

Carnegie Mellon University

**The Mechanism of Control in Organizations:
Essays on Imperfect Measures of Managerial Talent**

A DISSERTATION
SUBMITTED TO THE TEPPER SCHOOL OF BUSINESS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

Field of Accounting

by

Eunhee Kim

April 2017

Dissertation Committee:

Carlos Corona (co-Chair)
Jonathan Glover (co-Chair)
Pierre Jinghong Liang
Isa Hafalir

Copyright © 2017 Eunhee Kim

Acknowledgments

To be inserted

Contents

1	Introduction	1
2	The Market for Reputation: Repeated Matching and Career Concerns	5
2.1	Introduction	7
2.2	The Model	12
2.3	Analysis	19
2.3.1	Preliminaries: Complementarity and the Market Value for Managers . . .	19
2.3.2	Optimal Contract within a Firm-Manager Match	22
2.3.3	Repeated Matching Equilibrium	25
2.4	Applications	36
2.4.1	CEO Turnover and Firm Performance	37
2.4.2	Discussion	41
2.5	Conclusion	43
3	Career Concerns and Project Selection: The Role of Reputation Insurance	45
3.1	Introduction	48
3.2	Baseline Setup	54
3.3	Board of Directors' Monitoring as Insurance	57
3.4	Performance Disclosure Policy as Insurance	61
3.5	Severance Package as Insurance	65
3.6	Related Literature	68
3.6.1	CEO Turnover and Firm Performance	68
3.6.2	CEO Turnover, Reputation, and Project Selection	71
3.7	Conclusion	73
4	Generalists versus Specialists: When Do Firms Hire Externally?	77
4.1	Introduction	79
4.2	Related literature	83
4.3	The Model	86
4.3.1	Two Stage Production Process	88
4.3.2	Internal Promotion versus External Hiring	90
4.3.3	Incentive Contracts for Each Organization	92

4.4	Analysis	95
4.4.1	Management Effort Choice and Compensation	95
4.4.2	Organizational Efficiency	100
4.5	Empirical Implications	104
4.6	Conclusion	105
Appendices		107
A	Appendix for Chapter 1	109
B	Appendix for Chapter 2	121
C	Appendix for Chapter 3	131
References		137

Chapter 1

Introduction

Purpose and Scope of this Study

Managerial talent is neither observable nor divisible. Thus, performance measures that evaluate a manager's human capital are significant not only for compensating a manager through accurate appraisal, but also for inferring managerial talent. In practice, however, talent measures are often imperfect, thereby hindering firms from control of a manager's behaviors. As such, exploring the nature of imperfect measures and addressing how firms deal with it are important considerations in managerial accounting research. In this study, I investigate these issues through the lenses of both the market for managers and internal control. In particular, I examine how the market for managerial talent is influenced by imperfect talent measures and how such influence leads to a different matching of firms and managers. Then, taking the matching of firms and managers as given, I explore how firms make use of alternative contracting instruments to control managers' behaviors resulting from imperfect measures. The results provide novel explanations that

increase our understanding of 1) imperfect measures of managerial talent and 2) documented empirical evidence associated with managerial accounting research.

Outline of this Study

The unifying goal of this study is to better understand the impact of imperfect measures of managerial talent on a manager's behavior and how a firm attempts to control such behavior in both the market and firm levels. To distinguish different aspects of and issues developing from imperfect measures, Chapter 1 and Chapter 2 first discuss how an imperfect talent measure creates an agency problem and how firms respond to it when the measure is verifiable. Specifically, Chapter 1 aims to find a foundation of how imperfect talent measures influence the matching of firms and managers when managers have career concerns. Having found the tension from a manager's career concerns, Chapter 2 studies available, but less-understood contracting devices as internal control mechanisms that can serve as reputation insurance. Then, shifting the focus from a verifiable talent measure to an unverifiable measure, Chapter 3 examines, in the context of CEO hiring, how firms provide incentives to managers for developing firm-specific talent and the implications for subsequent firm performance and pay.

In Chapter 1, "The Market for Reputation: Repeated Matching and Career Concerns", I propose a multiperiod matching model of firms and managers to explain that labor market efficiency in sorting by imperfect measures may not guarantee economic efficiency in matching. In the model, firms compete for managerial talent and managers are concerned about their reputation. Due to the trade-off between match efficiency from productive complementarity and agency costs from managers' reputational concerns, assortative matching of firms and managers may fail. I

derive sufficient conditions for such failure with respect to size distributions of firms. The model can be applied to various agency problems with consideration of the labor market for managers, which will be particularly useful for analyzing cross-sectional patterns of two-sided matching, and aggregate firm performance and agency costs.

Motivated by the Chapter 1, in Chapter 2, “Project Selection and Career Concerns: The Role of Reputation Insurance”, I explore, in the context of CEO turnover, the latent aspects of existing practices in managerial accounting and control. In particular, this chapter asks how well different governance practices provide incentives for project selection when managers have career concerns and how such practices influence a firm’s decision of whether to replace their CEO. I show that a board of directors’ monitoring, performance disclosure policy, and a severance package serve as reputation insurance and mitigate a manager’s career concerns through different mechanisms. However, the incentive effects of reputation insurance are followed by a weakened turnover-performance relation. The board’s monitoring makes the relation weaker since the board’s information serves as a substitute for the project earnings. The non-disclosure of a CEO’s performance at departure weakens the relation due to information suppression. The presence of severance pay, on the other hand, creates performance tolerance for firms in order not to pay out, thereby lessening the turnover-performance sensitivity. I also provide empirical predictions related to the existing CEO turnover and governance practices based on the perspective of reputation insurance.

In contrast to Chapter 1 and Chapter 2, in Chapter 3, “Generalists versus Specialists: When Do Firms Hire Externally”, the discussion centers on the aspect of unverifiable talent measures. This chapter is inspired by puzzling observed associations among CEO appointments, pay, and

firm performance. In recent decades, the trend of external CEO hiring has increased, a practice often involving high outsider pay premiums. Most academics and practitioners ascribe the practice of outsider premiums to two factors: managerial talent and a match between a firm and CEO. However, this perspective seems to overlook that, after an outsider CEO is hired, firm performance often becomes unsatisfactory. To understand the missing link between CEO hiring choices, I consider CEO hiring as an incentive device for non-CEO employees for firm-specific talent acquisition. Specifically, I develop a multitask-multiagent team production model where each task sequentially requires a firm-specific talent and a management decision. Both internal promotion (the specialist CEO) and external hiring (the generalist CEO) provide incentives of talent acquisition to non-CEO employees but through different mechanisms. I identify conditions under which either internal promotion remains optimal or external hiring becomes optimal. This optimal contracting framework for multiple agents also explains why outsider CEOs appear to be paid more than insider CEOs, and how the performance of external hiring firms tends to be worse than the performance of internal promoting firms in spite of the higher pay.

Chapter 2

The Market for Reputation:

Repeated Matching and Career Concerns

Abstract

I propose a multiperiod matching model of firms and managers to explain that labor market sorting with imperfect measures may not guarantee economic efficiency in matching. In the model, firms compete for managerial talent and managers are concerned about their reputation. Due to the trade-off between match efficiency from productive complementarity and agency costs from managers' reputational concerns, assortative matching of firms and managers may fail. I derive sufficient conditions for such failure with respect to the size distributions of firms. The model can be applied to various agency problems with consideration of the labor market for managers, which will be particularly useful for analyzing cross-sectional patterns of two-sided matching, and aggregate firm performance and agency costs.

2.1 Introduction

This paper investigates how labor market sorting can impede economic efficiency in matching of firms and managers. In particular, I argue that performance-based repeated sorting creates managerial career concerns, thereby distorting equilibrium matching patterns. Understanding the characteristics of matched firms and managers (e.g., who hires which manager, or who works at which firm), is important not only for analyzing the labor market, but also for exploring the impact of endogenous matching on an individual firm's or manager's behaviors. Much is known about two-sided matching with fixed characteristics. However, in the context of firms and managers, the characteristics of at least one side of the match are not always fixed: managers may build their track records to form the perception of their talent. When the characteristics of one side of the market are endogenous, it is unclear whether matching patterns will be similar to the exogenous case. To answer this question, I propose a multiperiod matching model of firms and managers where firms compete for managerial talent and managers are concerned about their perceived talent (i.e., reputation).

I find that, even with productive complementarity between firms and managers, career concerns might lead to distortions in the matching of firms and managers. When firms and managers are productive complements, the benchmark efficient matching pattern is well known to be positive assortative (Becker (1973)): the best matched with the best and the worst matched with the worst. However, career concerns influence a manager's actions, which might not be in the best interest of a firm. Ex ante, this implies that a firm needs to bear agency costs in order to induce desirable actions from the manager. Thus, when a manager has career concerns, a firm

faces the trade-off between match efficiency from productive complementarity and agency costs from managerial career concerns. I derive sufficient conditions under which this trade off obtains non-assortative matching as an equilibrium.

The model combines four features. First, heterogeneous firms, which differ in size, compete for managerial talent. Second, each manager is of two types, good or bad, but the type is unknown to everyone, including the manager himself. The type characterizes a managerial talent in obtaining high quality information with a costly effort. Such information facilitates the manager's choice between a risky project and a safe project. Third, firm performance is publicly observed and is used by all market participants to update their beliefs about the talent of each manager, that is, the manager's reputation. Lastly, the update in manager reputation is followed by a rematching between firms and managers. While these features individually are not new, the interaction between these features shows that the labor market efficiency in sorting may not guarantee the economic efficiency in matching.

The underlying reason for this economic *inefficiency* is a distortion in a manager's preference for risk exposure. Due to complementarity between firm size and managerial talent, large firms are willing to pay more for managerial talent, which makes a manager's market wage determined by both a manager's reputation and firm size. Since a manager's project choice in a current period leads to his reputation update with different market wages in the next period, each manager has different induced preferences for the risky project. If the expected future wage upon the risky project is less than the future wage upon the safe one, a manager prefers to choose the safe project, and thus loses his incentive to acquire information. In this case, a firm needs to offer extra pay for information acquisition. If overcoming such a preference is too expensive relative

to marginal benefit of the manager's reputation, then a firm may find it profitable not to match with the career-concerned manager even if the manager's reputation is high.

Interestingly, such distortions in matching are affected by the distributions of firm size. The logic is the following. Since a manager's market wage is determined by firm size, a distribution of firm size forms a distribution of market wages for managers' reputation. Thus, depending on the shape of firm size distributions, managers' induced preferences for risk exposure can be heterogeneous, and even non-monotonic in a manager's reputation. In particular, while a high reputation manager may be willing to take the risky project, a medium reputation manager exhibits induced preference for the safe project if firm size increases much faster as size increases. Since the faster increase in firm size at the top of the distribution directly leads to the faster increase in market wage for a high reputation manager, the managers at the top may actively seek risk while the medium may not. This logic is reversed once the increase in firm size is slower as size increases. Consequently, distributions of firm size influence cross-sectional differences in managers' induced preferences for risk exposure, thus generating different distortions in matching patterns. I derive implications with respect to the firm size distributions that obtain non-assortative matching patterns.

Beyond just showing that managers' career concerns can create matching distortions, the model can be applied to various economic problems where one side's talent is traded, including matching of auditor and client, analyst and firm, and board of directors and firm. By introducing variations into the agency problem and/or the matching problem depending on a particular conflict or friction of interest, the model I propose provides a framework that enables analyzing the interactions between the market forces and agency problems. For example, one can analyze the

impact of the audit labor market on aggregate audit quality in the context of matching of auditors and clients. Given auditors' concerns about their reputations, the framework I propose can examine how the labor market for auditors influences the formation of firms and auditors, and how the market affects aggregate audit quality. The model can also be applied to analysts and their coverage firms. Given the analysts' concerns about their reputations, one can analyze the interplay between the labor market for analysts and agency conflicts within a match between analyst and coverage firms. In particular, one can investigate how aggregate forecast accuracy or the quality of analyst recommendations changes. To provide a direct application, in Section 4, I apply the model to the matching of firms and CEOs and offer new insight into CEO turnover-performance sensitivity.

The model in this paper is related to recent work on two-sided matching. Terviö (2008) develops a competitive assignment model to explain the observed levels of CEO pay. By considering the assignment of CEOs with different ability to firms of different sizes, Terviö shows how seemingly excessive levels of CEO pay can be derived from competitive market forces with fixed attributes of two sides and absence of agency problems. While I am following Terviö in determining a manager's market wage, in the model I propose, the attribute of a manager (i.e., reputation) evolves whenever its matched firm performance is realized, thus creating agency frictions in a dynamic matching framework.

Anderson and Smith (2010), the most closely related to this paper in terms of failure of assortative matching patterns, show that, with unknown ability but evolving reputation of two sides, matching patterns can be distorted in early periods. The trade-off of this failure is between the match efficiency and information learning. In Anderson and Smith, exogenous production

generates information about the two sides. Then, by matching with extreme reputation (0 or 1), one can learn more about his type through exogenous match outcome. In the framework I propose, a key distinction is that the match outcome is endogeneous. The project decision made by a manager determines both match outcome and the manager's new reputation. It is the endogenous production that creates agency frictions, thus lowering the match efficiency. Although the predictions of matching patterns in this paper are similar to Anderson and Smith's, the results are driven by different trade-offs.

Legros and Newman (2007) shows that in the context of non-transferability of utility, assortative matching patterns need type-payoff (as opposed to type-type) complementarity. That is, if a matched partner's exogeneously given transfer is too high, then positive assortative matching might not arise. Building on the result of type-payoff complementarity, I endogenize a matched partner's transfer and derive conditions under which assortative matching patterns can fail. In addition to these studies, there are some applied studies that are related to this paper, which I will discuss in more detail in Appendix.

The outline of the present paper is as follows. In Section 2, I describe the basic model setup for an individual firm and manager. After characterizing the economic ingredients, Section 3 analyzes a repeated matching problem with career concerns. In Section 4, I discuss potential applications of the baseline model. Since empirical predictions depend on specific applications, I defer a discussion of testable hypotheses until Section 4. Section 5 concludes.

2.2 The Model

The innovation of this paper is to endogenize the evolution of managerial reputation and to analyze how this endogeneity influences and is influenced by the labor market for managers. The model highlights the interaction between the perception of managers' types through performance information and induced risk preferences because of managerial career concerns. The economy consists of a continuum of agents (managers) and a continuum of principals (firms). The economy lasts for two periods. Within a period, the sequence of events is as follows: 1) at the beginning, the market-wide matching takes place with a single period contract between a matched principal and an agent; 2) the agent exerts effort to select an investment project; 3) the investment outcome is realized and payoffs are realized; and 4) both principal and agent return to the market for the next period matching (if this is the last period, the game ends). All players are risk neutral and share the same horizon with no discount factor. Also, agents are protected by limited liability.

Heterogenous Principals and Project Selection: The firms differ in their size represented by $S \in [S_{min}, \infty)$ with a well-defined smooth distribution function $G(S)$. To analyze matching patterns depending on the distributions of firm size, I do not impose any parametric assumptions on the size distribution. The firm size can be understood as a one-dimensional summary statistic that captures multi-attributes of firms with respect to performance.

Since each firm is uniquely characterized by its firm size, I will use the terms firm and principal interchangeably. The main task of each principal is to hire a manager, to design a single period take-it-or-leave-it contract, and to replace (or retain) the incumbent manager in order to

maximize the principal's payoff, which is modeled as the expected project return less the compensation for the manager. Let $y_0 > 0$ denote the principals' outside option in case they do not hire any manager. I assume that y_0 is sufficiently small that every firm wants to hire a manager from the market.

Each firm has the choice to invest in one of two projects: a risky project denoted as I_r or a safe project denoted as I_s . The safe project I_s will return a certain outcome, m , which can be interpreted as a status quo. The risky project will return either a success, $h > m$, or a failure, $l < m$. Without loss of generality, assume that the investment cost is the same for both projects, which is normalized to zero. The probability of success for the project I_r depends on a state variable consisting of $\{s_1, s_2, s_3\}$, where s_1 indicates I_r will generate h , s_2 and s_3 indicate I_r will generate l . The unconditional probability of each state is $Pr(s_1) = \alpha p$, $Pr(s_2) = \alpha(1 - p)$, $Pr(s_3) = 1 - \alpha$, where $\alpha \in (0, 1)$, and $p \in (0, 1)$. It is immediate to see that $Pr(h|\{s_1, s_2\}, I_r) = p$, $Pr(l|\{s_1, s_2\}, I_r) = 1 - p$, and $Pr(l|s_3, I_r) = 1$. The primitive parameters, α, p , are identical and independent for every firm in each period. To capture differences in firms, I assume the scale of operations (Sattinger (1993)). That is, a firm's project return is the project outcome (denoted as X) multiplied by its firm size, $S: S \times X$, where $X \in \{h, m, l\}$.

Managerial Talent and Information Acquisition: A manager selects a project if hired. The manager can exert effort at cost $c > 0$ to acquire information about a realized state and then select a project based on the signal that his effort generates. The signal that a manager can acquire is drawn from $\{\{s_1, s_2\}, \{s_3\}\}$. For convenience, let $r = \{s_1, s_2\}$, $s = \{s_3\}$. That is, the set of signals is coarser than the set of states. Let $v, \hat{v} \in \{r, s\}$ denote a partition of state variables and a manager's acquired information respectively. To capture the manager's talent, I assume that

there are two types of managers, $\tau = G$ and $\tau = B$, denoting a good and a bad type respectively. Hereafter, I will use the terms G -manager and a good type manager, and B -manager and a bad type manager interchangeably. The two types differ in their ability to acquire information about the realized state. By exerting effort, a good type manager knows if a realized state belongs to r or s (i.e., $\hat{v} = v$), but a bad type manager receives $\hat{v} = v$ with probability of $\beta \in (0, 1)$ and $\hat{v} \neq v$ with the complementary probability. Without effort, both types do not receive a signal. I assume that the ex ante probabilities of realization of signals is the same for both types so that the acquired signal per se does not communicate any information about the manager's type. This assumption is captured by setting $\alpha = 1/2$.¹ Table 2.1 describes the event trees for each type of manager. The derivation of each event probability is presented in Appendix. To make the project selection problem non-trivial, assume that $ph + (1 - p)l > m > l$, i.e., if $v = r$, then I_r is more profitable, and if $v = s$, then I_s is more profitable.

