < w_1^i$ for $\gamma < \underline{\gamma}$;
- (ii) For any $\gamma \in (0, 1)$, there exists $\underline{K} \in (0, \infty)$ with $w_1^{i+1} > w_1^i$ for $K > \underline{K}$, and $w_1^{i+1} < w_1^i$ for $K < \underline{K}$.

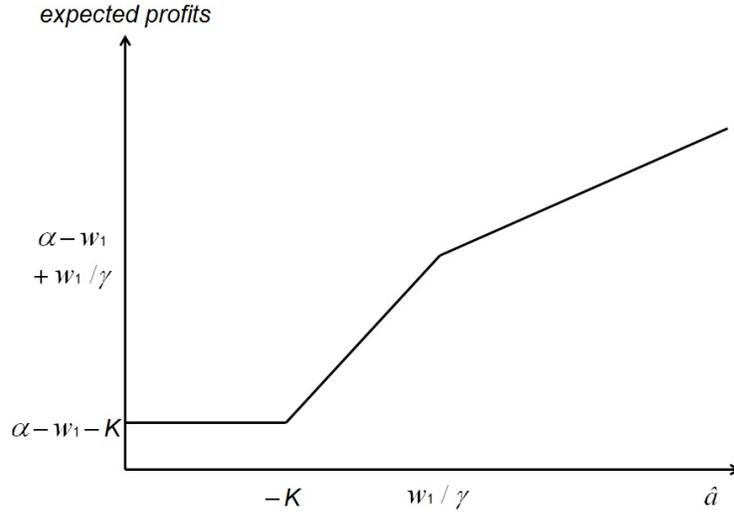


Figure 2: Expected second period profits as a function of the CEO's expected ability.

There are two forces at play. First, the fact that dismissal is costly ($K > 0$) implies that firms might be stuck with a CEO with a low estimated ability \hat{a} . Intuitively, costly dismissal ($K > 0$) introduces some concavity in the second period expected firm profits as a function of estimated ability \hat{a} , as depicted in Figure 2 (with $K = \infty$ and $\gamma < 1$, this function would be concave). Second, CEOs capture only a fraction γ of the gains from their estimated ability at the beginning of the second period, because of the imperfect transferability of managerial skills across firms ($\gamma < 1$). The higher γ , the less benefits firms reap from having a CEO with superior ability. Intuitively, the fact that the CEO captures some of the gains from his skills ($\gamma > 0$) also introduces some concavity in the aforementioned function depicted in Figure 2 (with $\gamma = 0$, this function would be convex). The slope of this function is 1 on $(-K, w_1/\gamma)$, and $1 - \gamma$ on $(w_1/\gamma, \infty)$.

In sum, an increase in either the dismissal cost (K) or the transferability of managerial skills (γ) increases the relative value of a CEO whose variance of ability is relatively low. In equilibrium, this type of CEO will be paid a higher salary if K and/or γ is sufficiently high. But this type of CEO is matched with a firm with poor monitoring capacity (low δ). Consequently, for sufficiently high values of K and/or γ , the CEO salary will be higher in firms with poor monitoring, which can be interpreted as bad governance.

3.3 Pay-for-luck and monitoring

The matching equilibrium and the associated contracts can explain the Bertrand and Mullainathan (2001) finding that firms without a large shareholder use contracts that display more pay for luck. This empirical fact is at the root of the managerial power theory, which argues that CEOs “set their own pay” in this type of firm (Bertrand and Mullainathan (2001)), and it has not been explained by any model of efficient contracting.

The link between pay-for-luck and poor monitoring capacity follows from endogenous matching and not (directly) from differences in monitoring capacity. For example, two firms with different δ 's but with identical CEOs would offer identical compensation packages. However, in the stable matching equilibrium, these two different firms are matched with CEOs of different types that require different contracts. As a consequence, the observed differences in CEO pay across firms with different monitoring capacities are explained by differences in CEO characteristics. Specifically, in a matching equilibrium, firms with worse monitoring (that is, with less concentrated ownership) are matched with CEOs with a lower σ_a^2 (“safe CEOs”). But Proposition 1 states that pay-for-luck is decreasing in σ_a^2 . Therefore, in equilibrium, there is more pay-for-luck in firms with worse monitoring. The following Proposition summarizes this point:¹⁸

Proposition 4. (i) The second period reservation wage is relatively more sensitive to luck than to performance for CEOs whose ability is more precisely estimated ex-ante:

$$\frac{dw_2^*(\hat{a}, L)}{dL} / \frac{dw_2^*(\hat{a}, L)}{dV} \quad \text{is (weakly) decreasing in } \sigma_a^2. \quad (19)$$

(ii) Firms with less concentrated ownership (low δ_i) offer contracts that display more pay-for-luck than firms with more concentrated ownership (high δ_i).

The intuition behind the result is well illustrated by the following extreme example. Suppose that the ability of a given CEO i is known: $\sigma_i^2 = 0$. Since $\hat{a} = \bar{a}$ with probability one, the second period reservation wage of this CEO depends only on future business conditions, or “luck” (L). Firm performance net of luck is not informative about the CEO’s ability. On the contrary, for of a CEO j with uncertain ability ($\sigma_j^2 > 0$), the second period reservation wage depends both on future business conditions and on the updated expected ability of the CEO. In the case of CEO i , the variability in

¹⁸The result is not purely driven by our modeling of monitoring. In a previous version of the paper, this result was obtained in a model where the cost of CEO dismissal varies across firms.

the second period pay of the CEO is fully attributable to luck. In the case of CEO j , it is attributable both to luck and to the updated CEO's ability. In a matching equilibrium, CEO i is hired by the firm with the worst monitoring ability (lowest δ), and he is exclusively paid for luck, whereas CEO j is paid both for luck and for performance.