Following the career concern literature (e.g., Holmstrom (1999)), I assume that each manager's type is unknown to everyone including themselves. All managers are endowed with an initial reputation that represents the probability of the agent being a good type. To see how career concerns differ depending on a manager's reputation, I assume that there are initially three levels of reputation, $0 < \gamma^l < \gamma^m < \gamma^h < 1$.² The total measure of managers in the labor market

¹That is, $Pr(\hat{v}|G) = Pr(\hat{v}|B)$ for all \hat{v} , $\Leftrightarrow \alpha = \alpha\beta + (1 - \alpha)(1 - \beta)$

²The assumption of this initial reputation distribution is for the sake of simplicity. The basic idea extends directly to a more general reputation distribution. The description of the economy where the initial endowment of reputation are continuously different is included in the appendix. The other assumption that the market's evaluation of each manager is characterized as a single dimensional characteristic is made for analytical simplicity. Without agency problems, the extension to multi-dimensional attributes is considered in Eisfeldt and Kuhnen (2013) and Pan (2015), and the multiple attributes are summarized by a single dimensional statistic through a linear combination of attributes

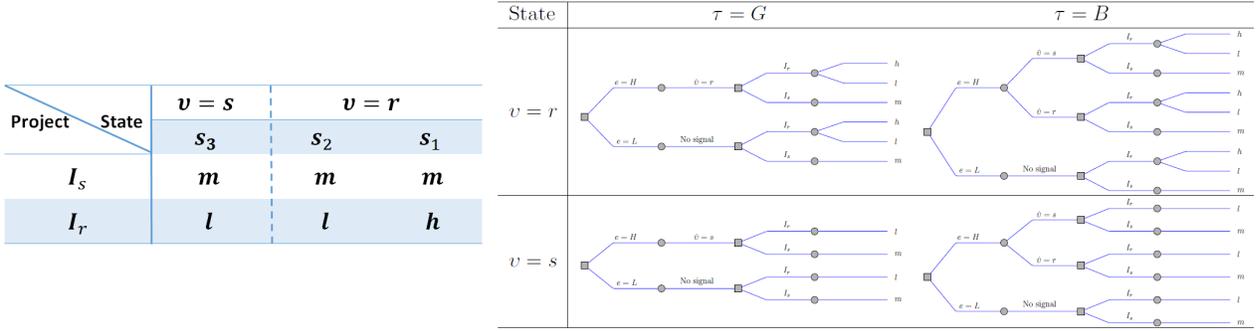


Figure 2.1: The Event and Decision Trees of Each Type for Each State

The left figure describes the project earnings depending on projects and states. The right figure presents event and decision trees and their outcomes depending on a manager's type and their decision in each state. Every outcome is feasible under both types, but given the structure of the signal for true states, h is more likely for G -managers, and l is more likely for B -managers.

is denoted as $\Gamma \in \mathbb{R}$ such that $\Gamma = 1 + \eta$, where $\eta \in (0, 1)$ denotes a measure of managers with reputation γ^l .³ After each manager's project outcome is realized, the manager's reputation is updated and a new reputation is used for the matching in the next period. Let $\{\gamma\}_t$ denote a set of all levels of manager's reputation in period $t = 1, 2$.

Let $\omega_t(\gamma)$ denote a manager's market value (or outside option) depending on the manager's reputation γ in period t .⁴ Hereafter, I will use the terms market value, outside option, and reputation premium interchangeably. Because high reputation implies that the manager is more likely to be a good type, the manager's outside option would be non-decreasing with reputation. Indeed, I shall show that the manager's market value, which is endogenously determined by the

³ η can be any positive number, but this is for the sake of simplicity in the benchmark matching pattern, which results in no matches with γ^l managers.

⁴The market value can be interpreted as the maximum periodic compensation that other firms are willing to pay to hire the manager or the expected payoff of an alternative job opportunity.

manager labor market and the firm size distribution ($G(S)$), is strictly increasing with reputation. To this end, every manager cares about their market perception, γ , as it determines not only his payoff today, but also his payoff tomorrow through rematching. The history of each manager's reputation is publicly observable. Thus, a manager's decision to exert effort and to choose a project is influenced by his career concern for the next period matching. Let $w_0 > 0$ denote a reservation utility for every manager: the periodic payoff for each manager must be greater than or equal to w_0 .

Repeated Matching and Equilibrium: At the beginning of each period, market-wide matching or rematching takes place. After the project outcome is realized at the end of period 1, the manager's reputation is revised, and the demand for rematching of managers and firms arises. From Figure 2.1, the probability of each project outcome differs depending on γ . Let $Pr(X|\gamma)$ denote probability of project outcome X conditional on the manager reputation γ . To describe the matching and rematching, consider the problem faced by firm S . Let $Y_t(S, \gamma)$ denote the expected project return for firm S in period t when it is matched with manager γ . That is,

$$Y_t(S, \gamma) = S \times \left(h \times Pr(h|\gamma) + m \times Pr(m|\gamma) + l \times Pr(l|\gamma) \right)$$

Let w^X denote a transfer upon the project outcome $X \in \{h, m, l\}$, and $E[w^X|\omega_t(\gamma)]$ denote the expected compensation cost given the market value of $\omega_t(\gamma)$. Then, firm S , taking the market value of each manager as given, chooses the optimal manager γ , to maximize its payoff which is expected project return net of the manager's compensation.

$$\max_{\gamma, w^X} Y_t(S, \gamma) - E[w^X|\omega_t(\gamma)]$$

The sequence of events in each period is summarized in the Figure 2.2. Assume that acquiring

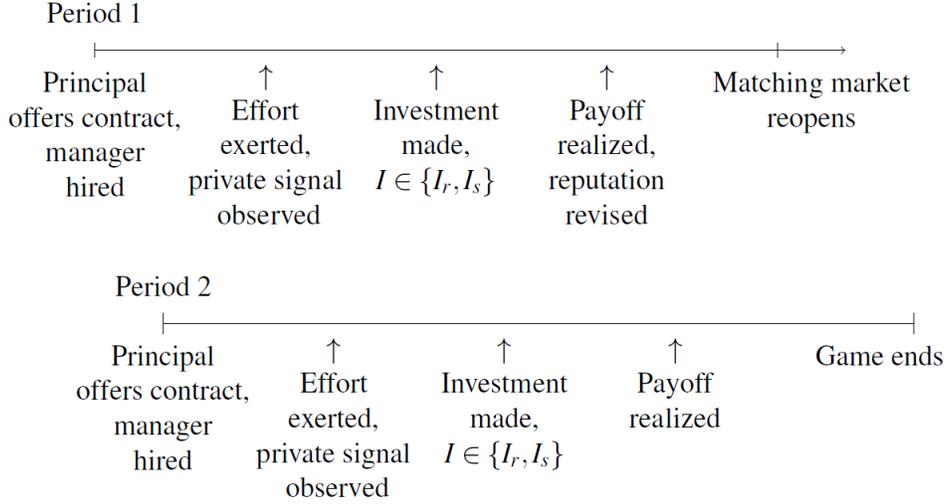


Figure 2.2: **Timeline**

information about the realized state is valuable enough that principals want to induce high effort from their matched managers.

Assumption 1. For $\forall \gamma \in \{\gamma\}_t, t = 1, 2,$

- $Pr(v = r|\hat{v} = r, \gamma)(ph + (1 - p)l) + Pr(v = s|\hat{v} = r, \gamma)l > m$
- $Pr(v = r|\hat{v} = s, \gamma)(ph + (1 - p)l) + Pr(v = s|\hat{v} = s, \gamma)l < m$
- $Y_t(S, \gamma) - E[w_t^X|\omega_t(\gamma)] \geq \max_x x \times \left(Pr(h)(S \times h - w_t^h) + Pr(l)(S \times l - w_t^l) \right) + (1 - x) \times (S \times m - w_t^m), \text{ for any } x \in [0, 1]$

Under the first and the second inequalities, a manager's information is valuable enough that selecting the project based on the manager's information is always efficient. Under the third inequality, it is optimal to induce $e = H$ to acquire the signal from the manager instead of randomly choosing the two projects without the manager's signal. The derivation of this assumption with respect to parameters is provided in Appendix. Then, an equilibrium in each period is defined as follows.

Definition 1. An equilibrium in period t is a tuple $(\{\omega_t^*(\gamma), w_t^X\}, \mu_t)$ that consists of a contract determined by their market value function $\omega_t^*(\gamma)$ for manager γ that induces them to exert effort, and a matching function of $\mu_t : [S_{min}, \infty) \rightarrow \{\gamma\}_t$ such that 1) a contract for manager γ induces effort and participation, 2) each firm chooses its manager optimally and each manager takes the best offer available or chooses not to participate, 3) the market clears.

For simplicity, I assume that if a matching function in period 2 gives the same reputation as the incumbent manager, then the principal retains the incumbent. To simplify analysis and to avoid any bargaining game within a match, I also assume that all the bargaining power is given to firms.⁵ I focus on an efficient equilibrium that maximizes the aggregate payoffs of the two sides.⁶ In the next section, I first solve for the market value for managers, and the corresponding contracts, and then solve for equilibrium matching patterns.

⁵Thus, in equilibrium, a firm will just pay the minimum required payment that is determined by the matching market to hire a particular manager.

⁶Notice that the matching function $\mu_t(S)$ can be found by considering the firms' future payoffs, that is, for $t = 1, 2$, $\mu_t(S) \in \operatorname{argmax} \sum_t Y_t(S, \gamma) - E[w^X | \omega_t(\mu_t(S))]$. Since firms' characteristic (i.e., size) is assumed to be fixed and does not change the next period outcome in that if $\mu_{t+1}(S) \neq \gamma_{incumbent}$, then the incumbent manager will not be assigned, and if $\mu_{t+1}(S) = \gamma_{incumbent}$, then the reassignment problem will anyway match them regardless of their history. Due to this static feature and the absence of a long-term contract, the equilibrium matching function is determined by maximizing the static payoff to the firms.

2.3 Analysis

2.3.1 Preliminaries: Complementarity and the Market Value for Managers

To find an equilibrium matching of firms and managers, I first show why firms compete for a manager's reputations. A high reputation means that the manager is more likely to be a good type, thus getting a more precise signal upon high effort. The following lemma confirms this.

Lemma 1. $\frac{\partial^2 Y}{\partial \gamma \partial S} > 0$, thus the project return (i.e., match output) exhibits complementarity.

Lemma 1 shows that there is complementarity between firm size and reputation. The complementarity indicates that large firms enjoy a greater return from hiring high reputation managers than small firms. This efficiency based on size also indicates that large firms are willing to pay more to bid away high reputation managers than small firms. The standard matching literature has shown that an efficient matching pattern shall be positive assortative, i.e., the largest firm is matched with the best reputation, the next largest firm is matched with the next best reputation, and so on. Since there are more managers, the matching shall clear managers from the top (there is no matched manager whose reputation is lower than any of unmatched manager given that both managers are willing to participate).

The market values for managers are determined by an equilibrium matching. The efficient matching of managers and firms then must satisfy two types of constraints: the sorting (SC) and the participation (PC) constraints. The sorting constraint states that each firm prefers the matched manager at their equilibrium market value to other managers. The participation constraint for

firms states that every firm's payoff from the equilibrium match must be greater than or equal to its payoff from no match. Similarly, the participation constraint for managers states that every manager's payoff from the equilibrium match must be greater than or equal to its payoff from no match. More formally,

$$Y_t(S, \gamma) - E[w_t^X | \omega_t(\gamma)] \geq Y_t(S, \gamma') - E[w_t^X | \omega_t(\gamma')] \quad \forall S, \gamma \quad (\text{SC}(S, \gamma))$$

$$Y_t(S, \gamma) - E[w_t^X | \omega_t(\gamma)] \geq y_0 \quad \forall S \quad (\text{PC-firm})$$

$$E[w_t^X | \omega_t(\gamma)] - c + \Pi_t^P(\gamma) \geq w_0 + \Pi_t^{NP}(\gamma) \quad \forall \gamma \quad (\text{PC-manager})$$

where $\Pi_t^K(\gamma)$, $K \in \{P, NP\}$ denotes a manager γ 's expected future market value contingent on the current reputation level of γ and the participation decision, P representing participation, NP denoting sitting out the matching market. Observe that a managers' expected payoff in period 1 includes the outcome of a matching in period 2 due to rematching. Also, the sorting constraints and the participation constraints for firms are static in that those constraints influence the current period equilibrium outcome.⁷ However, the participation constraints for managers not only influence the current period equilibrium but also are influenced by the next period equilibrium.

As a benchmark, I first derive the market value for a manager's reputation when there is no agency friction. Due to discrete structure of managerial characteristics, the process of market value determination is not the same as in Terviö (2008). Define $M(\gamma)$ as the measure of managers with reputations greater than or equal to γ . Let the set of managers be characterized as N tiers $1 > \gamma_1 > \gamma_2 > \gamma_3 > \dots > \gamma_N > 0$: there are N different levels of reputation. Then, $M(\gamma_i) > M(\gamma_j)$ for $i < j$. Due to the complementarity between firm size and reputation, it is efficient to assign high reputation to large firms. Then, solving for the equilibrium matching is identical

⁷This is because a firm can always hire a manager in the market in each period.

with finding the group of firms that will be matched with the same reputation managers. That is, matching is identified by characterizing the firm size thresholds that determine the group of firms for each reputation level: $\mu_t(S) = \gamma_i$ for all $S \in [S[i], S[i-1])$, $i = 2, \dots, N$ where $S[i] = G^{-1}(M(\gamma_i))$. Let the smallest firm within each group be a threshold firm (i.e., $S[i]$ for γ_i). Since firms have all the bargaining power, the market value for each manager is determined by binding sorting constraints that make a threshold firm S indifferent between hiring the equilibrium match and the next best match.

$$Y_t(S, \gamma) - E[w_t^X | \omega_t(\gamma)] = Y_t(S, \gamma') - E[w_t^X | \omega_t(\gamma')]$$

which yields $E[w_t^X | \omega_t(\gamma)] = Y_t(S, \gamma) - Y_t(S, \gamma') + E[w_t^X | \omega_t(\gamma')]$. Lemma 2 characterizes the market value based on this discussion.

Lemma 2. *Suppose that managers are characterized as $1 > \gamma_1 > \gamma_2 > \dots > \gamma_N > 0$. Without agency frictions, a reputation based compensation for γ_i and γ_{i+1} , $i = 1, 2, \dots, N-1$, satisfies,*

$$E[w_t^X | \omega_t(\gamma_i)] = (\gamma_i - \gamma_{i+1})FS[i] + E[w_t^X | \omega_t(\gamma_{i+1})].$$

Equivalently,

$$E[w_t^X | \omega_t(\gamma_i)] = \sum_i^{N-1} (\gamma_i - \gamma_{i+1})FS[i] + E[w_t^X | \omega_t(\gamma_N)].$$

where $F = (1 - \beta)(Pr(h)(h - l) + (Pr(m) - Pr(v = r))(m - l))$, $\omega_t(\gamma_N) = w_0$.

The endogenous reputation based market value captures a trade-off between the marginal benefit and the marginal cost of managers' reputation: the extra improvement in the project return due to the increase in reputation must be added to the wage required to hire the next alternative manager. It is worth emphasizing that the extra pay is not only driven by the probability of

being a good type, but also driven by the size of a firm that is indifferent to hiring either of two alternatives.⁸

Before I analyze an equilibrium matching, I shall show that there exists an equilibrium. Shapley and Shubik (1971) and Kaneko and Yamamoto (1986) have shown the existence of equilibria of a decentralized assignment problem.⁹ However, their standard proofs are not directly applicable to my economy because I introduce moral hazard through career concerns. In the next Section, I constructively show that there exists an equilibrium in every period in this decentralized repeated matching problem even with moral hazard through career concerns. Before proceeding further, I will first consider the incentive problem of a particular firm to find an optimal contract. In the following section, I will find a market equilibrium.

2.3.2 Optimal Contract within a Firm-Manager Match

This subsection finds an optimal contract between firm S and manager γ . Firm S finds (w_t^h, w_t^l, w_t^m) in period t by considering manager γ 's market value $\omega_t(\gamma)$ as given. As a benchmark, I first investigate the period 2 contract when there is no reputational incentive left. The required constraints

⁸An alternative way of deriving a market value function for managers is to assume a simple Nash bargaining (Firm's payoff, manager's payoff) = $(k \times Y_t(S, \gamma), (1 - k)Y_t(S, \gamma))$ where $k \in (0, 1)$. In this case, due to complementarity, matching with a large firm is strictly preferred for the same k . But, the competitiveness within each tier requires the equal treatment for identical reputation managers to have stable matching. More formally, for any firms S_1, S_2 that are assigned to the same γ , the equal treatment is characterized by $(1 - k_{S_1})Y_t(S_1, \gamma) = (1 - k_{S_2})Y_t(S_2, \gamma)$. It is clear that $k_{S_1} > k_{S_2}$ if $S_1 > S_2$. That is, the surplus split to manager γ decreases as its matched firm size increases.

⁹In a central assignment problem, equilibria are found by solving linear programming problem (Roth and Sotomayor (1992)), and there exists a solution consisting of only zero and one (Dantzig (1963)).

are as follows.

$$-c + \sum_{X \in \{h,l,m\}} Pr(X|\gamma)_{\gamma} w_2^X \geq \omega_2(\gamma) \quad (\text{IR})$$

$$-c + \sum_{X \in \{h,l,m\}} Pr(X|\gamma) w_2^X \geq \max_{x \in [0,1]} x w_2^m + (1-x)(Pr(h)w_2^h + Pr(l)w_2^l) \quad (\text{IC})$$

where $w_2^X \geq 0$. The (IR) constraint stipulates that the manager with γ reputation must be paid at least his outside option, $\omega_2(\gamma)$, the (IC) constraint requires that the manager prefers to exert effort to make an efficient investment choice rather than shirking and choosing a safe choice, choosing a risky project, or any combination of the two. In the appendix A.2.6, I show that, under the parametric Assumption 1, the incentive compatibility constraint needs only consider the safe choice without effort instead of mixing any combination of the two projects in period 1 and 2 (i.e., x is 1). Considering the (IR) and (IC) constraints with non-negative payments, the principal finds a contract to maximize,

$$\sum_{X \in \{h,l,m\}} Pr(X|\gamma)(S \times X - w_2^X)$$

Due to risk neutrality, there can be multiple solutions that generate the same payoff for the principal. The following Lemma 3 finds characteristics of an optimal solution.