3.4 Pay-for-performance and monitoring

Another major prediction of the managerial power theory is that the sensitivity of pay to performance is lower in firms with worse governance, and that the CEOs of these firms are punished less for bad performances. There is some empirical support for these predictions. For example, Hartzell and Starks (2003) find that the concentration of institutional investors, which may “result in greater monitoring and scrutiny of the CEO” (in the words of Bebchuk and Fried (2003)), is positively related to the pay-performance sensitivity (PPS) of CEO compensation. Bell and Van Reenen (2013) also find that the relation between pay and performance is more asymmetric in firms with a smaller base of institutional investors,¹⁹ in the sense that these firms' CEOs are even less punished for negative performances than they are rewarded for positive performances.

The model can explain this asymmetry, as well as the association between monitoring ability and pay-for-performance. Indeed, under certain conditions, a larger ownership stake δ_i is associated with less protection against negative outcomes and a higher PPS. Denoting by $PPS_{P,i}$ the PPS in firm i at the level of performance P , we summarize these results in the following Proposition:

Proposition 5. If the cost of dismissal K and the transferability of managerial skills γ are sufficiently large, then

(i) For any two given levels of performances \bar{P} and \underline{P} , with $\bar{P} > \underline{P}$, the ratio $\frac{PPS_{\bar{P},i}}{PPS_{\underline{P},i}}$ (if well-defined) is higher in firms with a larger δ_i ;

(ii) At any given level of the performance measure P , the sensitivity of pay to performance is higher in firms with larger δ_i .

The intuition behind the first result is the following: to the extent that the fixed wage w_1^* is higher in firms with worse monitoring (cf. Proposition 3), CEO pay is

¹⁹Aghion, Van Reenen, and Zingales (2013) argue that institutional investors are better at monitoring CEOs, while Hartzell and Starks (2003) argue that these institutions “serve a monitoring role”, and reference a number of other studies that support this conclusion.

sensitive to the outside option for a smaller range of poor performances. That is, CEOs in these firms are more protected against the negative consequences of a low first-period performance. The intuition for the second result is twofold. First, the first period performance of the firm is more informative for a CEO whose ability is less precisely estimated ex-ante (high σ_i^2). The outside option of such a CEO in the second period is therefore more sensitive to his first period performance, so that the type of firm he is matched with (a firm with a high δ_i) must increase the PPS on the upside (ψ) for retention purposes. Second, with a higher ψ and/or a lower w_1^* , the outside option of the CEO in the second period exceeds w_1^* for a larger range of performances. CEO pay must then be sensitive to performance for a larger range of performances.

In summary, the pay of CEOs at firms which better monitoring is sensitive to performance for a larger set of performances, and it is more sensitive to performance on that set. The model also predicts that the PPS is positively related to the variability of firm profits (which is increasing in σ_i^2), as in Inderst and Mueller (2010), but contrary to a standard model of incentive pay (e.g., Holmstrom and Milgrom (1987)).

3.5 Corporate governance spillovers

The matching model generates corporate governance spillovers, whereby an improvement in the monitoring capacity of a subset of firms has spillover effects on CEO compensation in firms with better monitoring.

Proposition 6. If δ_i increases for $i \in \{j, \dots, n\}$, with $j \leq l$, without changing the ranking of firms on that dimension, then CEO compensation increases in the set of firms with ownership structure $\{\delta_1, \dots, \delta_l\}$.

In equilibrium, firm expected profits are constrained by the competition for CEOs. The difference in expected profits between any given firm and the next firm with worse monitoring is increasing in the wedge between the monitoring intensities of the two firms. An improvement in the monitoring capacity of the worst firms reduces this wedge and therefore reduces expected profits in this subset of firms (except for firms that employ managers). This leads to an increase in the compensation of their CEOs. These firms, in turn, are willing to pay more to hire CEOs employed in firms with better monitoring. In equilibrium, CEO pay must therefore increase in all firms with better monitoring. In particular, an improvement in the monitoring capacity of the worst firms triggers an across-the-board increase in CEO pay. This result can explain that CEO pay rose as monitoring capacity improved in the U.S. (Huson, Parrino, and

Starks (2001)), even if this improvement only concerned firms with an initially low monitoring capacity.

By contrast, a firm-specific change in δ_i does not affect CEO pay in the firm in question. This is because second-period compensation w_2^* is not affected by δ_i , while w_1^* is only affected by δ_i to the extent that the reservation utility \bar{U} of the CEO depends on δ_i (the participation constraint is binding in equilibrium). But the reservation utility of the CEO is unrelated to firm-specific factors; it does not depend on the monitoring intensity of the firm. Interestingly, a model of managerial power would tend to make a different prediction. For example, in the model of Kuhnen and Zwiebel (2009), CEOs “receive further compensation for their entrenchment.” Given that changes in governance are often either correlated with other factors or not firm-specific, the related empirical evidence is scarce. This said, Cunat, Gine, and Guadalupe (2012) identify a firm-specific exogenous shock to governance, and find that the effect on the level of CEO pay is statistically insignificant – even though the effects on other firm-specific variables are large. This is in line with the prediction of our model that CEO pay is unrelated to *firm-specific* corporate governance.

4 Other predictions and empirical implications

In this section, we confront other predictions of the model to further empirical evidence.

Even though we do not explicitly derive predictions on this dimension, there are reasons to believe that the uncertainty about CEO ability σ_a^2 decreases over the tenure of a CEO, as more signals about his ability become available. An implicit prediction of our model is therefore that pay-for-luck should increase over CEO tenure. Bertrand and Mullainathan (2001) derive a similar prediction with the managerial power theory: insofar as CEOs with a longer tenure are more entrenched, they can extract more benefits in the form of asymmetric pay-for-luck.²⁰ This prediction, which is empirically validated, is thus common to the managerial power theory and the efficient contracting model.

The model does not generate any cross-sectional prediction regarding the frequency of firm-CEO separations depending on a measure of corporate governance. Firms with a more limited monitoring capacity hire CEOs with a more precisely estimated ability,

²⁰Garvey and Milbourn (2003) also find that relative performance evaluation, which consists in filtering out one type of exogenous shock, namely the market index, is stronger for younger CEOs. They interpret this finding as evidence that firms tend to let older (and more wealthy) CEOs hedge against market fluctuations themselves, since they are better able to do so than young CEOs.

which tends to reduce bad surprises and the associated dismissal. But CEOs with a more precisely estimated ability may be older, and therefore closer to retirement, which tends to increase voluntary turnover. A priori, it is not clear which effect dominates. This is in line with Kaplan and Minton (2012), who find no statistically significant relation between CEO turnover and corporate governance in the cross-section. On the contrary, with an across-the-board change in governance, which leaves the ranking of firms unchanged, there is no sorting effect. Therefore, only the direct effect is present, and the model predicts that an across-the-board improvement in monitoring capacity leads to more forced CEO turnover. This is again consistent with the evidence in Kaplan and Minton (2012), and with the predictions in Hermalin (2005).

According to Frydman (2007) and Frydman and Saks (2010), CEO pay was stable and pay dispersion across executives was low from the 1930s to the 1970s, but not in the following decades. The predictions of the model are consistent with these patterns. Indeed, the evidence in Frydman (2007) suggests that firm-specific skills predominantly mattered in the past, while general skills have become more important in the last decades, i.e., γ increased. In our model, with a low γ , CEO skills are not easily transferable, and CEO pay is largely unresponsive to CEO performance. CEO pay is therefore relatively stable over time. If in addition there is little dispersion in the *expected* ability of CEOs ex-ante (before the performances are realized), then the dispersion of CEO pay across firms is low – even if the *actual* dispersion of abilities across CEOs is large and is at least partly learned ex-interim. This is because the first period wage w_1^* depends on the expected ability \bar{a} , while the second period wage deviates from this level only to the extent that γ is substantial. On the contrary, the model predicts that a rise in γ should coincide with a higher PPS and a rise in ex-post pay dispersion among CEOs, consistent with what happened in the 1980s and 1990s.

It is noteworthy that Murphy and Zbojnik (2004) and Frydman (2007) relate the increase in the relative importance of general skills to the level of CEO pay and CEO turnover, whereas our model also relates it to the structure of CEO compensation and the PPS (which is increasing in γ). Relatedly, Cremers and Grinstein (2013) find that the level of CEO compensation is only slightly higher in industries with less firm-specific skills, i.e., where CEOs skills are more transferable. In our model, an increase in γ only increases expected CEO compensation to the extent that the utility function u is concave, ceteris paribus. This is to compensate the CEO for an increase in the variability of his second period pay, and for a greater difference between expected pay in the first and second period. If the risk aversion of CEOs is low (Graham, Harvey, and Puri (2013)), these effects will be positive but economically weak, consistent with

the evidence in Cremers and Grinstein (2013).

5 Conclusion

This paper develops a model of efficient contracting that addresses some fundamental arguments of the managerial power theory. The main contribution is to study the consequences of endogenous matching between firms with different capacities for monitoring and CEOs with a more or less uncertain ability. The model can explain that CEOs in firms with more diffuse ownership and worse monitoring (which can be interpreted as bad governance) receive higher salaries, are more paid for luck, and get a compensation that is less sensitive to performance, especially on the downside because of a greater asymmetry in the pay-for-performance relation. We also find that an improvement in the monitoring capacity of the worst firms has spillover effects that increase CEO pay in all firms, but that a firm-specific change in monitoring capacity does not affect CEO pay. These results contribute to a large recent literature which shows that the efficient contracting paradigm can actually explain a number of apparent anomalies (see Edmans and Gabaix (2009)). In particular, the ability of the model to explain many important stylized facts suggests that the retention motive is an important determinant of the structure (not just the level) of CEO compensation.

The results have obvious policy implications. If the “anomalies” related to CEO pay imply a corporate governance failure – that would affect a majority of large American firms and have wide ranging consequences, of which the observed anomalies would merely be a symptom – then it provides the grounds for a regulatory intervention, and a substantial change in firm governance. However, if this evidence can be comprehensively explained by a simple model of efficient contracting, then it is unclear whether such measures are necessary.

6 References

- Acharya, V., Volpin, P.F., 2010, Corporate governance externalities. *Review of Finance*, 14, 1-33.
- Akerberg, D.A., Botticini, M., 2002, Endogenous matching and the empirical determinants of contract form. *Journal of Political Economy*, 110 (3), 564-591.
- Aghion, P., Van Reenen, J., Zingales, L., 2013, Innovation and institutional ownership. *American Economic Review*, 103 (1), 277-304.
- Agrawal, A., Knoeber, C.R., Tsoulouhas, T., 2006, Are outsiders handicapped in CEO successions? *Journal of Corporate Finance*, 12 (3), 619-644.
- Bebchuk, L.A., Fried, J.M., 2003, Executive compensation as an agency problem. *Journal of Economic Perspectives*, 17 (3), 71-92.
- Bell, B.D., Van Reenen, J., 2013, Extreme wage inequality: pay at the very top. *American Economic Review*, 103 (3), 153-157.
- Bertrand, M., Mullainathan, S., 2001, Are CEOs rewarded for luck? The ones without principals are. *Quarterly Journal of Economics*, 116 (3), 901-929.
- Beaudry, P., DiNardo, J., 1991, The effect of implicit contracts on the movement of wages over the business cycle: evidence from micro data. *Journal of Political Economy*, 99 (4), 665-688.
- Burkart, M., Gromb, D., Panunzi, F., 1997, Large shareholders, monitoring, and the value of the firm. *Quarterly Journal of Economics*, 112 (3), 693-728.
- Cao, M., Wang, R., 2013, Optimal CEO compensation with search: theory and empirical evidence. *Journal of Finance*, 68 (5), 2001-2058.
- Core, J.E., Holthausen, R.W., Larcker, D.F., 1999, Corporate governance, chief executive compensation, and firm performance. *Journal of Financial Economics*, 51, 371-406.
- Cornelli, F., Kominek, Z., Ljungqvist, A., 2013, Monitoring managers: does it matter? *Journal of Finance*, 68 (2), 431-481.
- Cremers, K.J.M., Grinstein, Y., 2013, Does the market for CEO talent explain controversial CEO pay practices? *Review of Finance*, Forthcoming.
- Cunat, V., Guadalupe, M., 2009, Globalization and the provision of incentives inside the firm: The effect of foreign competition. *Journal of Labor Economics*, 27 (2), 179-212.
- Cunat, V., Gine, M., Guadalupe, M., 2012, Say pays! Shareholder voice and firm performance. Upjohn Institute working paper 13-192.
- Dicks, D.L., 2012, Executive compensation and the role for corporate governance