Lemma 3. *(Without Career Concerns) The optimal contract in period 2 is characterized as follows.*

$$\frac{w_2^h - w_2^m}{w_2^m - w_2^l} = \frac{Pr(l)}{Pr(h)} = \frac{1 - \alpha p}{\alpha p}$$

The above feature captures the incremental pay for each performance ($w_2^h - w_2^m$ and $w_2^m - w_2^l$), which I call pay performance sensitivity.¹⁰ When there is no career concern, it is clear that the

¹⁰The literature on CEO compensation defines PPS as the change in CEO pay for the change in shareholder wealth

pay performance sensitivity defined above is independent of reputation γ , rather it only depends on the characteristics of projects. This suggests that providing incentives to induce effort for information acquisition and to select the right project is not influenced by a manager's implicit incentives for their future market value.

Now, I consider the period 1 contract which will depend on a manager's implicit incentives as realized project outcome will change the manager's reputation, thus the market value. Let γ^X denote the updated reputation from project outcome X . Then, the (IR) and (IC) constraints are characterized as follows.

$$-c + \sum_{X \in \{h, m, l\}} Pr(X|\gamma)(w_1^X + \omega_2(\gamma^X)) \geq \omega_1(\gamma) + \sum_{X \in \{h, m, l\}} \omega_2(\gamma^X) \quad (\text{IR})$$

$$-c + \sum_{X \in \{h, m, l\}} Pr(X|\gamma)(w_1^X + \omega_2(\gamma^X)) \geq w_1^m + \omega_2(\gamma^m) \quad (\text{IC})$$

Then, the optimal compensation contract contingent on performance must satisfy the following.

Lemma 4. *(With Career Concerns) The optimal contract in period 1 is characterized as follows.*

$$\frac{w_1^h - w_1^m}{w_1^m - w_1^l} = \frac{\frac{cPr(l)}{Pr(h|\gamma) - Pr(h)(1 - Pr(m|\gamma))} - \overbrace{\left(\omega_2(\gamma^h) - \omega_2(\gamma^m)\right)}^{\text{Upside potential}}}{\frac{cPr(h)}{Pr(h|\gamma) - Pr(h)(1 - Pr(m|\gamma))} - \underbrace{\left(\omega_2(\gamma^m) - \omega_2(\gamma^l)\right)}_{\text{Downside potential}}}$$

(Jensen and Murphy (1990)). The performance sensitivity defined above implicitly considers the principal's wealth change as a linear function of $m - l$ and $h - m$. Also, this definition is useful to see how the shape of pay is more (or less) convex in the presence of different dynamic incentives. The brief argument to back up this definition comes from the definition of a convex function: a pay function $f(X_i)$ is convex if $\lambda f(h) + (1 - \lambda)f(l) > f(\lambda h + (1 - \lambda)l)$.

For λ such that $\lambda h + (1 - \lambda)l = m$, this definition is equivalent to $\frac{\lambda(f(h) - f(m))}{(1 - \lambda)(f(m) - f(l))} > 1$ for $\lambda = \frac{1}{2}$.

Without upside and downside potential, pay performance sensitivity reduces to the benchmark sensitivity in Lemma 3. Thus, the presence of career concerns from the upside and the downside potentials changes the shape of pay performance sensitivity. This also suggests that explicit incentives without considering a manager's implicit incentives may not induce a desirable action.

2.3.3 Repeated Matching Equilibrium

In this subsection, I find an optimal one-to-one matching between firms and managers.¹¹ Recall that the goal of this paper is to analyze the interaction between labor market for managers and managers' career concerns. Thus, the main analysis is a rematching equilibrium in period 1. I first describe reputation updating by Bayes' rule. Then I provide the main analysis of period 1 in tandem.

Reputation Update Depending on Performance Outcomes

Due to an optimal contract in Section 2.3.2, every matched manager exerts effort to acquire a signal before they select a project. Upon a performance outcome in period 1 with the current

¹¹In principle, a firm hires more than one manager, however I focus on one-to-one matching to highlight the main trade off between match efficiency and agency costs.

reputation γ , the principal and the market will update the manager's reputation as follows.¹²

$$Pr(\tau = G|h) = \frac{Pr(\tau = G)Pr(h|\tau = G)}{Pr(\tau = G)Pr(h|\tau = G) + Pr(\tau = B)Pr(h|\tau = B)} = \frac{1}{1 + \frac{1-\gamma}{\gamma} \frac{\alpha\beta p}{\alpha}}$$

$$Pr(\tau = G|l) = \frac{Pr(\tau = G)Pr(l|\tau = G)}{Pr(\tau = G)Pr(l|\tau = G) + Pr(\tau = B)Pr(l|\tau = B)} = \frac{1}{1 + \frac{1-\gamma}{\gamma} \frac{\alpha\beta(1-p) + (1-\alpha)(1-\beta)}{\alpha(1-p)}}$$

$$Pr(\tau = G|m) = \frac{Pr(\tau = G)Pr(m|\tau = G)}{Pr(\tau = G)Pr(m|\tau = G) + Pr(\tau = B)Pr(m|\tau = B)} = \frac{1}{1 + \frac{1-\gamma}{\gamma} \frac{(1-\alpha)\beta + \alpha(1-\beta)}{1-\alpha}}$$

Performance h always helps manager γ improve his reputation since $\beta < 1$. Due to the assumption of $\alpha = 1/2$, performance m maintains a manager's reputation. How l changes the reputation depends on the parameter values. The natural tendency is that l tarnishes the reputation. To capture a manager's concern for their downside, assume that l is sufficiently bad that even performance of h and l leads to reputation lower than two m s. This is summarized in the following assumption.

Assumption 2. $\frac{Pr(h|\tau=B)}{Pr(h|\tau=G)} \times \frac{Pr(l|\tau=B)}{Pr(l|\tau=G)} > \left(\frac{Pr(m|\tau=B)}{Pr(m|\tau=G)}\right)^2 \Leftrightarrow \beta(1 - \beta p) > 1 - p$

Depending on the matching outcome in period 1, $\{\gamma\}_{t=2}$ differs. For instance, if all γ^h and γ^m managers are matched and exert effort to select a project according to their signals, then there will be five tiers with 7 histories: $\gamma^{hh} > \gamma^{hm} = \gamma^{mh} > \gamma^{mm} > \gamma^{hl} > \gamma^{ml} = \gamma^l$. Or, if all γ^h managers are matched, but not all γ^m managers are matched (instead γ^l managers replace the unmatched γ^m), then there will be six tiers with 10 histories: $\gamma^{hh} > \gamma^{hm} = \gamma^{mh} > \gamma^{mm} = \gamma^m > \gamma^{hl} = \gamma^{lh} > \gamma^{ml} = \gamma^l > \gamma^{ll}$. The updated perception changes the likelihood of the manager being a good type. Then, the demand for rematching arises at the end of period 1.

¹²It is worth pointing out the difference between MacDonald (1982) and this paper in accumulating information. In MacDonald (1982), the extra signal that helps agents update their type is exogenously given. On the other hand, in this paper, the information that helps agents update their type comes from the task outcome, which in turn is influenced by the manager's endogenous choice.

As a benchmark, I first find a rematching equilibrium in period 2 (i.e., when there is no career concern).

Period 2 matching: Benchmark Matching Patterns Without Career Concerns

To analyze the interaction between matching efficiency and managerial career concerns, I first discuss period 2 rematching when managers have no career concerns. Characterizing equilibrium matching patterns considers the (IR) and (IC) constraints within a match and the (SC), (PC-firm) and (PC-manager) constraints across matches. Since a firm's outside option, y_0 , is sufficiently small; there are more managers than firms; all the bargaining power is given to firms, thus every firm hires a manager in equilibrium and so the (PC-firm) does not bind even for firm S_{min} . Moreover, the (IR) constraint is not critical in determining matching because it simply guarantees that a manager's expected payoff from joining a particular firm is at least greater than or equal to his market value (i.e., a manager's outside option) and managers are indifferent from switching the current match to the other if the expected payoff is the same. Lastly, a manager's payoff is only determined by his wage at the end of period 2 as this is the last period. Thus, as long as the expected payoff is greater than or equal to w_0 , a manager wants to join a firm. i.e., The participation constraints for managers do not bind for all managers but the lowest matched managers. Therefore, the remaining two constraints are the key in determining threshold firms and the market value for managers: the (IC) constraint within a match and the (SC) constraint across matches. Given Assumption 1, the (IC) constraint induces the signal acquisition effort from matched manager, which determines the agency costs for each manager. Taking into account the agency costs, the (SC) constraint, on the other hand, determines threshold firms and

the market values for managers.¹³

When there are no career concerns left, an efficient rematching equilibrium maximizes the total match surplus $\int Y_2(S, \mu(S))dG(S)$ subject to the (SC) constraints that determine $\omega_2(\gamma)$ with w_0 for the lowest matched managers.

$$Y_2(S, \gamma) - E[w_2^X | \omega_2(\gamma)] \geq Y_2(S, \gamma') - E[w_2^X | \omega_2(\gamma')] \quad \forall S, \gamma, \gamma' \in \{\gamma\}_{t=2} \quad (\text{SC}(S, \gamma))$$

where w_2^X , $X \in \{h, m, l\}$ is determined in Section 3.2.

If a firm's willingness to pay for manager γ covers the agency cost, then manager γ , if matched, will exert effort and select the project according to their signals. The match efficiency requires that all high reputation managers get matched until the market clears and the rest of the managers at the bottom are unmatched. Due to Lemma 2, the managers' market value in period 2 for each reputation tier is characterized until every firm is matched. Table 2.1 summarizes the competitively determined compensation for each reputation level. In an efficient equilibrium, the period 2 rematching pattern would look as if it takes as follows. All manager γ^{hh} are assigned to the largest firms $S \in [S[hh], \infty)$. Starting from the right below $S[hh]$, the next largest firms $S \in [S[hm], S[hh])$ are matched with the second top tier manager γ^{hm} until the measure of $M(\gamma^{hm}) - M(\gamma^{hh})$ firms are matched. Lemma 5 summarizes this discussion.

Lemma 5. *The rematching equilibrium in period 2 exhibits positive assortativity. The labor market clears from the top, and there is no matched manager whose reputation is smaller than any of the unmatched managers. The competitive market value is determined by Lemma 2.*

¹³The market value for the manager can be interpreted as a splitting rule of the match surplus between a firm and a manager.

Rank-order: $\gamma^{hh} > \gamma^{hm} = \gamma^{mh} > \gamma^{mm} > \gamma^{hl} > \gamma^{ml} = \gamma^l$

Reputation	Market value, $\omega_2(\gamma)$	Reputation	Market value, $\omega_2(\gamma)$
γ^{hh}	$(\gamma^{hh} - \gamma^{hm}) \cdot F \cdot S[hh] + \omega_2(\gamma^{hm})$	γ^{hl}	$(\gamma^{hl} - \gamma^m) \cdot F \cdot S[hl] + \omega_2(\gamma^{ml})$
γ^{hm}	$(\gamma^{hm} - \gamma^{mm}) \cdot F \cdot S[hm] + \omega_2(\gamma^{mm})$	γ^{ml}	w_0
γ^{mm}	$(\gamma^{mm} - \gamma^{hl}) \cdot F \cdot S[mm] + \omega_2(\gamma^{hl})$	γ^l	w_0

Table 2.1: Reputation based Market Value of managers

This table summarizes competitively determined market values for managers based on their reputation where $F = (1 - \beta)(Pr(h)(h - l) + (Pr(m) - Pr(v = r))(m - l))$ in case that $\forall \gamma^h, \gamma^m$ are matched in period 1.

Period 1 matching: Match Efficiency and Career Concerns

As in period 2, the two constraints are not important in determining matching in period 1: the (PC-firm) is not critical given that the market clears; neither is the (IR) as it simply guarantees the expected wage at least a manager's outside option. The (SC) is still the key in finding threshold firms. Contrary to period 2, however, the (PC-manager) is as important as the (IC) due to period 2 matching. To see this, recall the two constraints.

$$-c + \sum_{X \in \{h, m, l\}} Pr(X|\gamma)(w_1^X + \omega_2(\gamma^X)) \geq w_1^m + \omega_2(\gamma^m) \quad (\text{IC})$$

$$-c + \sum_{X \in \{h, m, l\}} Pr(X|\gamma)(w_1^X + \omega_2(\gamma^X)) \geq w_0 + \omega_2(\gamma) \quad (\text{PC-manager})$$

The left hand sides are the same each other. The right hand sides, on the other hand, share the same market value in period 2 as $\omega_2(\gamma^m) = \omega_2(\gamma)$. If the (IC) binds (i.e., it is costly to upset manager γ 's preference for I_s without effort), then manager γ has incentive to maintain his current

reputation, which will be also possible in case that he sits out the market. The following lemma shows that the matching in period 1 can be characterized by focusing on the (PC-manager).

Lemma 6. *Suppose there are more manager γ than firms that are willing to match with them. If the incentive compatibility constraint for manager γ binds, then so does the manager's participation constraint.*

To see how future compensation influences manager γ^i 's incentives to take I_r ex ante, it is convenient to rearrange the manager's participation constraint in period 1 to derive a reputation threshold above which a manager has incentive to take I_r by acquiring information through his effort.

$$R^*(i) \equiv \frac{1}{1 - \beta} \frac{D(i)}{U(i)} - \frac{\beta}{1 - \beta}$$

where $U(i) = p(\omega_2(\gamma^{ih}) - \omega_2(\gamma^{il}))$, $D(i) = \omega_2(\gamma^{im}) - \omega_2(\gamma^{il})$. The derivation of the reputation threshold is released to the appendix A.2.3. $U(i)$ represents a period 2 compensation difference between period 1 performance of h and l , and $D(i)$ captures a period 2 compensation difference between period 1 performance of m and l . For example, manager γ^i prefers to take I_r if his current reputation is greater than his reputation threshold $R^*(i)$. Observe that $R^*(i)$ increases in $D(i)$ and decreases in $U(i)$. That is, high upside potential $U(i)$ lowers $R^*(i)$, and high downside potential increases $R^*(i)$. Thus depending on the relative magnitude of upside and downside potentials, and current reputation, manager γ^i 's risk-taking incentive differs.

To see how the matching (threshold firms and the market values) is determined, consider manager γ^i with preference for I_s (i.e., career concerns) which leads to $R^*(i) > \gamma^i$. This implies that extra rents are necessary for manager γ^i to provide incentive to take I_r . However, the extra payoff must satisfy the (SC) constraints, i.e., it is incentive compatible for firms to pay extra

payoff to their equilibrium matched manager rather than switching their matching partners. Since the market value for manager γ^i is determined by the indifference condition of the threshold firm $S[i]$ that is matched with manager γ^i , the extra payoff for manager γ^i is no longer incentive compatible for firm $S[i]$. Instead, it shall be shifted toward a firm larger than firm $S[i]$ to increase the size of a threshold firm, thereby increasing the market value for manager γ^i . Lastly, to make this shift stable in a market equilibrium, no manager γ^i has incentive to deviate from this shift, i.e., those unmatched γ^i , if any, must be indifferent from participation and sitting out. This stability requirement makes the (PC) for manager γ^i bind, which will not happen in period 2 as long as manager γ^i is not the lowest reputation manager.

To summarize the key mechanism, manager γ^i 's career concerns (induced preference for I_s due to period 2 matching) requires increases in current compensation to mitigate career concerns. The increases in current compensation shall satisfy the (SC) constraints, which shifts the smallest firm size to increase the current market value for manager γ^i . Since not all manager γ^i may get matched due to this shift, the stability requires that, between participation and waiting for the next period, manager γ^i is indifferent, thereby creating the option value of sitting out. The following proposition summarizes this discussion.

Proposition 1. *Let $\omega_1^{SC}(\gamma^i, S[i]), \omega_1^{PC}$ denote the period 1 market value for manager γ^i that is characterized by a threshold firm $S[i]$'s sorting constraint, and manager γ^i 's participation constraint, and let $\omega_1^*(\gamma^i)$ denote an equilibrium market value. Then, in period 1, for all $\gamma^i \neq \gamma^l$, $\omega_1^*(\gamma^i) = \omega_1^{SC}(\gamma^i, S[i]) \geq \omega_1^{PC}$ and the equality holds for $S[i] > G^{-1}(1 - M(\gamma^i))$.*

In what follows, I will investigate matching patterns to derive conditions where positive assortativity may or may not arise as an equilibrium pattern. Since the key trade-off and economic

tension come from cross-sectional comparison between match efficiency and career concerns, and the parameters for investment projects are independent and identical across firms, the focus here is on the conditions with respect to the distributions of firm size and aggregate economy size.

Positive Assortative Matching: γ^h, γ^m **matched**, γ^l **unmatched** The standard prediction of most existing matching models with productive complementarity is positive assortative. Any deviation from this can be improved by rematching firms and managers assortatively. Even with agency costs due to career concerns, if the match efficiency loss of any firms is strictly greater than agency costs, then an equilibrium matching pattern shall be as follows: all γ^h managers are assigned to $\forall S \in [S[h], \infty)$, and all managers γ^m to the rest of the firms. The following proposition demonstrates this.

Proposition 2. (*Assortative Matching*) *The matching pattern is positive assortative if match efficiency losses are too big: every manager γ^h is matched with the large firms, every manager γ^m is matched with the rest of the firms, and γ^l remains unmatched, i.e., there is no matched manager whose reputation is less than any of the unmatched manager in the large economy.*

Failure of Assortativity and Option Value of Waiting Productive complementarity prefers a positive assortative matching pattern where the high reputation is matched with large firms and the low reputation is matched with small firms. This assortative matching pattern can change as the agency frictions due to career concerns become larger. The source of friction comes from the option value of maintaining the current reputation. The option value depends on how a manager can be valued in period 2. Clearly, the positive option value is only available for manager γ^h or

γ^m as manager γ^l has no incentive to maintain its lowest reputation. Thus, I consider two cases: option value for manager γ^m and for manager γ^h . To explore the impact of firm size distributions on managers' career concerns, and potential distortions in matching patterns, I derive sufficient conditions for such distortions with respect to future wage, and then provide implications for distributions of firm size.