regulation. *Review of Financial Studies*, 25 (6), 1971-2004.

Dittmann, I., Maug, E., 2007, Lower salaries and no options? On the optimal structure of executive pay. *Journal of Finance*, 62 (1), 303-343.

Edmans, A., Gabaix, X., 2009, Is CEO pay really inefficient? A survey of new optimal contracting theories. *European Financial Management*, 15 (3), 486-496.

Edmans, A., Gabaix, X., 2011, Tractability in incentive contracting. *Review of Financial Studies*, 24 (9), 2865-2894.

Eisfeldt, A., Kuhnen, C., 2013, CEO turnover in a competitive assignment framework. *Journal of Financial Economics*, 109 (2), 351-372.

Falato, A., Li, D., Milbourn, T., 2012, CEO pay and the market for CEOs. Working paper, Federal Reserve Board.

Frydman, C., 2007, Rising through the ranks: The evolution of the market for corporate executives, 1936-2003. Working paper, MIT.

Frydman, C., Jenter, D., 2010, CEO compensation. *Annual Review of Financial Economics*, 2, 75-102.

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

Gabaix X., Landier A., 2008, Why has CEO pay increased so much? *Quarterly Journal of Economics*, 123 (1), 49-100.

Garvey, G., Milbourn, T., 2003, Incentive compensation when executives can hedge the market: evidence of relative performance evaluation in the cross section. *Journal of Finance*, 58 (4), 1557-1582.

Garvey, G. and Milbourn, T., 2006, Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad. *Journal of Financial Economics*, 82 (1), 197-225.

Graham, J.R., Harvey, C.R., Puri, M., 2013, Managerial attitudes and corporate actions. *Journal of Financial Economics*, 109 (1), 103-121.

Hall, B.J., Murphy, K.J., 2003, The trouble with stock options. *Journal of Economic Perspectives*, 17 (3), 49-70

Harris M., Holmstrom B., 1982. A theory of wage dynamics. *Review of Economic Studies*, 49 (3), 315-333.

Hartzell, J.C., Starks, L.T., 2003, Institutional investors and executive compensation. *Journal of Finance*, 58 (6), 2351-2374.

Hermalin, B.E., 2005, Trends in corporate governance. *Journal of Finance*, 60 (5), 2351-2384.

Hermalin, B.E., Weisbach, M.S., 1998, Endogenously chosen boards of directors

and their monitoring of the CEO. *American Economic Review*, 88 (1), 96-118.

Holmstrom, B., Ricart I Costa, J., 1986, Managerial incentives and capital management. *Quarterly Journal of Economics*, 101 (4), 835-860.

Holmstrom, B., Milgrom P., 1987, Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, 55 (2), 303-328.

Huson, M.R., Parrino, R., Starks, L.T., 2001, Internal monitoring mechanisms and CEO turnover: A long-term perspective. *Journal of Finance*, 56 (6), 2265-2297.

Inderst, R., Mueller, H.M., 2010, CEO replacement under private information. *Review of Financial Studies*, 23 (8), 2935-2969.

Kaplan, S., Minton, B., 2012, How has CEO turnover changed? *International Review of Finance*, 12 (1), 57-87.

Kuhnen, C.M., Zwiebel, J., 2009, Executive pay, hidden compensation and managerial entrenchment. Working paper, Northwestern University.

Lazear, E.P., 2004, Output-based pay: incentives, retention or sorting? *Research in Labor Economics*, 23, 1-25.

Legros, P., Newman, A., 2007. Beauty is a beast and frog is a Prince: Assortative matching with nontransferabilities. *Econometrica*, 75 (4), 1073-1102.

Murphy, K.J., Zabojnik, J., 2004, CEO pay and appointments: a market-based explanation for recent trends. *American Economic Review*, 94 (2), 192-196.

Murphy, K.J., 2012, Executive compensation: where we are, and how we got there, in *Handbook of the Economics of Finance*, Elsevier Science North Holland.

Oyer, P., 2004. Why do firms use incentives that have no incentives effects? *Journal of 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. *Journal of Financial Economics*, 76, 99-133.

Rajgopal, S., Shevlin, T., Zamora, V., 2006, CEOs' outside employment opportunities and the lack of relative performance evaluation in compensation contracts. *Journal of Finance*, 61 (4), 1813-44.

Roth, A., Sotomayor, M., 1989, The college admission problem revisited. *Econometrica*, 57 (3), 559-70.

Shen, W., Cannella, A.A., 2002, Revisiting the performance consequences of CEO succession: The impacts of successor type, postsuccession senior executive turnover, and departing CEO tenure. *Academy of Management Journal*, 45 (4), 717-733.

Shleifer, A., Vishny, R., 1986, Large shareholders and corporate control. *Journal of Political Economy*, 94 (3), 461-488.

Taylor, L., 2013, CEO wage dynamics: estimates from a learning model. *Journal of Financial Economics*, 108 (1), 79-98.

Tervio, M., 2008, The difference that CEOs make: an assignment model approach. *American Economic Review*, 98 (3), 642-668.

7 Appendix

Proof of Lemma 2:

Recall the optimal effort of the shareholder: $M(K, \sigma_a^2) = \int_{-\infty}^{-K} (-K - x) \hat{\varphi}(x) dx$.

The problem is globally concave, so that the following first-order condition is necessary and sufficient

$$e^* = \delta_i M(K, \sigma_a^2) / c. \quad (20)$$

Since $M(K, \sigma_a^2) > 0$, it then follows that e^* is strictly increasing in δ_i . Finally, our assumption that $c > M(K, \sigma_a^2)$ guarantees that $e^* \in (0, 1)$.

To prove that e^* is increasing in σ_a^2 , we will show that $M(K, \sigma_a^2)$ is increasing in σ_a^2 . Let $\sigma_i^2 > \sigma_j^2$. Then we have to show that

$$\int_{-\infty}^{-K} (-K - x) \hat{\varphi}^i(x) dx > \int_{-\infty}^{-K} (-K - x) \hat{\varphi}^j(x) dx.$$

Integrating by parts, we get

$$\begin{aligned} & \int_{-\infty}^{-K} (-K - x) \hat{\varphi}^i(x) dx - \int_{-\infty}^{-K} (-K - x) \hat{\varphi}^j(x) dx \\ &= \int_{-\infty}^{-K} \left(\hat{\Phi}^i(x) - \hat{\Phi}^j(x) \right) dx. \end{aligned}$$

Given that an increase in variance of the normal distribution is a mean preserving spread, we get that $\hat{\Phi}^i(x) > \hat{\Phi}^j(x)$ for $x < \bar{a}$. Since the mean of both distributions is $\bar{a} > 0$, we have that $\hat{\Phi}^i(x) - \hat{\Phi}^j(x) > 0$ for $x < -K < \bar{a}$. The result follows. ■

Proof of Lemma 3:

The optimal long-term contract solves the following optimization problem:

$$\begin{aligned} & \min_{\langle w_1, w_2(\hat{a}, L) \rangle} w_1 + \int \int w_2(x, y) dF_{\hat{a}}(x) dG(y) \\ & w_2(\hat{a}, L) \geq W_2(\hat{a}, L) \quad \text{for all } \hat{a}, L \\ & u(w_1) + \int \int w_2(x, y) dF_{\hat{a}}(x) dG(y) \geq \bar{U}. \end{aligned}$$

The first-order conditions with respect to w_1 and $w_2(\hat{a}, L)$ are respectively:

$$1 - \mu u'(w_1) = 0 \quad (21)$$

$$1 - \lambda(\hat{a}, L) / (f_{\hat{a}}(\hat{a}) g(L)) - \mu u'(w_2(\hat{a}, L)) = 0 \quad \text{for all } \hat{a}, L, \quad (22)$$

where $\lambda(\hat{a}, L)$ and μ are respectively the (nonnegative) Lagrange multipliers associated with the constraints (12) and (13), where $\lambda(\hat{a}, L) \geq 0$ satisfy the complementary slackness condition:

$$\lambda(\hat{a}, L)(W_2(\hat{a}, L) - w_2(\hat{a}, L)) = 0 \quad \text{for all } \hat{a}, L. \quad (23)$$

Since the second-order conditions for minimization are satisfied, this immediately yields the form of the optimal long-term contract described in Lemma 3. The second period wage is equal to the reservation wage if the reservation wage is larger than the first-period wage, or it is equal to the first-period wage. ■

Proof of Proposition 1:

The comparative static results follow immediately from the performance measure P and the optimal contract as defined in Lemma 3.

To establish (iii), when $\hat{a} < 0$, $w_2^* = w_1^*$ and the second period-wage is constant. For any given value of $\hat{a} > 0$, there exists $\underline{L} \equiv w_1^* / \gamma \hat{a}$ such that $\frac{dw_2^*}{dL} = \gamma \hat{a} > 0$ if $L > \underline{L}$ and $\frac{dw_2^*}{dL} = 0$ if $L < \underline{L}$. ■

Proof of Proposition 2:

We will use the condition in Proposition 1-ii in Legros and Newman (2007) to prove that in a stable matching firms with better governance match with riskier CEOs.

Consider two firms with ownership structures $\delta_i > \delta_{i+1}$ and two CEOs indexed by “risk” $\sigma_i^2 > \sigma_{i+1}^2$. Recall that the choice of monitoring effort for firm δ_j is characterized by $e^* = \delta_j M(K, \sigma_a^2) / c$. Monitoring is all the more valuable that the CEO is more risky: compared to a firm whose shareholder does not monitor, a firm δ_j whose shareholder monitors with effort e^* makes $\delta_j^2 M^2(K, \sigma_a^2) / c$ extra profits, where $M(K, \sigma_a^2)$ is increasing in σ_a^2 (cf. Lemma 2). This means that the match $\{(\delta_i, \sigma_i^2), (\delta_{i+1}, \sigma_{i+1}^2)\}$ is more efficient than the other match.

Consider now the long-term contracts offered by the firm δ_{i+1} to both types of CEOs that lead to the same expected profit for the firm. To prove the matching configuration, we just need to show that if the firm δ_i designs long-term optimal contracts for the CEOs such that CEOs are indifferent between the contracts offered by the two firms, it would make larger expected profits by hiring the more risky CEO.

To start with, note that the expected utility of a CEO for a given long-term contract depends only on w_1^* . This means that the firm δ_i would propose the same long-term

contract as the firm δ_{i+1} if it needs to provide the same expected utility to the CEO. Since $M(K, \sigma_a^2)$ increases in σ_a^2 , the firm δ_i would make more profits by contracting with the riskier CEO. ■

Proof of Proposition 3:

The ranking of firms is such that $\delta_1 > \delta_2 > \dots > \delta_l > \dots > \delta_n$, and the ranking of CEOs is such that $\sigma_1 > \sigma_2 > \dots > \sigma_l$. According to Proposition 2, firm l is matched with CEO l , firm $l - i$ is matched with CEO $l - i$, and firm 1 is matched with CEO 1.