Some γ^l replaces γ^m : Distortion at the middle First, consider the case where manager γ^m has positive option value of maintaining his current reputation γ^m . This implies that the (IC) binds, thus the market value must increase to provide incentives for information acquisition. The extra pay for information acquisition itself does not mean that some firms will deviate to match with manager γ^m . For those firms to find it profitable to match with γ^l instead of γ^m , the extra pay has to be greater than the incremental marginal benefit of matching with manager γ^m relative to manager γ^l .

Lemma 7. *Let $S[m]$ denote an equilibrium threshold firm for γ^m . In period 1, manager γ^m 's career concerns lead to $S[m] > S_{min} = G^{-1}(1 - M(\gamma^m))$ if $Pr(h|\gamma^m) \times U(m) + \omega_1^{SC}(\gamma^m, S_{min}) < Pr(l|\gamma^m) \times D(m)$.*

Where $U(m) = S[mh] \times F(hm)$, $D(m) = S[mm] \times F(mm)$ denote an upside potential and downside potential respectively, and $S[i], F(i)$ denote an equilibrium threshold firm for γ^i and expected project return that is contributed to manager γ^i . I now characterize an equilibrium that exhibits a hole at the reputation γ^m . Let $U(h), D(h)$ denote an upside and downside potential for manager γ^h , where $U(h) = S[hh] \times F(hh)$, $D(h) = (S[hm]F(hm) + S[mm]F(mm))$.

Proposition 3. (Failure of Assortativity at the Middle) *There exists a stable matching where some*

γ^m get unmatched if $Pr(h|\gamma^h) \times U(h) + \omega_1(\gamma^h) \geq Pr(l|\gamma^h) \times D(h)$ and Lemma 7 holds: every manager γ^h is matched with the large firms $S \in [S[h], \infty)$, some manager γ^m is matched with $S \in [S[m], S[hh])$ where $S[m] > S_{min}$, all remaining firms $S \in [S_{min}, S[m])$ are matched with γ^l . The unmatched is indifferent from participation and sitting out.

Implications for distributions of firm size The characteristic of the sufficient condition is that 1) there are sufficiently large threshold firms ($S[hh]$) that determine the market value for γ^{hh} , and 2) the sizes of other smaller threshold firms, ($S[hm], S[mm]$) are not large enough. Observe that the presence of large firms is not sufficient to obtain the above equilibrium. This is because even if there exist large firms $S[hh]$, if firm $S[hm]$ (the matched firm size in case that manager γ^h maintains his reputation) is also large, the manager γ^h is less willing to take the I_r . Moreover, too many large firms (i.e., firm $S[hm]$ large enough) do not generate the above result because manager γ^m may have incentive to take the I_r . Thus, it is the distribution of firms that matters. In particular, the curvature at the top of the firm size distribution that is very convex supports the above equilibrium.

Proposition 4. *The hole at the middle equilibrium is supported by a distribution of firm size that indicates that firm size increases faster (i.e., its slope) as size increases.*

Some γ^m replaces γ^h : Distortion at the top Next, consider the case where manager γ^h has positive option value of maintaining his current reputation γ^h . This implies that the manager's (IC) binds, thus the market value for manager γ^h has to increase to provide incentives for information acquisition. Similar with manager γ^m case above, the extra pay for information acquisition does not mean that some firms will deviate to match with manager γ^h yet. For those

firms to find it profitable to match with γ^m instead of γ^h , the extra pay has to be greater than the incremental marginal benefit of matching with manager γ^h relative to manager γ^m .

Lemma 8. *Let $S[h]$ denote an equilibrium threshold firm for γ^h . In period 1, manager γ^h 's career concerns lead to $S[h] > G^{-1}(1 - M(\gamma^h))$ if $Pr(h|\gamma^h) \times U(h) + \omega_1(\gamma^h) < Pr(l|\gamma^h)D(h)$.*

where $U(h) = S[hh] \times F(hh)$, $D(h) = \left(S[hm]F(hm) + S[mm]F(mm) \right)$. The following proposition characterizes an equilibrium that exhibits a hole at the reputation γ^h .

Proposition 5. *(Failure of Assortativity at the Top) There exists a stable matching where some γ^h get unmatched if $Pr(h|\gamma^m) \times U(m) + \omega_1^{SC}(\gamma^m, S_{min}) \geq Pr(l|\gamma^m) \times D(m)$ and Lemma 8 holds : some manager γ^h is replaced by γ^m to be matched with the large firms, and the rest of firms are matched with manager γ^m and γ^l until every firm is matched with a manager. The unmatched γ^h is indifferent from participation and sitting out.*

The characteristic of the sufficient condition is that 1) the threshold firm ($S[hh]$) that determines the market value for γ^h is not large enough relative to $S[hm]$, and 2) the threshold firm $S[hm]$ is relatively larger than $S[mm]$. Similar with the previous case, the presence of large firms is not sufficient to obtain the above equilibrium. In particular, if firm $S[hm]$ is large enough, then manager γ^h 's incentive to keep his reputation becomes larger, thus the manager becomes less willing to take the I_r . However, given that firm $S[hm]$ is big enough, firm $S[mm]$ is sufficiently small so as to create manager γ^m 's risk-taking incentive. This suggests that the curvature at the middle of the distribution of firm size that is very convex supports the above equilibrium.

To summarize, the intuition for the above two propositions relies on the trade-off between matching efficiency and the required pay. The required pay is driven by the interaction between managers' career concerns and the firms' competition for managerial talent. Due to comple-

mentarity between firm size and managerial talent, the market value for managers is a function of threshold firm size and manager's reputation. This suggests that the distribution of firm size influences the distribution of market wage for managers' reputation. Managers' concerns about the next period market wages in turn create induced preferences for risk exposure, which differs across managers depending on the shape of the firm size distribution (thus wage distribution). The market for managers sorts them by their perceived talent (reputation) whenever the manager's performance is available. However, the results imply that the interaction between the market for managers and agency conflicts do not always guarantee economic efficiency in matching outcomes.

2.4 Applications

The baseline model that I proposed in this paper can be applied to various economic problems including matching of lender and borrower, analyst and firm, auditor and client, and board of directors and firm. Introducing variations into the agency problem and/or the matching problem depending on a particular conflict or friction of interest can provide a framework that enables to analyze the interactions between the market forces and agency problems. The agency costs and the potential distortions in matching patterns that this paper has focused on can be applied to exploring match performance impact on rematch of firms and manager. In particular, how strong past performance is associated with a firm's (or a manager's) decision of match dissolution can be one direct application of the baseline model and its trade-off in this paper.

In addition to this direct application of managerial turnover-performance sensitivity, one can also analyze the audit labor market effect on aggregate audit quality in the context of matching

of auditors and clients. It is reasonable to assume that talented auditors are capable of auditing complex and large transactions, which can be interpreted as productive complementarity between auditor reputation and firm size (and/or transaction complexity). Given auditors' concern about their reputation (perceived talent), how the auditor labor market influences which firm matches with which auditor, and how the auditor market affects aggregate audit quality are important questions to understand. By similar logic, the baseline model can be also applied to analysts and their coverage. The productive complementarity between an analyst's talent in forecasting and firm size (with presumption that large firms being difficult to forecast) is easily justified. Also, the labor market for analysts is clearly driven by their track record that forms their reputation, and better reputation analysts perform better (Stickel (1992)). Given this one-side concern about their reputation and productive complementarity between the two-sides, one can analyze the formation of analysts and coverage firms, and based on such formation, how aggregate forecast accuracy or recommendation quality changes. Other potential applications include matching of lenders and borrowers, and a board of directors and firms.

Based on its direct implication, I explain in more detail about how to apply the baseline model to CEO-firm match to understand the well-known, yet puzzling evidence of weak turnover performance sensitivity.

2.4.1 CEO Turnover and Firm Performance

One of the central roles of corporate boards is to replace or retain their CEOs. While firm performance is negatively related to CEO turnover, it has been extensively documented that this negative association is economically small (e.g., Murphy (1999), Brickley (2003), Larcker and

Tayan (2015)).¹⁴ These findings have been rationalized as arising from weak internal monitoring mechanisms resulting from flawed governance structures (Hermalin and Weisbach (1998), Taylor (2010)). However, recent substantial changes in corporate governance have not altered the association between performance and CEO turnover (Huson et al. (2001), Bhagat and Bolton (2008), Kaplan and Minton (2012)).¹⁵ However, this empirical regularity can be explained by the distortions in matching patterns. To put roughly, a dynamic labor market competition for CEO talent yields a non-monotonic association between a CEO's perceived talent and a CEO's preference for risk exposure. This in turn, distorts a firm's preference for CEO talent, thereby weakening the observed association between firm performance and CEO turnover.

To see this, consider the baseline setup. But to analyze the impact of past performance and career concerns on a firm's turnover decision, suppose that the repeated matching economy lasts for three periods and that initially every CEO is endowed with the identical reputation γ . Since everyone is identical ex ante, without loss of generality, assume that the period 1 matching is random. Other than these modification, the structure of the game is identical with the baseline setup. In the model, we have started from assuming that there are managers of $\gamma^h, \gamma^m, \gamma^l$ in period 1. Manager γ^i can be interpreted as a manager whose initial period performance is i , and

¹⁴The association between performance and CEO turnover has been largely studied and the minor impact of firm performance has been well-known. See Coughlan and Schmidt (1985), Warner et al. (1988), Jensen and Murphy (1990), Puffer and Weintrop (1991), Murphy and Zimmerman (1993), Jenter and Lewellen (2010), and Dikolli et al. (2014).

¹⁵Both Bhagat and Bolton (2008) and Kaplan and Minton (2012) find evidence that only board independence increases turnover sensitivity to a certain performance measure (e.g., industry adjusted stock return in Kaplan and Minton). However, the change in other proxies for governance quality, including CEO-Chair duality and governance indices, do not change turnover-performance sensitivity.

the period 1 matching game in the baseline setup is interpreted as a firm's decisions of replacing or retaining the incumbent CEO, and as a CEO's decision of whether to remaining or leaving the existing employer firm (or the market) this period.

Before discussing the results, it is worth noting that the matching outcome in the model can identify only involuntary turnover events: CEOs leaving for other firms will not be identified, but only those who remain unmatched in the market will be identified as involuntary turnover. This feature of identification is also consistent with empirical research. Often the challenge in the literature is to distinguish involuntary CEO turnover events from voluntary ones. Thus, in the event of CEO departure, only those departing CEOs who do not have a next job are classified as forced turnover (Parrino (1997)). However, this identification invites some caution in interpretation. Because, in the model, it is incentive compatible for those unmatched better reputation managers (either γ^h or γ^m) to sit out, a departing CEO without a job does not have to be the outcome of involuntary turnover. That is, those departing CEOs without a job may not involuntarily step down. This *misclassified* forced turnover at the better performance levels may weaken the association between turnover and performance. Therefore, for a better identification of turnover, one needs to expand the career horizon of departing CEOs or to collect more information regarding the turnover events.¹⁶

Even with this caution, the findings, Proposition 1 to 4, provide implications for CEO turnover performance sensitivity. First, under the conditions described in Proposition 2, all γ^h and γ^m managers get matched, and all γ^l managers are unmatched. This suggests that performance l

¹⁶Kaplan and Minton (2012) also discuss a similar argument that those departing CEOs that are classified as voluntary turnover (at poor performance level) may not be voluntary. The results of misclassification may also lead to a weak association between turnover and performance.

always leads to turnover, which generates strong impact of performance in predicting turnover. Since this strong association is not empirically observed, to test this model, the firm size conditions described in Proposition 2 should be empirically falsified on average. Or, in a country or industry where the distributional properties in Proposition 2 are observed, the model predicts that the turnover pattern should be strongly associated with performance.

Indeed, the empirically well-documented firm size distribution exhibits a power law (Ijiri and Simon (1977), Axtell (2001), Gabaix (2016)), which is related to the Proposition 3. In Proposition 3, all γ^h , some γ^m managers get matched, but the other γ^m managers sit out. Instead some γ^l managers will get matched. Thus, the association between performance l and turnover is weakened by the measure of matched manager γ^l s. Moreover, the association between performance m and turnover increases by amount of the measure of manager γ^m s sitting out the matching market (recall, in previous case, the relation between m and turnover is zero). Overall, the lack of turnover in performance l and the excess turnover in performance m jointly weakens the impact of performance in predicting turnover, which is consistent with empirical regularity. Interestingly, the sufficient conditions that derives such matching pattern in Proposition indicate a characteristic where there are large number of small and medium firms and small number of large firms, which is satisfied by a power law. Since the conditions with respect to firm size is only a sufficient condition, Proposition 3 does not conclude the distribution in the real world to be a power law distribution. However, at least, the model confirms that, once the empirically well supported distribution is considered, it will generate empirically well documented turnover performance relation.

Lastly, in Proposition 4, all γ^m managers get matched, but only some γ^h get matched. i.e.,

some γ^h managers sit out due to its option value. Thus, performance h predicts some turnover events, while performance m faces no turnover, and performance l faces turnover. This U-shaped turnover performance relation is not empirically observed (at least in the U.S.). Neither is the distributional properties of firm size in Proposition 4. However, what the baseline model predicts is that under the economy where there are not many large firms, and firm size is almost homogeneous, the cross-sectional turnover-performance pattern will be close to U-shaped, and there will be turnover at the top performance.

Overall, depending on the distributions of firm size (i.e., size of economy), the model predicts that the relation between CEO turnover and performance differs, either strong, weak, or U-shaped. However, across all these three propositions, the measure of unmatched CEOs is the same. This feature allows one to compare the shape of turnover performance relation and to test the predictions depending on firm size distributions. In particular, by splitting turnover data into industries or groups of similar firm size distributions, or comparing different countries with different shapes of firm size distributions, the baseline model that highlights the impact of interplay between the labor market and managerial career concerns can be tested.

2.4.2 Discussion

This section discusses some interpretations of the key trade-off in this model. The trade-off between match efficiency versus the career concern driven agency costs in a firm-manager matching game can be naturally captured by a general two sided matching game (say A and B) where players B can choose whether to participate in a matching game that involves some kind of observable experiment. The experiment technology is not match specific, but rather solely based

on player B 's choice, and the result of experiment changes the characteristics of player B . The experimental results are directly related to match output with two players being productive complements, players divide the match output within a match with the presence of outside option, and matching game is repeated. Given that players A have sole bargaining power, the equilibrium splitting rule of match output is determined by player A 's willingness to pay to match, and player B 's willingness to participate and experiment. Since player B will get better terms in case that many players A compete for B , naturally player B wants to look better. The incentive of looking better arises only if player B expects that it will increase (or not decrease) their value in the future. Thereby, in an interim period, player B may decide not to participate if option value of maintaining their characteristic is high. Whether offering extra payoff to attract certain characteristic of player B depends on player A 's willingness, which will depend on their match efficiency. If the marginal benefit for player A is not big enough relative to marginal cost, then player A will seek to match with the next best player B , which goes against assortativity.

The purpose of introducing agency friction is to explore the formation of firms and managers and to provide insights into managerial accounting and control. The agency friction through costly actions and/or misalignment of a firm's and manager's incentives is one of the most elemental issues in organizations. Which contract is offered and how internal control mechanisms are designed would look like an outcome of an individual firm and manager's problem. However, the formation of firms and managers is closely tied to the labor market for managers, thus the endogeneity of matching is critical in understanding such contracting outcomes. Including firm and/or manager fixed effects could partially mitigate potential issues arising from endogeneity. But considering the matching market explicitly (e.g., a distribution of firm size, supply of

managers) will allow researchers to better control endogeneity concerns as well as better predict internal control mechanisms.

2.5 Conclusion

I conclude this paper by pointing out a limitation of the modeling assumption about the fixed characteristics of firm size. The research question I try to answer is how managerial career concerns interact with the formation of firms and managers. To address the question by deriving a manager's career concerns endogenously, I focus on a manager's future compensation, a function of a manager's reputation and matched firm size. Since a manager's reputation evolves, thereby making the manager's future compensation evolve, I assumed away *the evolution of firm size*. The match output by a particular firm size and a manager's reputation can potentially change the size characteristic of a firm. For example, by taking a positive NPV risky project, a firm's positive (negative) earnings can make the firm size bigger (smaller). To focus on the trade-off between match efficiency and agency costs, I have assumed that the match output is consumed by the matched firm and manager in current period, thus no impact on the growth of firm size. The stationarity in firm size distributions would partially justify the assumption in this paper, however, the growth aspect due to match output will create different incentive for firms in deciding their willingness to match and pay.

As an extension, one could potentially incorporate a richer framework by considering the evolution of firm size depending on the match outcome. This is clearly affected by a characteristic of a matched manager that the firm has had, and now a firm also has future concerns due to the impact of the current match on their size growth. To pursue this extension, one needs to address

several issues, including the magnitude of growth or decline. For instance, the growth potential may exhibit a scale of operations as in project return. In that case, a large firm's growth (decline) is bigger than a small firm's due to success (failure), leading to either a heavier competition toward managerial talent at the top or a large firm's strategic choice of safe project to maintain their size. If the former (the latter) is dominant, a firm's growth potential can make match efficiency bigger (smaller) than agency costs. Considering these issues might allow one to develop a better understanding of interactions between agency problems and the market for managers. While it invites many directions for extension, the repeated matching model with moral hazard that I propose will be a useful framework to examine the labor market forces in managerial accounting research.

Chapter 3

Career Concerns and Project Selection:

The Role of Reputation Insurance

Abstract

Motivated by the Chapter 1, I explore, in the context of CEO turnover, the latent aspects of existing corporate governance practices, including a board of directors, performance disclosure policy, and severance pay. In particular, I ask how well different governance practices provide incentives for project selection when managers have career concerns and how such practices influence a firm's decision of whether to replace their CEO. I show that a board of directors monitoring, performance disclosure policy, and a severance package serve as reputation insurance and mitigate a manager's career concerns through different mechanisms. However, the incentive effects of reputation insurance are followed by a weakened relation of turnover to performance: the board's monitoring serves as a substitute for performance; the non-disclosure of a CEOs performance at departure causes misclassification errors ; the presence of severance pay, on the other hand, creates performance tolerance for firms in order not to pay out. Based on the perspective of reputation insurance, I also provide empirical predictions related to the existing governance and CEO turnover practices.

3.1 Introduction

Chapter 1 considers how career concerns are derived by the market for managers and the associated inefficiency in the formation of firms and managers: Managerial career concerns influence agency problems, thus creating distortions in matching decisions; particularly, in the context of a firm-CEO match, it even distorts a firm's decision of whether to replace or retain its CEO. The natural follow-up question is whether there exist any corporate governance mechanisms that attempt to reduce such inefficiency. In this chapter, I take up this question in the context of a firm-CEO match. I argue that existing institutions can serve as insurance for a manager to resolve potential agency frictions arising from a manager's career concerns. The goal here is to compare different mechanisms of corporate governance as insurance for managers and to derive empirical implications for CEO turnover practices.

Building on the baseline model in Chapter 1, I present a series of governance models where a career concerned CEO needs to exert effort for information acquisition and takes a project, either the safe or the risky. The first model explores a board of directors in which directors monitor a CEO for the purpose of talent evaluation. The second model investigates a firm's disclosure policy of a CEO's performance. The third model studies ex ante severance pay agreement. I show that these governance practices—strengthening monitoring, decreasing disclosure, and designing severance pay that is less tied to performance—serve as reputation and reduce potential frictions due to career concerns, thus arising as part of optimal contracting devices.

In the first model, the firm appoints a board of directors at costs. The costs depend on the monitoring quality of a board: the better quality the more expensive. When a firm decides

whether to replace the incumbent CEO, the board acquires information about the incumbent CEO's talent to make a better replacement decision. I show that the board's monitoring can function as insurance for the CEO by supporting him when there is a poor performance outcome. The intuition is because the board's information about the CEO's talent deters the reputation reduction at poor performance, which in turn, incentivizes the CEO ex ante to make an efficient project decision that is potentially risky. Therefore, the expected future monitoring (thus, better information about a manager's talent) mitigates career concerns and reduces agency conflicts.

However, the board's better monitoring can potentially weaken the impact of performance on a firm's turnover decision in the following period. When the board's monitoring becomes better, they acquire more precise information which is potentially superior to noisy performance outcomes. Consequently, the board relies more on their own information rather than the realized performance, thereby weakening the impact of performance on their turnover decision. This result suggests that weak turnover-performance sensitivity might not indicate weak corporate governance, but instead can be an optimal decision made by a board of directors (Fisman et al. (2013), Laux (2015)). The model predicts that the association between turnover and performance depends on the board's monitoring quality; the better monitoring leads to more (ex ante profitable) corporate risk taking behaviors. Since the benefit of mitigating career concerns (for match efficiency) is bigger for large firms, the model also predicts that large firms tend to rely on better monitoring as insurance for their executives as opposed to career concern-driven inefficient turnover.

In the model of performance disclosure, the firm promises not to disclose a realized performance at the CEO's departure. If a firm's replacement decision of its incumbent CEO does not

fully reveal the actual performance, say medium performance of the safe choice and the poor performance of the risky choice, then the market forms an expectation over the departing CEO's reputation. That non-disclosure does not fully reveal the realized performance is only for large firms that can replace medium and poor performing CEOs to high performers. In such a case, a firm's non-disclosure not only increases the incentive to take the risky project, but also reduces the CEO's preference for the safe one, thus efficiently providing incentive for risk taking. Consequently, the promise of non-disclosure serves as reputation insurance, thus mitigating the CEO's career concerns.

As in the board's monitoring, a firm's disclosure policy as insurance weakens the association between turnover and performance. The intuition here is that, although the realized performance is known inside the firm, the market cannot observe which performance leads to a CEO departure (i.e., information suppression). That is, an individual firm's disclosure policy for the sake of insurance creates negative externality in sorting in the market. In principle, this requires one to speculate root causes of turnover. This speculation, by its nature, is subject to classification error, which potentially weakens the association between turnover and performance. The model predicts that the more opaque firms are in disclosing performance, the more managers will take risky investments. Since performance non-disclosure serves as insurance only for large firms, large firms tend to be more vague when it comes to disclosure of departing executives' performance than small firms.

Lastly, in the third model, the firm promises to pay an ex ante severance pay agreement at the CEO's departure. The severance pay has a direct interpretation as insurance payment at the CEO dismissal, which involves his reputation reduction. Although the CEO's market wage following

poor performance will decrease, the severance pay at departure substitutes for the reduction of his market wage. Thus, the expected future severance as insurance in turn incentivizes the CEO to take the potentially profitable risky project. Put differently, the direct insurance payment through severance agreement mitigates managerial career concerns. Similar to the first two models of governance, the ex ante severance agreement can weaken the turnover-performance sensitivity due to performance tolerance. Although the severance agreement incentivizes the CEO ex ante, the incumbent firm may decide not to replace the CEO ex post so as to save severance pay if the benefit of replacement is not big enough.

In practice, most academics and practitioners criticize executives' severance packages for interrupting incentives (e.g., Bebchuk and Fried (2009)). However, in the model I propose here, offering severance pay that is not tied to performance can be optimal ex ante by virtue of reputation insurance. When a firm wants to encourage appropriate risk taking, providing such insurance can effectively incentivize the CEO. The empirical implications of this model are that the presence of a severance agreement tends to follow more corporate risk taking behaviors; the impact of performance on turnover becomes weaker when an executive's contract has severance arrangements; large firms tend to provide a bigger severance package than small firms.

The remainder of the chapter proceeds as follows. Since the goal of this chapter is to investigate reputation insurance aspects of governance practices, I take a manager's career concerns that are derived in Chapter 1 as given throughout this chapter. Thus, building on the model of a CEO-firm match in Chapter 1, Section 2 briefly provides the baseline setup. Then, Section 3 develops and discusses the board's monitoring as insurance. Section 4 considers the disclosure policy as insurance. Section 5 introduces a severance agreement as insurance. Since CEO turnover has

been extensively studied in the literature, I will discuss more about the existing related works in Section 6. Section 7 concludes.

3.2 Baseline Setup

Consider a firm (principal) and a CEO (agent). The game lasts for two periods. Within a period, the sequence of events is as follows: 1) the firm offers a contract with governance device (either monitoring, disclosure policy, or severance pay); 2) the CEO accepts or rejects the contract, if he rejects, the game is over, if he accepts, then he exerts effort to select an investment project; 3) the project earnings and payoffs are realized; and 4) both the firm and CEO return to the market for the next period production (if this is the last period, the game ends). All players are risk neutral and share the same horizon with no discount factor.

As in Chapter 1, the firm is characterized as its size S and the CEO is characterized as its reputation γ . The firm has the choice to invest in one of two projects: a risky project denoted as I_r or a safe project denoted as I_s . The safe project I_s will return a certain outcome, m , which can be interpreted as a status quo. The risky project will return either a success, $h > m$, or a failure, $l < m$. Without loss of generality, assume that the investment cost is the same for both projects, which is normalized to zero. The probability of success for the project I_r depends on a state variable consisting of $\{s_1, s_2, s_3\}$, where s_1 indicates I_r will generate h , s_2 and s_3 indicate I_r will generate l . The unconditional probability of each state is $Pr(s_1) = \alpha p$, $Pr(s_2) = \alpha(1 - p)$, $Pr(s_3) = 1 - \alpha$, where $\alpha \in (0, 1)$, and $p \in (0, 1)$. It is immediate to see that $Pr(h|\{s_1, s_2\}, I_r) = p$, $Pr(l|\{s_1, s_2\}, I_r) = 1 - p$, and $Pr(l|s_3, I_r) = 1$. I assume the scale of operations (Sattinger (1993)). That is, a firm's project return is the project outcome (denoted

as X) multiplied by its firm size, $S: S \times X$, where $X \in \{h, m, l\}$.

The CEO's effort is modeled as gathering information to select the project. The information (signal) that the CEO can acquire is drawn from $\{\{s_1, s_2\}, \{s_3\}\}$. For convenience, let $r = \{s_1, s_2\}$, $s = \{s_3\}$. Let $v, \hat{v} \in \{r, s\}$ denote a partition of state variables and the CEO's acquired information respectively. There are two types of CEOs, $\tau = G$ and $\tau = B$, denoting a good and a bad type respectively. The two types differ in their ability to acquire information about the realized state. By exerting effort, a good type manager knows if a realized state belongs to r or s (i.e., $\hat{v} = v$), but a bad type manager receives $\hat{v} = v$ with probability of $\beta \in (0, 1)$ and $\hat{v} \neq v$ with the complementary probability. Without effort, both types do not receive a signal. I assume that the ex ante probabilities of realization of signals is the same for both types so that the acquired signal per se does not communicate any information about the manager's type. This assumption is captured by setting $\alpha = 1/2$.¹ Assume that both the firm and the CEO does not know his type (Holmstrom (1999)).

To make the project selection problem non-trivial, assume that $ph + (1 - p)l > m > l$, i.e., if $v = r$, then I_r is more profitable, and if $v = s$, then I_s is more profitable. Let $Pr(X|\gamma)$ denote probability of project outcome X conditional on the CEO reputation γ . Let $Y(S, \gamma)$ denote the expected project return for the firm with the CEO γ . That is,

$$Y(S, \gamma) = S \times \left(h \times Pr(h|\gamma) + m \times Pr(m|\gamma) + l \times Pr(l|\gamma) \right)$$

Let w^X denote a transfer upon the project outcome $X \in \{h, m, l\}$, and $E[w^X|\omega_t(\gamma)]$ denote the expected compensation cost given the market value for the CEO, $\omega_t(\gamma)$ in period $t = 1, 2$. Then, the firm, taking the market value of each manager as given, chooses the optimal compensation

¹That is, $Pr(\hat{v}|G) = Pr(\hat{v}|B)$ for all \hat{v} , $\Leftrightarrow \alpha = \alpha\beta + (1 - \alpha)(1 - \beta)$

and a governance device, to maximize its payoff which is expected project return net of the manager's compensation.

$$\max_{w^x, \mathcal{G}} Y(S, \gamma) - E[w^X | \omega_t(\gamma), \mathcal{G}] - C(\mathcal{G})$$

where $\mathcal{G}, C(\mathcal{G})$ denote the choice of governance device and corresponding cost respectively.

3.3 Board of Directors' Monitoring as Insurance

In practice, a firm's board of directors or compensation committee evaluates their CEO based on their own appraisal policy. To capture this practice, consider a board of directors that can observe a noisy signal about a CEO's type. Here, the board of directors can be seen as a monitoring technology for a firm to provide better information about a manager. Although the board is a common feature that every organization has, their performance can differ depending on their ability. Thus, I assume that a firm can incur cost to set up the board that is good at monitoring the CEO. That is, with some probability, the principal gets a signal S_G representing a good type, and with a complementary probability, the principal receives nothing.

$$b = Pr(s_G | \tau = G) > Pr(s_G | \tau = B) = 1 - b > 0$$

Assume $b > \frac{1}{2}$. The parameter b can be considered as a board's competence in monitoring. The characteristic of this signal is closer to the extra information introduced in MacDonald (1982). To highlight the insurance role of monitoring, assume that there is no commitment problem, and the principal will disclose the observed signal truthfully if received. Note that conditional upon the CEO's type, the signal is independent of a performance outcome. To gauge how the principal's

monitoring plays a role as insurance, I assume that the monitoring technology gives a signal upon the poor outcome.² Moreover, it is assumed that the competence of monitoring is limited so that $E[\omega_3(\gamma)|e = L, I_r, b] < E[\omega_3(\gamma)|e = L, I_s, b]$ for $\forall \gamma, b$. The monitoring technology always has a positive type 1 and type 2 error. The expected reputation change with the board's extra signal is summarized in the following lemma.

Lemma 9. *Suppose that $b > \frac{1}{2}$. Let $\gamma_{b+}^l, \gamma_{b-}^l$ denote an updated reputation when a principal receives s_G and nothing upon a poor outcome. Then, the ex ante expected reputation change upon a poor outcome is characterized as follows.*

$$E[\gamma^l|b] = \left(b(Pr(l, \tau = G) - Pr(l, \tau = B)) + Pr(l, \tau = B) \right) \gamma_{b+}^l + \left(-b(Pr(l, \tau = G) - Pr(l, \tau = B)) + Pr(l, \tau = G) \right) \gamma_{b-}^l$$

$$\text{where } \gamma_{b+}^l = \frac{1}{1 + \frac{1-\gamma}{\gamma} \frac{((1-p)\beta + (1-\beta)) \frac{1-b}{b}}, \gamma_{b-}^l = \frac{1}{1 + \frac{1-\gamma}{\gamma} \frac{((1-p)\beta + (1-\beta)) \frac{b}{1-b}}.$$

The expected reputation change conditional on the realization of l is the following.

$$E[\gamma|b, l] = (\gamma b + (1 - \gamma)(1 - b)) \gamma_{b+}^l + (\gamma(1 - b) + (1 - \gamma)b) \gamma_{b-}^l$$

Albeit helpful to gauge the incumbent CEO, this monitoring technology comes at cost $C(b)$ with $C'(b) > 0, C''(b) > 0$. This cost can be interpreted as the firm's investment in developing a better measurement system for a CEO performance review or hiring competent board members who will provide a more precise evaluation about the incumbent CEO. Thus, the necessary and sufficient condition for the principal to invest in monitoring technology at the beginning of period

²A board of directors is more likely to evaluate (monitor) the incumbent CEO upon poor performance. Relaxing this assumption will have a distraction from failing to receive a signal upon h or m .

2 is,

$$Y_2(S, \gamma) - E[w_2^X | \omega_2(\gamma)] - C(b) \geq \max_{\gamma'} \left\{ Y_2(S, \gamma') - E[w_2^X | \omega_2(\gamma')] \right\} \quad (3.1)$$

The left-hand side of (3.1) is the principal's payoff from the match with CEO γ at monitoring cost $C(b)$ and the right-hand side is the principal's payoff of hiring a different CEO or retaining the incumbent but paying more to mitigate career concerns. Basically, the above condition requires that the principal finds it optimal to invest in monitoring technology to incentivize the CEO instead of hiring an alternative CEO, or paying the CEO extra premium. Lemma 10 shows the existence of b^* and its characteristics.

Lemma 10. *(Optimal Monitoring Intensity and Externality on Compensation) There exists a unique $b^* \in (\frac{1}{2}, 1)$ that mitigates the CEOs' career concerns. Introducing monitoring technology shifts the compensation for CEOs downward.*

The reasoning behind uniqueness is that the principal will choose b just enough to bind the (IC) constraint for the career-concerned CEO since the monitoring is costly. On the other hand, the intuition for the externality on compensation is that introducing the monitoring technology substitutes for extra pay.

The natural investigation is to see how this monitoring technology affects the market's reliance on firm performance to update the reputations of CEOs. Proposition 6 finds when the market prefers monitor to firm performance.

Proposition 6. *(Monitor as Substitute) As board competency, b , increases, the market prefers monitoring to performance. More formally, if the equilibrium choice of $b^* > \frac{\frac{Pr(l|\tau=G)}{Pr(l|\tau=B)}}{\frac{Pr(l|\tau=G)}{Pr(l|\tau=B)} + 1}$, then firms prefer to retain γ_{b+}^l if an alternative CEO is γ .*

Proposition 6 is also consistent with recent findings on the relation between governance

and turnover-performance sensitivity and the implications for firm performance (Fisman et al. (2013)). Fisman et al. find that the weak turnover-performance sensitivity of a weak board that protects a poor performing incumbent CEO can lead to a better subsequent performance. In my model, the weak turnover-performance sensitivity comes from a superior monitor that generates better information about the incumbent CEO's type than firm performance. It is worth pointing out the difference between Crémer (1995), Laux (2015) and this paper. In Crémer (1995) and Laux (2015), the benefit of retention (in case of a success) substitutes for monetary incentives to induce effort from the agent under a no-monitoring regime, but the presence of monitoring eliminates such substitution, thereby increasing incentive costs. However, in this paper, monitoring occurs only if a CEO takes I_r and faces a failure. Thus, the benefit of monitoring can occur only when the CEO acquires a signal (i.e., $e = H$), not when the CEO shirks (and chooses I_s).³

3.4 Performance Disclosure Policy as Insurance

This section applies the main model by introducing performance disclosure as a choice variable to show that a firm's disclosure policy on performance can act as reputation insurance, thus part of a contract. The motivation for this extension is that an exact reason or cause for departure is often not explicitly disclosed.⁴ To see how performance disclosure plays as reputation insurance for

³Since $E[\omega_3(\gamma)|e = L, I_r, b] < E[\omega_3(\gamma)|e = L, I_s, b]$ for any feasible b , a CEO's best project choice upon $e = L$ is still I_s . Thus, monitoring relaxes (IC) constraint. However, if the monitoring benefit is sufficiently high that $E[\omega_3(\gamma)|e = L, I_r, b] \geq E[\omega_3(\gamma)|e = L, I_s, b]$, then the same friction as in Crémer (1995) and Laux (2015) will appear.

⁴See, for instance, Parrino (1997), Jenter and Lewellen (2010). For an anecdotal evidence, when the Metropolitan Transportation Authority (MTA) replaced Arthur Leahy in January 2015, the MTA disclosed, "On his watch, Metro

CEOs, assume that a principal can credibly commit whether to disclose performance information to the market. To avoid the market's inference from the realized pay, assume that w_2^X is also not disclosed upon departure.

For performance non-disclosure to create reputation insurance, it shall be the case that the market cannot infer the exact performance outcome at the CEO departure without any stated reasons. Otherwise, the promise of non-disclosure does not provide any insurance benefit. Lemma 11 finds a condition in which performance non-disclosure can play as insurance.

Lemma 11. *Performance non-disclosure can be used as insurance in period 2 only if firms are sufficiently large, $S \in [S[hl], S[h]]$.*

The mechanism of non-disclosure as reputation insurance is the reduction of the downside potential by aggregating the two outcomes. By aggregating, the expected payoff of choosing the risky project increases, but the payoff of the safe project decreases. Thus, the firm relaxes the (IC) constraint effectively by simultaneously balancing the left hand side and the right hand side. There is no direct cost borne by the principals, thus no direct compensation is given to the CEOs. Hence the promise of non-disclosure makes only those firms better off. However, the non-disclosure creates rematch friction in the last period due to the aggregation of two outcomes. This aggregation can lead to a potential mismatch between firms and CEOs, thus incurring mismatch costs in period 3. Lemma 12 finds the expected mismatch costs in period 3 due to performance non-disclosure.