Consider firm l . The contract it is offering its CEO is designed to make firm $l + 1$ indifferent between attracting CEO l and hiring a manager, that is, to ensure that firm $l + 1$ would make negative expected profits by giving a better offer to CEO l . Similarly, the contract of CEO $l - 1$ offered by firm $l - 1$ is designed so that firm l makes the same expected profit with the optimal contract it offers CEO l and the expected profit it would make by attracting CEO $l - 1$.

That is, for any $i \in (1, \dots, l-1)$, the CEO contracts are determined so that firm $i + 1$ is indifferent between hiring CEO i or CEO $i + 1$. In addition, since the contract is fully determined by the wage w_1 , the wages of CEO i and CEO $i + 1$, denoted respectively by w_1^i and w_1^{i+1} , are determined so that firm $i + 1$ is indifferent between hiring CEO i or CEO $i + 1$. So assume for now that $w_1^i = w_1^{i+1} \equiv w_1$. Then the profit of firm $i + 1$ must be higher when it employs CEO $i + 1$, otherwise the matching equilibrium is not stable. We now find conditions under which the profit is higher when firm $i + 1$ employs CEO i with $w_1^i = w_1^{i+1}$, so that to restore the matching equilibrium it has to be that $w_1^i > w_1^{i+1}$. Similarly, when the profit is higher when firm $i + 1$ employs CEO $i + 1$, the opposite condition on wages obtains.

Remember that $\hat{\varphi}^i$ denotes the p.d.f. of \hat{a} for a CEO of type σ_i . We also call a^* the updated ability such that $w_1 = \gamma a^*$. It represents the lowest updated ability for which the participation constraint is binding. The ex-ante profit of the firm is piecewise linear, but the slope changes at $\hat{a} = -K$ and at $\hat{a} = a^*$ (as depicted in Figure 2).

The profit of firm $i + 1$ over both periods when it employs CEO i at wage w_1^i and the monitoring intensity is e^* is

$$\begin{aligned} \Pi(\delta_{i+1}, \sigma_i) &= \alpha + \gamma \bar{a} - w_1 + \int_{-K}^{a^*} [\alpha - w_1 + x] \hat{\varphi}^i(x) dx + \int_{a^*}^{\infty} [\alpha - w_1 + \gamma a^* + (1 - \gamma)x] \hat{\varphi}^i(x) dx \\ &+ e^*(\delta_{i+1}, \sigma_i) \int_{-\infty}^{-K} [\alpha - w_1 - K] \hat{\varphi}^i(x) dx + (1 - e^*(\delta_{i+1}, \sigma_i)) \int_{-\infty}^{-K} [\alpha - w_1 + x] \hat{\varphi}^i(x) dx \end{aligned}$$

The profit of firm $i + 1$ over both periods when it employs CEO $i + 1$ at wage w_1^i is:

$$\begin{aligned} \Pi(\delta_{i+1}, \sigma_{i+1}) &= \alpha + \gamma \bar{a} - w_1 + \int_{-K}^{a^*} [\alpha - w_1 + x] \hat{\varphi}^{i+1}(x) dx + \int_{a^*}^{\infty} [\alpha - w_1 + \gamma a^* + (1 - \gamma)x] \hat{\varphi}^{i+1}(x) dx \\ &+ e^*(\delta_{i+1}, \sigma_i) \int_{-\infty}^{-K} [\alpha - w_1 - K] \hat{\varphi}^{i+1}(x) dx + (1 - e^*(\delta_{i+1}, \sigma_{i+1})) \int_{-\infty}^{-K} [\alpha - w_1 + x] \hat{\varphi}^{i+1}(x) dx \end{aligned}$$

We want to show that $\Pi(\delta_{i+1}, \sigma_i) < \Pi(\delta_{i+1}, \sigma_{i+1})$ when $\gamma > \underline{\gamma}(K)$, or $K > \underline{K}(\gamma)$.

We have

$$\begin{aligned} \Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1}) &= \int_{-\infty}^{-K} (-K - x) (e^*(\delta_{i+1}, \sigma_i) \hat{\varphi}^i(x) - e^*(\delta_{i+1}, \sigma_{i+1}) \hat{\varphi}^{i+1}(x)) dx \\ &+ \int_{-\infty}^{a^*} x (\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x)) dx + \int_{a^*}^{\infty} [(1 - \gamma)x + \gamma a^*] (\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x)) dx \end{aligned}$$

We first prove the following intermediate result:

$$\int_{-\infty}^{-K} (-K - x) \hat{\varphi}^i(x) dx > \int_{-\infty}^{-K} (-K - x) \hat{\varphi}^{i+1}(x) dx.$$

Integrating by parts, we get

$$\begin{aligned} &\int_{-\infty}^{-K} (-K - x) \hat{\varphi}^i(x) dx - \int_{-\infty}^{-K} (-K - x) \hat{\varphi}^{i+1}(x) dx \\ &= \int_{-\infty}^{-K} (\hat{\Phi}^i(x) - \hat{\Phi}^{i+1}(x)) dx. \end{aligned}$$

Given that an increase in variance of the normal distribution is a mean preserving spread, we get that $\hat{\Phi}^i(x) > \hat{\Phi}^{i+1}(x)$ for $x < \bar{a}$. Since $\bar{a} > 0$, we have that $\hat{\Phi}^i(x) - \hat{\Phi}^{i+1}(x) > 0$ for $x < -K < \bar{a}$. The result follows.