Lemma 12. *By relying on performance non-disclosure, the expected rematch distortion in period buses are more accessible, more punctual, and cleaner". However, the departing CEO's performance had been under confidential review by the board and was not disclosed (LATimes, Jan.2015).*

3 is

$$\frac{Pr(m|\gamma) \times Pr(l|\gamma)}{(Pr(m|\gamma) + Pr(l|\gamma))^2} \left(\int_{S \in H_1} SdG - \int_{S \in H_2} SdG \right) \frac{\partial}{\partial S} Y_3(S, \gamma)$$

where $H_1 = [S[l], S[hl]]$, $H_2 = [S[hl], S[m]]$. The rematch distortion always exists regardless of the parameter values.

Here, $\frac{Pr(m|\gamma) \times Pr(l|\gamma)}{(Pr(m|\gamma) + Pr(l|\gamma))^2}$ denotes the mismatch probability. With this probability, CEO γ^m , who is supposed to match with $S \in H_2$, is assigned to $S \in H_1$; CEO γ^l , who is supposed to match with $S \in H_1$, is assigned to $S \in H_2$. The mismatch distortion is the negative consequence of mis-assignment due to complementarity between firm size and reputation. In this case, it is the performance non-disclosure (i.e., information suppression) upon replacement that weakens the relation between performance and the turnover pattern. Contrary to the board's monitoring technology, information suppression endogenously creates period 3 match friction. Interestingly, although only large firms enjoy a benefit from the reputation insurance without incurring costs, the distortion happens to smaller firms. Proposition 7 highlights this result.

Proposition 7. *(Endogenous Match Frictions and Negative Externality) Performance non-disclosure provides reputation insurance, thus potentially mitigating a non-assortative rematching pattern in period 2. However, the consequence of match distortion always arises for small firms in period 3, $S \in [S[ND], S[m]]$.*

It is worth noting that the departing CEO with performance m ex post does not want to pool with a poor performing CEO, thus revealing his realized performance. However, given that every CEO has incentive to defend himself and protect their reputation, the departing CEO's argument about his/her undisclosed performance cannot be credible to the market. Therefore, the market still takes the expected value of reputation over performance of m and l , and the departing CEO's

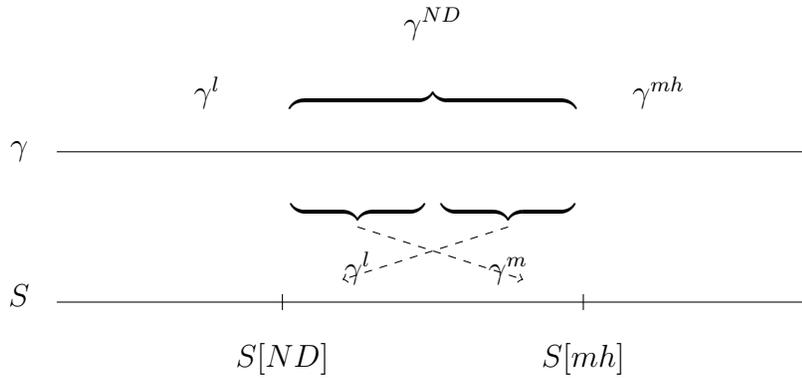


Figure 3.1: **Potential mis-assignment of CEOs and Firms due to non-disclosure in period 2.**

actual performance will not be revealed.

3.5 Severance Package as Insurance

In practice, some CEOs get paid upon their departure, and sometimes this is precommitted when the CEOs are newly appointed, or when the contracts for the incumbent CEOs are renegotiated (Rau and Xu (2013)). Since the payment is followed by separation, the ex ante agreement of severance pay can potentially provide insurance upon a reputational shock. But the mechanism of severance pay as insurance differs from the previous institutions: the ex ante severance pay agreement directly offers payment upon a negative shock to reputation as in standard insurance.

To capture this feature, suppose that principals can offer severance pay upon replacement (both forced and voluntary), and that this pay is legally binding.⁵ As insurance, the ex ante agreement of severance pay is attractive only if the realization of the payment upon a bad event

⁵In practice, if the severance package is specified in a CEO's employment contract, then it is legally binding.

(i.e., separation) is expected. Given the stationary distribution of firm size, the chance of separation is deterministic contingent on a performance outcome. The deterministic feature confounds severance pay with an extra transfer upon a poor outcome without replacement, or with extra pay upfront. Thus, to distinguish the severance pay from other transfers, suppose that the investment opportunity is subject to shock. The process governing the transition of the investment opportunity follows a discrete time Martingale process and is assumed to be exogenous. More formally,

$$\alpha_{t+1} = \alpha_t + \epsilon_{t+1}, E[\alpha_{t+1}|\alpha_t] = \alpha_t$$

where ϵ_t follows a well-defined symmetric distribution with support $(-\frac{\alpha_1}{2}, \frac{\alpha_1}{2})$ and variance σ^2 . Let EAS denote an agreed severance pay level, and assume that it is expected that the incumbent CEO will be replaced upon l . For severance pay to have an insurance effect for CEO γ , it must satisfy the following.

$$E[\tilde{Y}_3(S, \gamma^X) - \tilde{\omega}_3(\gamma^X)|\alpha_2] \geq E[\tilde{Y}_3(S, \gamma^l) - \tilde{\omega}_3(\gamma^l) + EAS|\alpha_2]$$

Basically, replacement with severance upon l should be incentive compatible for the principal ex ante. This is because, as insurance (payment after a bad event), the CEO certainly prefers the insurance to be paid instead of being retained by the incumbent firm. Now, a natural investigation is when and with what level such ex ante severance pay is agreed in a credible way. Lemma 13 summarizes the result for this investigation.

Lemma 13. *(Optimal Ex ante Severance Pay) If the principal wants to offer a severance pay agreement, then the severance pay committed by sufficiently large firms create an insurance effect. i.e., Not every firm is free to use ex ante severance pay as insurance. This promise of*

severance is only credible in a relatively less volatile economy: $\sigma^2 \leq \sigma^{FEAS}$.

The next investigation finds the characteristics of such ex ante severance pay agreement and the consequence thereof. As in performance non-disclosure, the severance pay as insurance can create match friction endogenously.

Proposition 8. (*Performance Tolerance*) *Suppose the severance pay is credibly promised ($\sigma^2 \leq \sigma^{FEAS}$). As the last period shock tends to be extreme (i.e., $\alpha_3 \downarrow 0$, or $\alpha_3 \uparrow 1$), then, $\forall S \in [S[m], S[m] + f(EAS, \alpha_3)]$ find it profitable to retain their CEOs upon l , where $f(EAS, \alpha_3) = \frac{EAS}{\alpha_3(1-\alpha_3)\Delta}$.*

Here, $\Delta = p(\alpha_3 + (1 - \alpha_3)(\gamma^m + \gamma^l))(\gamma^m - \gamma^l)$ captures the marginal benefit of CEO γ^m relative to γ^l in terms of project selection efficiency. The intuition behind Proposition 8 is that the principal might find it profitable to retain the poor performing incumbent CEO instead of paying the severance pay and hiring a new CEO. That is, introducing ex ante severance pay agreement might reduce the non-assortative rematching outcome in period 2. However, this result comes at the potential expense of ex post rematching distortion in period 3.

3.6 Related Literature

3.6.1 CEO Turnover and Firm Performance

Much of the literature on CEO turnover attempts to identify the driving forces that lead to the replacement of CEOs and the associated consequences. One recurring stylized fact regarding CEO turnover is that a firm's financial performance surrounding the forced resignation is inversely related to the likelihood of turnover (e.g., Coughlan and Schmidt (1985); Warner et al. (1988);

Jensen and Murphy (1990), Murphy and Zimmerman (1993); Murphy (1999) (survey); Brickley (2003); Kaplan (2012)). However, it has also been documented that the economic magnitude of the turnover-performance relation is arguably small (See discussion in Murphy (1999); Brickley (2003)). Among others, Puffer and Weintrop (1991) find evidence that the expected firm performance matters in deciding replacement of the CEO. i.e. The turnover occurs if firm performance fails to meet the expected performance. Dikolli et al. (2014) examine the effect of CEO tenure on the turnover-performance sensitivity. They suggest that, rather than the managerial entrenchment story, as tenure increases, a CEO's ability is likely to be revealed, thereby reducing a board's incentive to replace an incumbent CEO. Meanwhile, Jenter and Lewellen (2010) argue that the weak turnover-performance relation results from the misclassification of succession types. Even though the existing empirical studies have attempted to identify the exact reasoning of weak-turnover performance relation, the exact mechanism has yet to be identified.

From a theoretical perspective on CEO turnover, several recent papers attempt to explain the association between executive turnover and performance. In the context of a dynamic optimization framework, but with a similar driving force on the turnover, Garrett and Pavan (2012) focus on the effect of evolving productivity to explain cross-sectional difference of turnover policies across firms. They argue that, contrary to the managerial entrenchment story, it is optimal to have different levels of turnover-performance sensitivity depending on the CEO's tenure, thus providing a potential justification on the weak turnover-performance sensitivity. In the context of board independence, Laux (2008) studies how board independence affects CEO turnover when a CEO has private information about his ability. While independent boards will effectively replace the incumbent CEO, the independent board's aggressive turnover decision increases required sev-

erance pay (information rents) to induce truthful information. Laux (2008) shows that a less independent board can arise as optimal as a commitment not to replace the CEO aggressively. In a similar setting, Laux (2014) explores the impact of a CEO's misreport on a CEO turnover pattern. To induce truthful reporting about an earnings signal in order to make a better turnover decision, a board has to provide severance pay which weakens the CEO's effort incentive ex ante. In both settings, the weak turnover-performance sensitivity can endogenously arise as optimal contracting if the cost of inducing effort is greater than the benefit of a better turnover decision that involves severance pay to elicit truthful reporting. All these studies provide meaningful insights into how the CEO succession and performance are related. This paper, thus, seeks to extend the existing literature by providing another rationale to better understand the CEO succession and performance relation.

3.6.2 CEO Turnover, Reputation, and Project Selection

There are also several studies that examine the managerial project selection problem taking into account reputational incentives, and its impact on the decision of CEO replacement. Casamatta and Guembel (2010) analyze the relation between CEO turnover and the change of organizational strategies. They show that it is better to replace the incumbent CEO upon poor performance when it is also better to change the organizational strategy chosen by the incumbent. The intuition is the incumbent CEO's incentive to sabotage in order to prove that his previous choice of the strategy is correct. With information asymmetry on the agent's types, Hirshleifer and Thakor (1992) and Hirshleifer (1993) consider a similar set up with the present paper, but their focus is on the interaction between capital structure and managerial reputation incentives rather than

designing an optimal contract to mitigate the agency issue which is one goal of the present paper. Among other results, Hirshleifer and Thakor (1992) show that the inept type of agent takes a safe project since the success likelihood is higher although the outcome level is mediocre. Prendergast and Stole (1996) and Sliwka (2007) also consider the agent's project choice with information asymmetry, but the key tension of these two is the agent's distortion in project selection in order to signal his type: Prendergast and Stole (1996) shows that the agent sometimes uses his private signal heavily or less heavily over time depending on his incentive to signal whether he learns something or he learns everything; Sliwka (2007) finds the condition for when the agent becomes hesitant to change his previous selected project. On the other hand, using career concern (no information asymmetry), Boot (1992) shows how the distortion in investment takes place on account of the agent's signaling incentive.

However, the above literature assumes away the issue of designing explicit contract.⁶ The interactions between implicit and explicit incentives have been explored by Gibbons and Murphy (1992), Meyer and Vickers (1997), Autrey et al. (2007), Autrey et al. (2010). Gibbons and Murphy (1992) show that explicit incentive should increase as career concern declines, and Meyer and Vickers (1997) find that relative performance evaluation can either enhance or reduce the efficiency of implicit incentives. Autrey et al. (2007) examine how mandated disclosure of performance affects explicit incentive contracts, and Autrey et al. (2010) show that career concerns can either mitigate or magnify the performance aggregation costs in a multitask setting. In all these papers however, firms are identical, thus the market value for the agent (after the realiza-

⁶The exception is Casamatta and Guembel (2010) that finds optimal compensation contract, but their key focus is on the change of organizational strategy, and its impact on the turnover decision.

tion of period 1 performance) is independent of firm size distribution in the economy. My study complements to this literature by highlighting the cross-sectional differences in career concerns (i.e., implicit incentives) across CEOs due to heterogeneous firms and by examining the interplay between such career concerns and market-wide rematching patterns (the consequence of CEO turnover).

3.7 Conclusion

In this chapter, I develop a model of governance practices to explore the latent aspects as insurance for managers. Based on the project selection model with the CEO's career concerns in Chapter 1, I introduce three control devices separately to see how a firm effectively incentivizes its CEO to resolve agency conflicts arising from career concerns. In particular, when a CEO is concerned about the market perception of their ability (i.e., reputation) that is formed by firm performance, the CEO, a rational payoff maximizer, will be tilted toward protecting his perceived talent, which may not be the best interest of a firm. The models I propose also explore how these devices influence the association between firm performance and the pattern of CEO turnover. Although the presence of reputation insurance reduces inefficient replacement decisions *ex ante*, each mechanism create other channels that weaken the impact of firm performance on CEO turnover. By providing new perspectives of existing governance practices, this chapter also provides testable empirical predictions, and implications for corporate governance and investment behaviors.

To conclude this chapter, it is useful to point out a limitation and potential extension of the model. For instance, in a model of a board's monitoring, I abstract away from any potential

collusion between the board and the CEO by considering the board as a special monitoring technology since I focus on the role of extra information in resolving career concerns and making a better turnover decision. Introducing a potential collusive behavior, such as a side contract, between the board and the CEO might reduce a firm's incentive for incurring the costs for the board. Even further, a firm might have incentive to have the board with imprecise monitoring in order to prevent such collusive behavior. As an another extension, considering the combination of the three devices will shed additional light on the governance practices. In practice, every organization has a choice to have all these forms of governance. Depending on organizational characteristics or culture, what features of firms lead either of devices to be the preferred choice will be also interesting to consider.

Chapter 4

Generalists versus Specialists: When Do Firms Hire Externally?

Abstract

In recent decades, the trend of external CEO hiring has increased, a practice often involving high outsider pay premiums. Most academics and practitioners ascribe the practice of outsider premiums to two factors: managerial talents and a match between a firm and a CEO. However, this perspective seems to overlook that, after an outsider CEO is hired, firm performance often becomes unsatisfactory. To understand the missing link among CEO hiring, the pay premium and firm performance, this paper develops a multitask-multiagent team production model where each task sequentially requires a firm-specific skill and a management resource allocation. By analyzing a multiagent incomplete contracting problem, this paper identifies conditions under which either internal promotion remains optimal or external hiring becomes optimal. This optimal contracting approach for multiagent also explains why outsider CEOs appear to get paid more than insider CEOs and how the performance of external hiring firms tends to be worse than the performance of internal promoting firms in spite of the higher pay.

4.1 Introduction

The issue of finding the best CEO candidate for a firm and determining his/her appropriate compensation has attracted the attention of academics and practitioners alike. Undoubtedly, if an existing employee is qualified enough, then a firm will fill the CEO position with this qualified current employee. However, if there is no qualified internal candidate, a firm will likely hire the most appropriate person from outside the firm even though it requires a higher pay premium. Existing literature helps explain such external hiring trend by describing how a firm's hiring decision and the level of offered compensation are determined by managerial talents, the labor market condition for executives, and a match between a firm and a CEO (Murphy and Zabochnik (2004); Murphy and Zabochnik (2007)). This literature has also pointed out how CEOs' general and transferable skills have become more important than firm-specific skills, thus making a firm seek out a CEO who owns such transferable skills (Murphy and Zabochnik (2004); Murphy and Zabochnik (2007); Custódio et al. (2013)).

However, even though external CEOs get paid high premiums, these external hiring firms often exhibit unsatisfactory subsequent performance. While the existing literature explains the demand for outside CEOs and the associated pay premium, a major limitation is that it overlooks this mismatch between firm performance and pay premium (Zajac (1990); Shen and Cannella (2002); Karaevli (2007) for survey; Kale et al. (2009)). If the CEO's talent leads to an outsider CEO premium, but does not generate better performance, then the question becomes why would a firm want to hire an outsider CEO. This paper aims to fill this gap by considering CEO hiring choices as incentive devices for non-CEO employees to explain this mismatch. When an organi-

zation wants its employees to acquire firm-specific skills, features of firm specificity becomes a key friction: by its nature, a firm-specific skill is unverifiable, thus uncontractible, although who owns firm-specific skills is observable within an organization (Prendergast (1993)). As a solution to this incomplete contract problem, the organization can make use of CEO hiring choices to create firm-specific skill acquisition incentives for their employees. Since firm performance is affected not only by a CEO's behavior but also by non-CEO executives (hereafter, subordinates), considering how a firm motivates its subordinates is necessary to understand the missing link between CEO appointments and firm performance.

By its nature, a firm-specific skill is unverifiable, implying a wage contingent on the skill is not credible *ex ante*. Thus employees lack incentive to acquire the firm-specific skill. Moreover, the skill is only meaningful for the current firm, making the skill nontransferable, which further reduces the skill acquisition incentive. It is well known that this type of incomplete contract problem can be solved by asset ownership and/or property rights (Grossman and Hart (1986); Hart and Moore (1990)); however the characteristics of firm-specific skills imply that, although the skill is owned by agents, the ownership over the skill may not incentivize agents because the firm-specific skill has less value outside the firm (i.e., unverifiable and nontransferable skill).¹ Another implicit assumption in this argument is that an employee has no bargaining power over his firm specific skill. The statement that the supply of subordinates labor market is sufficiently high and inelastic can support the argument. Therefore, if a firm-specific skill is necessary, an organization needs to find ways to incentivize its subordinates.

¹This argument implicitly considers firm-specific skills as assets. Hence, the control right over the firm-specific skill does not change the outside option.

This paper argues that the decision of whether a CEO is hired internally or externally creates subordinates' incentive to develop firm-specific skills, and this can account for the consequence of subsequent firm performance. To illustrate this idea, I offer an optimal contracting view based on a sequential team production model with two hiring choices: internal promotion and external hiring. More specifically, a firm-specific skill developed by subordinates first affects production technology of each task permanently. Then, a superior will finalize the productivity by allocating resources into tasks. Both internal promotion and external hiring utilize the competition between subordinates, however, the key difference of hiring choices is whether subordinates compete for a superior position or for the resources from a superior.

By considering multiple agents' incentive problems, this paper compares efficiency of the two distinctive mechanisms (internal promotion and external hiring) contingent on the benefit of investment for firm-specific skills. In general, internal promotion succeeds in providing firm-specific skill development incentives and adequate management allocation incentives by aggregating two incentive problems using a promotion bonus (or management slack). This aggregation is enabled by a rank order tournament of realized firm-specific skills, thereby inducing an investment incentive to acquire unverifiable firm-specific skills. The management slack in this case is necessary since winning the tournament itself may not be enough if the reward is not big enough. In the model, a superior who has acquired a firm-specific skill can save his cost of managing his own skill (task). The cost saving from the expertise is a two-edged sword. On one hand, the superior can effectively manage his own task. On the other hand, the expertise benefit creates favoritism. With these benefit and cost, this management slack by aggregating the two incentive problems is an efficient mechanism to the multiagent incomplete contracting

problem if such favoritism creates enough slack for the superior. Since enough management slack incentivizes the subordinates effectively, to the extent that the slack is big enough relative to the superior's biased allocation, the incentive aggregation remain as optimal for the multiagent incomplete contracting problem.

However, external hiring becomes optimal if such management slack incurs too much distortions in management. If the slack is not enough, a firm needs to induce more bias for enough slack, thereby creating distortions in management allocation. In this case, it is optimal for the principal to use external hiring in which the externally hired agent makes a management resource allocation decision contingent on realized skills. Although external hiring separates two incentive problems (skill investment and management allocation), which usually makes the total compensation costs expensive, external hiring can be optimal if the overall costs are exceeded by the management slack required.

This dominance between internal promotion and external hiring is characterized by the efficiency of firm-specific skill on management. If the impact of the expertise on managing the superior's own skill is high, this creates large slack, thus incentivizing the subordinates (in anticipation of such slack). The benefit of expertise naturally leads to relatively lower pay for an internal hire. However, as the benefit from managing its own skill becomes less effective, the required pay for the internal CEO increases and management efficiency decreases, thereby making external hiring optimal. As a result, from an optimal contracting point of view, the pay for outsider CEOs might not be a *de facto* premium, but rather an optimal wage contingent on firm-specific skills and internal production technology (i.e., cost saving from expertise). Moreover, the external hiring firm's unsatisfactory performance might not be attributed to the externally

hired CEO. If an optimal mechanism is external hiring, but a firm relies on internal promotion, then a firm's payoff would be much lower than the payoff under the external hiring due to the biased allocation for the sake of management slack. This approach incorporating the provision of incentives inherent in hiring choice and management decision offers a way to understand cross-sectional differences in CEO selections and to provide testable implications of CEO hiring practices: why some firms find their CEOs from inside (or outside), why outside hiring seems to involve higher pay, and why external hiring firms are more likely to face diminished performance despite the higher pay.

It is worth pointing out that this paper does not simply argue that the factors, found by existing studies including managerial ability, the match and labor market competition, cannot explain the current practice of CEO appointments and compensation. Rather, this paper takes a complementary position to this existing literature to provide a better understanding of the hiring practice and firm performance. Furthermore, overcoming unverifiableness in firm-specific skills per se is not the only key argument of this paper. The main point is that we must understand the existing practice of CEO hiring, pay, and firm performance based on the holistic perspective of mechanism design to the multiagent incomplete contracting problem and distinct incentive effects thereof. The paper proceeds as follows. Section 2 relates this paper to the existing literature. Section 3 describes the model. Section 4 analyzes the model and finds an optimal form of CEO appointments. Section 5 extends the main model and Section 6 provides empirical implications. Conclusion is presented in Section 5. All proofs are in the appendix.

4.2 Related literature

The fundamental research goal of this paper is to understand the association between CEO appointments, the pay and firm performance, in particular, the missing link between the outsider pay premium and the firm's unsatisfactory performance. To achieve this goal, this paper considers CEO appointments as mechanism to solve contract incompleteness inherent in firm specificity: internal promotion as a tournament mechanism and external hiring as a multitask mechanism. Thus, this section is dedicated to review these two strands of literature, tournaments and the multitask agency literature.

As originally investigated in Lazear and Rosen (1981) and Green and Stokey (1983), a rank-order tournament has been analyzed as an implementation device to incentivize agents to work. Lazear and Rosen (1981) has shown that the tournament scheme can perform identically as the piece rate scheme under the condition of no common shock and agents' risk neutrality, and further explained why the prize level is convex increasing in ranks. Meanwhile, Green and Stokey (1983) has examined the tournament scheme in a setting where agents are subject to idiosyncratic and common shock and identified the conditions under which the tournament dominates the piece rate scheme. Since then, accounting, economics, finance, and the strategic management literature have studied tournaments to understand executive compensation and found theoretically consistent results. (O'Reilly III et al. (1988); Rees (1992); Main et al. (1993); Eriksson (1999); Kale et al. (2009)).

The follow-up literature on tournaments has discovered other positive and negative aspects of tournaments. As a negative aspect, the tournament mechanism can be problematic as a moti-

vating device due to collusion and/or sabotage(Lazear and Rosen (1981); Dye (1984)). However, other incentive devices overcoming this negative aspect are also suggested. For example, Chen (2005) has shown that external hiring might resolve such sabotage incentive of internal employees. In his model, sabotage is costly to employees, so the principal's commitment to use external recruitment with target promotion level reduces the marginal benefit of sabotage since this negative activity cannot affect external candidates. Meanwhile, as the present paper will point out, Malcomson (1984) has highlighted another benefit of a rank-order tournament that overcomes information asymmetry between an employer and an employee. When the outcome is subjectively assessed by the employer, thus failing at creating work incentive ex ante, a rank-order tournament recovers the employees' work incentive given that the employer fixes the total promotion level, thus tournament prize level, upfront. In a rather different context, Fairburn and Malcomson (2001) has also a similar feature with the present paper in that they connect the unverifiable performance measure to the tournament scheme. In their model, promotion is subject to bribes by subordinates to influence the performance assessor i.e., the manager. By linking the manager's pay to the outcome after promotion, it reduces the manager's incentive of accepting bribes and manipulating the performance assessments.

Conceptually, the internal promotion in this paper has an analogous effect of renegotiation in Hermalin and Katz (1991). Hermalin and Katz (1991) has verified that renegotiation can be beneficial if there is informative but unverifiable signal about the agent's action. This is because, based on the interim signal, the principal can offer a new contract that is mutually beneficial. Thus, the presence of renegotiation with respect to the unverifiable signal induces the agent's effort. Similarly, the internal promotion (i.e., a tournament) based on unverifiable firm-

specific skill can successfully motivate subordinates to invest in the first period. However, the difference between Hermalin and Katz (1991) and this paper is that the renegotiation combined with unverifiable interim information induces the agent's effort, while the internal promotion in this paper is a device to overcome such unverifiableness of the skill.

The multitask agency literature, another main strand of related research to this paper, has emphasized the issue of multidimensionality of tasks (Holmstrom and Milgrom (1991); Holmstrom (1999); Dewatripont and Tirole (1999); Dewatripont et al. (2000) for survey; MacDonald and Marx (2001)). As Holmstrom and Milgrom (1991) has highlighted, the multidimensionality might be problematic due to effort substitution. In particular, the effort substitution is mainly attributed to the precision of performance measure: the problem arises if the precision of performance measure for each task differs. That is, if the wage is sensitive to outcomes of tasks, then an agent would want to exert his effort on more precisely measured tasks rather than exerting effort on noisily measured tasks. In this circumstance, Holmstrom and Milgrom (1991) have found that, depending on the principal's desired effort level and the functional form of agent's effort cost, the incentive scheme might be high-powered or low-powered on a better measured task. As an extension, MacDonald and Marx (2001) have developed an applied model of effort substitution to understand the convex shape of CEO pay. Since the agent might specialize on one task among many to save his effort cost, which makes the moral hazard problem more severe, the optimal reward for medium outcome must be low enough so that the agent has no incentive to exert the effort only on his specialized task, and the optimal reward for full success must be high enough to induce high effort on every task.

However, unlike these traditional multitask agency literature, this paper highlights a hidden

characteristic of multitask under the hierarchical structure, namely attention getting competition. As this paper will show, under the sequential team production setting, the presence of effort substitution by a superior in period 2 can motivate subordinates in period 1. The superior's discriminating effort strategy with respect to unverifiable skills thus resolves both subordinates' moral hazard and the principal's ex post opportunism.

In terms of a general framework, Chan (1996) and Harris and Helfat (1997) relate to this paper in that they compare internal promotion to external hiring. Chan (1996) has shown analytically how external hiring reduces work incentive attributed to internal promotion and finds how a firm can regain the lost work incentive. Chan (1996) has suggested a competitive handicap can be another incentive device to recover the lost tournament incentive: external recruitment occurs only if the quality of external candidate is significantly high. Meanwhile, Harris and Helfat (1997) has empirically investigated pay difference between internally and externally appointed CEOs, and connects the pay premium to different skill sets of those CEO candidates. However, both of these papers are silent about subsequent firm performances: Chan (1996)'s main focus is about tournament incentive per se, not about the pay premium; Harris and Helfat (1997)'s focus is to examine the relationship between CEO types and initial cash compensation.

From a methodological point of view, the present paper's approach based on tournaments and multitask agency is not new. However, deviating from the existing literature, this paper offers a novel explanation using these two traditional theories to understand CEO appointments and subsequent firm performance. By considering CEO selections as organizational incentive devices, this paper is able to explain how organizations motivate their employees to acquire firm-specific skills, thereby filling the missing link between the CEO pay premium and unsatisfactory

subsequent performance.

4.3 The Model

In this section, I develop a model in which a principal hires multiple agents to complete firm production. The sequential production process involves two stages: the first stage requires investments to develop firm-specific skills, and the second stage requires a management decision to determine the outcomes of multiple tasks. The investment for skills can be broadly interpreted as the development of firm-specific human capital, productivity, ideas or any actions that can improve firm revenue, but only meaningful in the present firm.² Since this paper focuses on an agent's willingness to acquire firm-specific skills, assume that these skills cannot be trained by a firm.³

Firm-specific skills affects production technology permanently; however, the skills, per se, do not generate revenue without a management decision: the gains from the skills rely on how those skills are managed. Based on this sequential production, assume that the principal organizes jobs through a two-levels hierarchy: agents in the lower level conduct firm-specific skill investment jobs in stage 1, and an agent in the higher level conducts the management job for achieving successful outcomes in stage 2. Although these hierarchically designed jobs seem not unrealistic,

²The skill might also be used in another firm, however due to firm specificity, the skill can be more efficiently used in the present firm. Incorporating this possibility using a parameter would not change the qualitative setting but incur complexity.

³In practice, many firms have a job training system to educate their employees about their operations and/or tasks, however, there seems no such training system for top executives, for example vice presidents, to motivate them to acquire firm specialized skills.

the assumption will be discussed in the conclusion section to argue that this job design is indeed efficient.

Now, the principal needs to deal with two incentive problems: she should first provide enough incentive to subordinate agents in the first stage to make them invest; second, she must give a superior agent appropriate incentive to induce the best management decision to maximize firm revenue. Depending on how the principal incentivizes her agents, the firm has different hiring structures: internal promotion or external hiring. Combining two incentive problems and introducing different hiring structures under a sequential team production setting are intended to develop a model that incorporates organizations' different ways of motivating their agents to acquire firm-specific skills. This holistic approach provides a way to better understand the organizational choice of CEOs through the lens of incentive devices for subordinates to acquire firm-specific skills, namely competition to be a superior (tournament) versus competition to attract management attention (multitask).

4.3.1 Two Stage Production Process

This section describes two stage production process more specifically to highlight key agency issues in each stage.

First stage - Firm-specific skills development: There are two main problems in stage 1, the moral hazard problem and the enforceability problem⁴, contributed to the unverifiability of actions and skills.

⁴The term of enforceability is in the sense of Malcomson (1984). Kahn and Huberman (1988) calls this a bilateral moral hazard problem and Prendergast (1993) refers this to a dual moral hazard problem.

At the beginning of stage 1, the principal hires one agent for each task to develop firm-specific skills, $\theta \in \{H, L\}$ where $H > L > 0$. Assume that the skills for task 1 and task 2 are independent of one another. This job (hereafter, investment) incurs the disutility of $I \in \{c, 0\}$, $c > 0$ that is borne by an agent to whom the task is assigned in stage 1. The choice of investment determines the likelihood of skill parameters : $Pr(\theta = H|I = c) = p$, $Pr(\theta = H|I = 0) = q$, and $p > q$, i.e., the realization of $\theta = H$ is more likely under the agent's costly investment. Assume that the firm-specific skill for task i permanently changes the task i 's production technology, which will be specified soon. Assume also that $\theta = H$ increases the efficiency of the production technology.

However, the agent's investment choice is unobservable, thus uncontractible. Assume further that the realized skill is soft information which cannot be verified by a third party, thus it is subject to the principal's ex post opportunism.⁵ Since the skill and the expertise are firm-specific, they do not generate any value if the agent leaves the current firm after the investment. This implies that the entire bargaining power over the skill is vested in the principal, thereby further reducing the agent's investment incentive ex ante. Putting this all together, the fixed wage fails to induce the investment in stage 1 due to the standard moral hazard problem, but the wage contingent on the realized skill parameter is not credible ex ante. Therefore, a contract for subordinates is required to resolve both moral hazard and the enforceability problems.

Second stage - Management effort: Contrary to stage 1, management decision problem in stage 2 is only subject to the traditional moral hazard problem.

After each skill parameter is realized at the end of stage 1, but before the stage 2 begins, the principal designates a superior who needs to manage two skills realized in stage 1 to generate

⁵Employees' expertise or firm-specific skills seem difficult to say to be hard information.

revenue. The superior in stage 2 makes a management decision. The management decision is to allocate superior's management resource (hereafter, effort) into the two tasks. Assume that a superior's effort is normalized to 1. Since the realized skills in stage 1 are used in stage 2 combined with the management effort, the superior's effort allocation is contingent on the realized skills, $m(\theta_1, \theta_2) = (m_1, m_2)$, $m_1 + m_2 \leq 1$. The allocated management effort determines the outcome of each task, t_1, t_2 . The outcome is either S or F , where S denotes success and F denotes failure. This management job incurs costs, $c(m, \theta) = (m_1 + m_2) \cdot d$, where $d > 0$ denotes the cost of management for a unit of management effort. Assume, as in standard moral hazard setup, that m is unverifiable, thus uncontractible.

Production Technology: Let $f(\theta, m)$ denote a production technology upon θ, m : $f(\theta, m) = Pr(t = S | \theta, m) \in (0, 1)$. Both θ, m might influence the production function differently, but to highlight the basic intuition in this paper, I will focus on the effect of firm-specific skill and management effort are additive.⁶ That is,

$$f(\theta, m) = f(\theta + m)$$

It is assumed conventionally that, $f(\theta + m)$ is differentiable, non-decreasing with respect to $(\theta + m)$.

For simplicity, assume that the market supply of agents is unlimited in order to guarantee that the principal can find an outsider without other frictions, as the competition in the CEO labor market is not the main focus of this paper.

Contracts: Let $X_{(t_1, t_2)}$ denote revenue contingent on the outcome of tasks, (t_1, t_2) . Although the agents' investment choice and management effort allocation are not verifiable, the revenue

⁶The multiplicative case is released to the appendix.

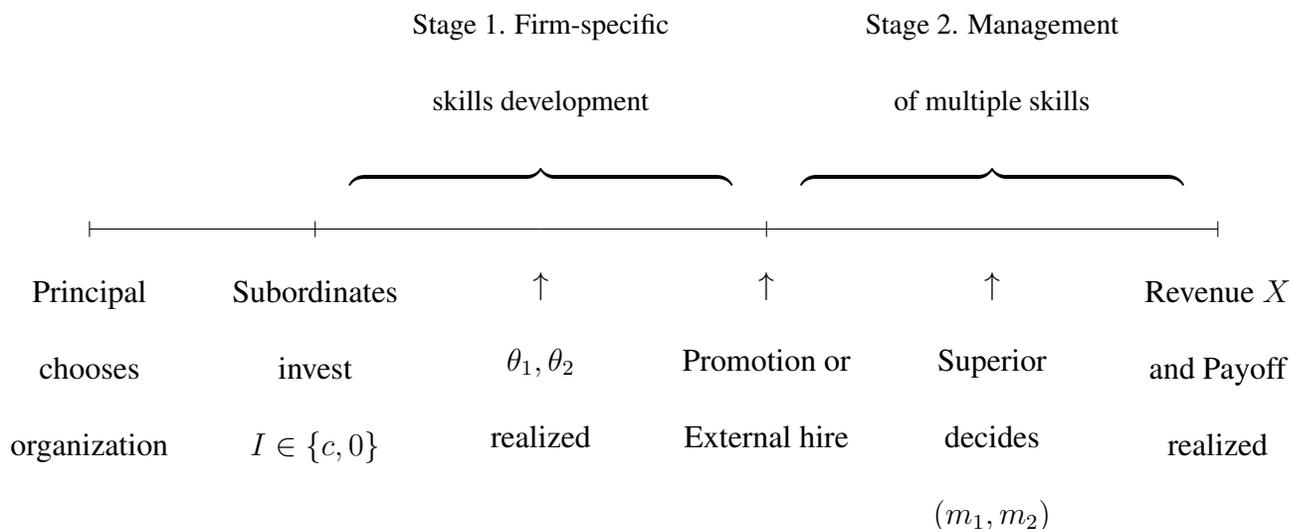


Figure 4.1: **Timeline.**

$X_{(t_1, t_2)}$, is verifiable, thus contractible. For the sake of notation, let $X_{(S, F)} = X_{(F, S)} = X_S$. Assume that the principal values full success of both tasks than the sum of an individual success, i.e. $X_{(S, S)} > 2X_S$. To summarize, the principal designs a contract including hiring choice that incentivizes agents to induce firm-specific investments in the first stage and to make the best management decision in the second stage to maximize her net payoff. Figure 4.1 depicts the sequence of events.