First, assume that $K = \infty$. We have that

$$\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1}) = \int_{-\infty}^{a^*} x (\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x)) dx + \int_{a^*}^{\infty} [(1 - \gamma)x + \gamma a^*] (\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x)) dx.$$

We are just comparing the expected values under the two distributions of a function that is equal to x for $x < a^*$ and equal to $(1 - \gamma)x + \gamma a^*$ for $x > a^*$. This function is piecewise linear and concave. Since φ^i is a mean-preserving spread of φ^{i+1} , second-order stochastic dominance gives the result that $\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1}) < 0$. This implies that $w_1^{i+1} > w_1^i$ when $K = \infty$.

Second, assume that $K = 0$. Then

$$\begin{aligned} \Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1}) &= \int_{-\infty}^0 x((1 - e^*(\delta_{i+1}, \sigma_{i+1}))\hat{\varphi}^i(x) - (1 - e^*(\delta_{i+1}, \sigma_{i+1}))\hat{\varphi}^{i+1}(x))dx \\ &\quad + \int_0^{w_1/\gamma} x(\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x))dx + \int_{w_1/\gamma}^{\infty} [(1 - \gamma)x + w_1](\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x))dx. \end{aligned}$$

From Lemma 2, $e^*(\delta_{i+1}, \sigma_i) > e^*(\delta_{i+1}, \sigma_{i+1})$, so that we have

$$\int_{-\infty}^0 x e^*(\delta_{i+1}, \sigma_i) \varphi^{i+1}(x) dx > \int_{-\infty}^0 x e^*(\delta_{i+1}, \sigma_{i+1}) \varphi^{i+1}(x) dx$$

and $\lim_{\gamma \rightarrow 0} \int_{w_1/\gamma}^{\infty} [(1 - \gamma)x + w_1](\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x))dx = 0$. Thus,

$$\begin{aligned} &\lim_{\gamma \rightarrow 0} \Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1}) \\ &> \int_{-\infty}^0 x((1 - e^*(\delta_{i+1}, \sigma_{i+1}))(\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x)))dx + \int_0^{\infty} x(\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x))dx \end{aligned}$$

We are now comparing the expected values under the two distributions of a function that is equal to $x(1 - e^*(\delta_{i+1}, \sigma_{i+1}))$ for $x < 0$ and equal to x for $x > 0$. This function is piecewise linear and convex. Second-order stochastic dominance gives the result that $\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1}) > 0$ when $K = 0$ and $\gamma \rightarrow 0$, in which case we must have $w_1^{i+1} < w_1^i$.

Third, we now show that $\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1})$ is decreasing in γ and in K , which completes the proof. We have:

$$\begin{aligned} \frac{d}{d\gamma} \{\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1})\} &= \int_{w_1/\gamma}^{\infty} \frac{\partial}{\partial \gamma} \{[(1 - \gamma)x + w_1](\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x))\} dx \\ &= - \int_{a^*}^{\infty} x(\hat{\varphi}^i(x) - \hat{\varphi}^{i+1}(x)) dx < 0. \end{aligned}$$

The first equality follows from the fact that the terms coming from differentiating with respect to γ in the bounds of the integrals cancel out. This shows that $\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1})$ is decreasing in γ .

We also have that:

$$\begin{aligned} \frac{d}{dK} \{\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1})\} &= \int_{-\infty}^{-K} -(e^*(\delta_{i+1}, \sigma_i)\hat{\varphi}^i(x) - e^*(\delta_{i+1}, \sigma_{i+1})\hat{\varphi}^{i+1}(x))dx \\ &\quad + \int_{-\infty}^{-K} (-K - x) \left(\frac{d}{dK} e^*(\delta_{i+1}, \sigma_i)\hat{\varphi}^i(x) - \frac{d}{dK} e^*(\delta_{i+1}, \sigma_{i+1})\hat{\varphi}^{i+1}(x) \right) dx \end{aligned}$$

Recall from Lemma 2 that:

$$e^*(\delta_{i+1}, \sigma_i) = \frac{\delta_{i+1}}{c} \int_{-\infty}^{-K} (-K-x)\hat{\varphi}^i(x)dx,$$

so that

$$\frac{d}{dK} e^*(\delta_{i+1}, \sigma_i) = -\frac{\delta_{i+1}}{c} \hat{\Phi}^i(-K).$$

Substituting, we get

$$\begin{aligned} & \frac{d}{dK} \{\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1})\} \\ &= -e^*(\delta_{i+1}, \sigma_i) \hat{\Phi}^i(-K) + e^*(\delta_{i+1}, \sigma_{i+1}) \hat{\Phi}^{i+1}(-K) \\ & \quad - \frac{\delta_{i+1}}{c} \hat{\Phi}^i(-K) \int_{-\infty}^{-K} (-K-x)\hat{\varphi}^i(x)dx + \frac{\delta_{i+1}}{c} \hat{\Phi}^{i+1}(-K) \int_{-\infty}^{-K} (-K-x)\hat{\varphi}^{i+1}(x)dx \\ &= -\frac{2\delta_{i+1}}{c} \left(\hat{\Phi}^i(-K) \int_{-\infty}^{-K} (-K-x)\hat{\varphi}^i(x)dx - \hat{\Phi}^{i+1}(-K) \int_{-\infty}^{-K} (-K-x)\hat{\varphi}^{i+1}(x)dx \right). \end{aligned}$$

Since $-K < \bar{a}$, we have

$$\hat{\Phi}^i(-K) > \hat{\Phi}^{i+1}(-K),$$

Moreover, as proved in Lemma 2,

$$\int_{-\infty}^{-K} (-K-x)\hat{\varphi}^i(x)dx > \int_{-\infty}^{-K} (-K-x)\hat{\varphi}^{i+1}(x)dx,$$

So that

$$\hat{\Phi}^i(-K) \int_{-\infty}^{-K} (-K-x)\hat{\varphi}^i(x)dx - \hat{\Phi}^{i+1}(-K) \int_{-\infty}^{-K} (-K-x)\hat{\varphi}^{i+1}(x)dx > 0$$

Thus,

$$\frac{d}{dK} \{\Pi(\delta_{i+1}, \sigma_i) - \Pi(\delta_{i+1}, \sigma_{i+1})\} < 0. \quad \blacksquare$$

Proof of Proposition 4:

(i) Either $W_2 < w_1^*$, in which case $\frac{dw_2}{dL} = 0$ and $\frac{dw_2}{dV} = 0$, so that $\frac{\frac{dw_2}{dL}}{\frac{dw_2}{dV}}$ is independent of σ_a^2 . Or $W_2 \geq w_1^*$, in which case $\frac{dw_2}{dL} = \psi\eta$ and $\frac{dw_2}{dV} = \psi$, so that $\frac{\frac{dw_2}{dL}}{\frac{dw_2}{dV}} = \eta$, which is positive and decreasing in σ_a^2 (as shown in Proposition 1).

(ii) follows directly from assortative matching and (i). \blacksquare

Proof of Proposition 5:

We begin by proving result (ii). For any given firm, the PPS at P is either equal to zero or to ψ , and it is equal to zero if and only if $P < \kappa$, where $\kappa = \frac{w_1^*(1-\psi)}{\psi}$ (as derived in section 2.3). To prove result (ii), we show that ψ is increasing in δ_i , and that κ is decreasing in δ_i for sufficiently high K and γ .

First, (17) shows that ψ is increasing in σ_i^2 , and from Proposition 2 proves that CEOs with a high σ_i^2 are matched with firms with a high δ_i , so that ψ is increasing in δ_i . Second, for any given firm, κ is decreasing in ψ and increasing in w_1^* . In addition, for sufficiently high K and γ , Proposition 3 states that w_1^* is lower at firms with a higher δ_i . Since ψ is increasing in δ_i , it follows that, for sufficiently high K and γ , κ is decreasing in δ_i .

We now prove result (i). If for any δ_i we have $\text{PPS}_{\bar{P},i} = 0$, then the ratio $\frac{\text{PPS}_{\underline{P},i}}{\text{PPS}_{\bar{P},i}}$ is not well-defined. Given the form of the contract, for a given firm i the ratio is well-defined if and only if $\bar{P} > \kappa_i$. For a well-defined ratio, for any given firm i there are two intervals. For $\underline{P} < \kappa$, the ratio is equal to zero. For $\underline{P} > \kappa$, the ratio is equal to one. Result (i) then follows from the fact that κ is decreasing in δ_i . ■

Proof of Proposition 6:

The ranking of firms is $\delta_1 > \delta_2 > \dots > \delta_l > \dots > \delta_n$ and the ranking of CEOs is $\sigma_1 > \sigma_2 > \dots > \sigma_l$.

According to Proposition 2, firm l is matched with CEO l , firm $l - 1$ is matched with CEO $l - 1$, and firm 1 is matched with CEO 1. We consider an increase in δ_i for firms $i \in \{j, n\}$, which does not affect the ordering of firms and leaves the matching unchanged.

Consider firm l . Before the changes in monitoring capacity, the contract it was offering to its CEO was designed to make firm $l + 1$ indifferent between attracting CEO l and hiring a manager, that is, to ensure that firm $l + 1$ would make negative expected profits by giving a better offer to CEO l . Similarly, the contract of CEO $l - 1$ offered by firm $l - 1$ was designed so that firm l makes the same expected profit with the optimal contract it offers CEO l and the expected profit it would make by attracting CEO $l - 1$.

For any given i , when δ_i increases, the monitoring of firm i changes independently of the contracts, which leads to an increase in the expected profits of firm i .

The indifference condition that determines the wage w_l is:

$$\Pi(\delta_{l+1}, \sigma_l, w_1^l) = \Pi(\delta_{l+1}, 0, w_1^{l+1})$$

where $\Pi(\delta_l, \sigma_l, w_1^l)$ corresponds to the profit of firm δ_{l+1} that employs CEO σ_l and $\Pi(\delta_{l+1}, 0, w_1^{l+1})$ the profit of firm δ_{l+1} that employs a manager. We have

$$\begin{aligned} \Pi(\delta_{l+1}, \sigma_l, w_1^l) &= (\alpha + \gamma\bar{a} - w_1^l) + \int_{-K}^{w_1^l/\gamma} [\alpha - w_1^l + x] \hat{\varphi}^l(x) dx \\ &\quad + \int_{w_1^l/\gamma}^{\infty} [\alpha - w_1^l + \gamma a^* + (1 - \gamma)x] \hat{\varphi}^l(x) dx \end{aligned}$$

$$+ e^*(\delta_l + 1, \sigma_l) \int_{-\infty}^{-K} [\alpha - w_1^l - K] \hat{\varphi}^l(x) dx + (1 - e^*(\delta_{l+1}, \sigma_l)) \int_{-\infty}^{-K} [\alpha - w_1^l + x] \hat{\varphi}^l(x) dx$$

and $\Pi(\delta_{l+1}, 0, w_1^{l+1})$ does not depend on δ_{l+1} since managers are not monitored. The profits $\Pi(\delta_{l+1}, \sigma_l, w_1^l)$ are increasing in δ_{l+1} . To restore the indifference condition, there must be an increase in w_1^l .

Now consider CEO σ_{l-1} . His wage w_1^{l-1} is determined by the indifference condition:

$$\Pi(\delta_l, \sigma_l, w_1^l) = \Pi(\delta_l, \sigma_{l-1}, w_1^{l-1}).$$

A higher δ_l increases the expected profits firm l relatively more when it employs a riskier CEO, and w_l has increased to react to the competition of firm $l + 1$. These two facts make it more attractive for firm l to compete for CEO $l - 1$ relative to CEO l . In response, to maintain assortative matching, firm $l - 1$ needs to adjust the contract it offers to CEO $l - 1$ by increasing his first-period wage so that firm l is indifferent between hiring CEO l or CEO $l - 1$. By induction, this process leads to an increase in the wage of all CEOs. ■