4.3.2 Internal Promotion versus External Hiring

This section provides the key mechanism of each hiring structure to understand different incentive effects. To overcome contract incompleteness inherent in unverifiable and non-transferable firm-specific skills through competition, the principal can use either internal promotion or external hiring: internal promotion introduces competition over the skills directly and external hiring

relies on competition over the skills indirectly to attract management effort.

Internal Promotion: If the principal decides to use internal promotion, she designates one of subordinates as a superior depending on the rank order of θ s: If $(\theta_1, \theta_2) = (H, L)$, then the subordinate 1 is promoted; if $\theta_1 = \theta_2$, then the principal randomly picks one of the two subordinates. The promoted agent, superior, receives a decision right to allocate management effort in stage 2. After the outcome is realized at the end of stage 2, the promoted agent gets the compensation, and the unpromoted agent earns the subordinate's wage. Although θ is soft information, this internal promotion based on the rank order can resolve the principal's ex post opportunism given that the self-commitment property of a tournament.⁷

Subordinates are promised to receive the promotion bonus, $\bar{U} > 0$ if promoted, otherwise, they will get v .⁸ Since the contract contingent on (θ_1, θ_2) is not credible ex ante, the principal specifies the compensation contract for the internal superior taking into account the worst case, (θ_L, θ_L) . This generates management slack in stage 2 depending on the realized skills in stage 1. In essence, the expected slack plays a role as a promotion bonus.⁹ Assume that the principal has full commitment power to stick to the initially offered contract.

External Hiring: External hiring finds a superior from outside. Contrary to the internal promotion, the principal specifies the compensation contract for the superior after the state, (θ_1, θ_2) , is realized. Then, the newly hired agent determines the management allocation decision. Since

⁷Malcomson (1984) shows that tournament incentives can resolve the problem of unverifiable performance measure given that the prize level is fixed in advance.

⁸It is also possible to compensate the unpromoted agent contingent on the final outcome of the task. But that makes the analysis complicated without getting much clear insights.

⁹The introduction of promotion bonus seems reasonable from a practical perspective in that the bonus is quite common in practice for internally promoted CEOs (Equilar (2013)).

the subordinate's incentive in external hiring is provided through management effort externality, assume that, contrary to internal promotion, the subordinates get paid contingent on the outcome of their own task, which is determined by the allocated management decision. Since the choice of hiring structure is known ex ante, and management allocation is affected by the realized skills, thus, this creates competition between the subordinates in order not to lose the superior's effort to another subordinate. Since the main focus of this paper is to highlight two different incentive devices to overcome contract incompleteness, assume that there is no negative activity such as sabotage.

Ultimately, the external hiring utilizes the indirect competition between subordinates to attract more effort from a superior contrary to the internal promotion exploiting the direct competition between subordinates to be a superior.

4.3.3 Incentive Contracts for Each Organization

Depending on the principal's choice of organization, the program for optimal contract differs: internal promotion relies on incentive aggregation and external hiring relies on incentive separation. For expositional ease, I will alternately use internal promotion and IP, similarly external hiring and EH.

Internal Promotion: Suppose the principal chooses internal promotion. Then, at the beginning, the principal decides the compensation contract upon the final outcome that characterizes promotion bonus \bar{U} and losing prize v to induce the investment in stage 1. Given the promotion bonus from the expected compensation in stage 2, the following condition characterizes the

incentive compatibility constraint in stage 1.

$$\begin{aligned} & p(\text{win}|I = c) \cdot \bar{U} + (1 - p(\text{win}|I = c)) \cdot v - c \\ & \geq p(\text{win}|I = 0) \cdot \bar{U} + (1 - p(\text{win}|I = 0)) \cdot v \end{aligned} \quad (\text{IC-1})$$

Note that it is straightforward to see $p(\text{win}|I = c) = \frac{1}{2}$, $p(\text{win}|I = 0) = \frac{1-p+q}{2}$. Then, the (IC-1) constraint yields the relation between the promotion bonus, \bar{U} , and the losing prize, v .

$$\bar{U} \geq v + \frac{2c}{p - q} \quad (4.1)$$

Intuitively, the promotion bonus, \bar{U} , increases in the losing prize, v , which is well-known in a traditional tournament model. There is also an individual rationality constraint that attracts agents to participate, however, as long as the promotion bonus is satisfied with the above condition, it does not bind.¹⁰ To minimize compensation costs, it is always optimal to set $v = 0$.

Since the expected management slack indeed provides investment incentive in stage 1 under IP, the principal solves the following program when offering a contract before the state is realized. In equilibrium, the superior's contract will exhaust all feasible management resource for the two tasks. Thus $m_2 = 1 - m_1$. All the bold symbols denote vectors.

$$\max_{m,w} E[\mathbf{X} - \mathbf{w}] \quad (\text{IP})$$

$$E[\mathbf{w}] - (m_1 + m_2)d \geq \bar{u} \quad (\text{IR-2})$$

$$m \in \text{argmax}_n E[\mathbf{w}] - (n_1 + n_2)d \quad (\text{IC-2})$$

$$E_\theta[\mathbf{w}|\mathbf{m}] - (m_1 + m_2)d - E_{\theta_L}[\mathbf{w}|\mathbf{m}] \geq \frac{2c}{p - q} \quad (\text{Slack})$$

¹⁰Given that the initial reservation utility is normalized to zero, the (IR-1) constraint is, $p(\text{win}|I = c) \cdot \bar{U} + (1 - p(\text{win}|I = c)) \cdot v - c \geq 0$.

Note that two tasks are independent,

$$E[\mathbf{X} - \mathbf{w}] = \Pr(\mathbf{X}_{SS}|\theta, \mathbf{m})(\mathbf{X}_{SS} - \mathbf{w}_{SS}) + \Pr(\mathbf{X}_S|\theta, \mathbf{m})(\mathbf{X}_S - \mathbf{w}_S) \quad (4.2)$$

where $Pr(X_{SS}|\theta, \mathbf{m}) = f(\theta_1 + m_1)f(\theta_2 + m_2)$, $Pr(X_S|\theta, \mathbf{m}) = f(\theta_1 + m_1)(1 - f(\theta_2 + m_2)) + (1 - f(\theta_1 + m_1))f(\theta_2 + m_2)$. Recall that the contract contingent on θ is not credible ex ante, thus \mathbf{w} is determined based on the worst case, i.e. θ_L , hence the expected promotion bonus (management slack) is computed as follows.

$$\bar{U} = \sum_{\theta} Pr(\theta, \text{promoted})(E[w|\theta] - E[w|\theta_L]) \quad (\text{Promotion bonus})$$

Basically, the promotion bonus refers to additional payoff in addition to the compensation that is determined based on θ_L .

External Hiring: Now, suppose that the principal decides to use external hiring. Recall that in the external hiring case, subordinates get motivated by the fact that the outcome of their task is determined by the externally hired superior's management decision which is contingent on θ . i.e., it is the full anticipation of discriminating management decision that creates the investment incentives in case of $\theta_1 \neq \theta_2$. For expositional convenience, assume that, in case of $\theta_1 \neq \theta_2$, it is the task 1 whose H .

Then, at the beginning, principal decides v to induce investment from subordinates.

$$Pr(X_S|I = c)v - c \geq Pr(X_S|I = 0)v, \quad (\text{IC-1})$$

which yields,

$$v \geq \frac{c}{Pr(X_S|I = c) - Pr(X_S|I = 0)} \quad (4.3)$$

Again, given that the reservation utility is normalized to zero, the individual rationality constraint (IR-1) is satisfied as long as v satisfies the above condition.

To determine the externally hired superior's compensation, it is optimal for the principal to specify the compensation contract after the skills are realized. i.e., after the state information for production is realized, the principal solves,

$$\max_{m,w,v} E[\mathbf{X} - \mathbf{w} - v] \quad (\text{EH})$$

$$E[\mathbf{w}] - (m_1 + m_2)d \geq \bar{u} \quad (\text{IR-2})$$

$$m \in \text{argmax}_n E[\mathbf{w}] - (n_1 + n_2)d \quad (\text{IC-2})$$

where v satisfies (4.3).

Now, the principal's problem is complex because the incentive problem in stage 2 is inter-linked with the incentive problem in stage 1. Hence, depending on the choice of organization and a desirable management decision from the principal's perspective, the corresponding pay will differ. The solution strategy is as follows. I first solve an optimal management allocation problem in stage 2 for each production technology considering the realized state in stage 1 as given. Then, I will determine the cheapest incentive compatible pay for an internally promoting firm and an externally hiring firm to implement the optimal management decision. After finding the implementation problem, I solve the principal's optimization problem to determine which organization is better than the other. This highlights the benefit and cost of aggregation of the two incentive problems. Then, I compare the efficiency of the two organizational types based on their performance. As a finalizing step for empirical puzzles, I specify the production function and parameter values to offer implications on the demand for outsider CEOs and on the association

among CEO appointment, pay and firm performance.

4.4 Analysis

4.4.1 Management Effort Choice and Compensation

Management Effort Allocation Problem

In this section, I characterize management effort choice under different production technologies, taking as given the firm specific skills. Then, I find optimal compensation for an internally promoted superior and an externally hired superior. As the trade-off that this paper aims to come across is between incentive aggregation (IP) and incentive separation (EH), introducing allocation distortion would make this key trade-off less clear. Thus, to isolate this trade-off, assume that X is sufficiently high compared to w, v, d , so that the principal finds the first-best level of management allocation to maximize the expected revenue. Or, consider X as surplus taking into account compensation paid. Then

$$E[\mathbf{X}] = f(\theta_1 + m_1)f(\theta_2 + m_2)(X_{SS} - 2X_S) + (f(\theta_1 + m_1) + f(\theta_2 + m_2))X_S \quad (\text{II})$$

Concave Production Technology: Given the assumption that $X_{SS} - 2X_S > 0$, the second order derivative of the above objective is always negative. The first order condition then characterizes the first best efficient management effort allocation,

$$m_1 = \frac{1 - \theta_1 + \theta_2}{2}, \quad m_2 = \frac{1 + \theta_1 - \theta_2}{2} \quad (4.4)$$

For instance, if $(\theta_1, \theta_2) = (\theta_H, \theta_L)$, then $m_1 < m_2$. This will generate equal likelihood of getting success by implementing balanced management allocation, which I shall call inverse

discrimination.¹¹ Since external hiring induces investment incentive through the anticipation of managerial effort discrimination, this immediately yields the following observation.

Observation 1. *If $f(\cdot)$ is concave, then the external hiring is not efficient in creating investment incentive in stage 1.*

Convex Production Technology: Now, (Π) attains its minimum at the (4.4). Thus, $(1,0)$ is always optimal.¹² This seems potentially inefficient especially in case of (θ_H, θ_H) . One might suggest hiring two CEOs to manage two tasks instead of shutting down one or selling one task to other firm. This can increase organizational efficiency further, and I will investigate these possibilities in the extension. The following lemma summarizes the discussion above.

Lemma 14. *If $f(\cdot)$ is concave, $m_1 = \frac{1-\theta_1+\theta_2}{2}$, $m_2 = \frac{1+\theta_1-\theta_2}{2}$, suggesting that the management resource allocation shows inverse discrimination. If $f(\cdot)$ is convex, then $(1,0)$ is optimal.*

Optimal Compensation

This subsection finds an optimal compensation to implement different management allocations that are determined by production technology.

Concave Production: One can easily see that if $f(\cdot)$ is concave, it is optimal to pay only for X_{SS} as balanced management allocation is efficient. For notational convenience, let $f(m^* + \theta) = f_\theta^B$, denoting a success likelihood of a single task upon a balanced allocation. Recall that the principal needs to set the pay taking into account the worst case, thus $w_{SS} = \frac{d}{(f_\theta^B)^2}$. Then, the promotion

¹¹I call this balanced as it achieves equal likelihood of getting success in both tasks, and I call this inverse discrimination as more management effort is exerted into an inferior task.

¹²If (θ_H, θ_L) , then clearly $(1,0)$ is optimal. If (θ_H, θ_H) or (θ_L, θ_L) , then the superior randomly chooses one, say task 1, then allocate all the resources into that task.

bonus becomes,

$$\bar{U} = \xi^{SS} w_{SS}$$

where $\xi^{SS} = Pr(X_{SS}, (\theta_H, \theta_H), \text{promoted}) + Pr(X_{SS}, (\theta_H, \theta_L), \text{promoted})$ denotes ex ante probability of getting slack upon the outcome of X_{SS} . Recall that to induce firm-specific skill investment, $\bar{U} \geq \frac{2c}{p-q}$. This committed promotion bonus (management slack) might outweigh the investment cost in stage 1. Then, the anticipation of getting this slack in stage 2 flows into the stage 1, which I refer to spillover effect. Hence, if $\xi^{SS} w_{SS} \geq \frac{2c}{p-q}$, then the expected compensation cost for the internally promoted CEO is,

$$E^{IP}[W(\theta_L)] = \eta^{SS} \frac{d}{(f_L^B)^2} \quad (\text{Spillover})$$

where $\eta^{SS} = Pr(X_{SS}, (\theta_H, \theta_H)) + Pr(X_{SS}, (\theta_H, \theta_L)) + Pr(X_{SS}, (\theta_L, \theta_L))$ denotes ex ante probability of getting X_{SS} . On the other hand, if $\bar{U} < \frac{2c}{p-q}$, the compensation w_{SS} needs to be adjusted so that $\bar{U} = \frac{2c}{p-q}$, thus making $w_{SS} = \frac{1}{\xi^{SS}} \frac{2c}{p-q}$. That is, the demand for creating investment incentive in stage 1 spill backs to stage 2, which I refer to spillback effect. Therefore, the expected compensation cost becomes,

$$E^{IP}[W(\theta_L)] = \frac{\eta^{SS}}{\xi^{SS}} \frac{2c}{p-q} \quad (\text{Spillback})$$

The following proposition summarizes the discussion.

Proposition 9. *If $f(\cdot)$ is concave, the principal always prefers IP to EH, and an optimal compensation contract for an internally promoted superior is,*

$$w_{SS} = \max \left\{ \frac{d}{(f_L^B)^2}, \frac{1}{\xi^{SS}} \frac{2c}{p-q} \right\}$$

where ξ^{SS} denotes a likelihood of getting slack.

Convex Production: If $f(\cdot)$ is convex, then since (1,0) is optimal, the problem reduces to 1) inducing effort of 1, 2) allocating all the effort to θ_1 . Formally,

$$f(\theta_1 + 1)w^S - d \geq \max \{f(\theta_1 + m)w^S - md, f(\theta_2 + n)w^S - nd\} \quad \text{for } \forall m, n \in [0, 1)$$

Given that $\theta_1 \geq \theta_2$, the second term of the right hand side is always dominated by the first term. Moreover, conditional on that $f(\cdot)$ is convex, any $m < 1$ is suboptimal. Thus, the incentive compatibility boils down to binary effort choice problem which yields $w^S = \frac{d}{f(1+\theta) - f(\theta)}$. Again, for notational convenience, let $f(1 + \theta) = f_\theta^U, f(\theta) = f_\theta$. Then, the expected compensation cost for internal promotion is,

$$\begin{aligned} E^{IP}[W(\theta_L)] &= \eta^S \frac{d}{f_L^U - f_L} && \text{(if } \bar{U} \geq \frac{2c}{p-q} \text{)} \\ &= \frac{\eta^S}{\xi^S} \frac{2c}{p-q} && \text{(Otherwise)} \end{aligned}$$

where $\eta^S = Pr(X_S, (\theta_H, \theta_H)) + Pr(X_S, (\theta_H, \theta_L)) + Pr(X_S, (\theta_L, \theta_L))$ denotes ex ante probability of getting X_S , and $\xi^S = Pr(X_S, (\theta_H, \theta_H), \text{promoted}) + Pr(X_S, (\theta_H, \theta_L), \text{promoted})$ denotes ex ante probability of getting slack upon the outcome of X_S .

Recall that the compensation for the external CEO is determined after the effect of firm-specific skills are realized. Thus, $w_S = \frac{d}{f_\theta^U - f_\theta}$ for a realized θ . Then, the ex ante expected compensation cost becomes,

$$\begin{aligned} E^{EH}[W(\theta)] &= p^2 f_H^U \frac{d}{f_H^U - f_H} + 2p(1-p) f_H^U \frac{d}{f_H^U - f_H} + (1-p)^2 f_L^U \frac{d}{f_L^U - f_L} \\ &= \eta_H^S LL_H + \eta_L^S LL_L + d \end{aligned}$$

where η_θ^S denote a success likelihood upon θ , and LL_θ denote a limited liability rent upon θ . Contrary to IP, EH requires for the principal to pay the subordinates upon the realized outcome.

Recall that the principal needs to pay the subordinate only if his task succeeds. To determine v , recall (4.3) and the convex production function, thus $Pr(X_S|I = c) - Pr(X_S|I = 0) \equiv \Delta$ is,

$$\Delta = \frac{p - q}{2} \left(\underbrace{f(1 + H) - f(L)}_{\text{Intrinsic incentive}} + \underbrace{p(f(H) - f(L))}_{\text{Negative externality}} + \underbrace{(1 - p)(f(1 + H) - f(1 + L))}_{\text{Positive externality}} \right)$$

Hence, $v = \frac{c}{\Delta}$. Since the realized skill affects the likelihood of success, the subordinates have intrinsic incentive in acquiring high skill. This incentive is further strengthened by externality by the competitor subordinate's skill: negative externality occurs due to the competitor's high skill, and positive externality occurs due to the competitor's low skill. Then, the expected compensation cost is,

$$E^{EH}[v] = \eta^S v$$

The following result summarizes the above discussion.

Proposition 10. *If $f(\cdot)$ is convex, then an optimal compensation contract for an external hire is*

$w_S = \frac{d}{f_{\theta_1}^U - f_{\theta_1}}$ *for a realized θ_1 , and for an internally promoted superior,*

$$w_S = \max \left\{ \frac{d}{f_L^U - f_L}, \frac{\eta^S}{\xi^S} \frac{2c}{p - q} \right\}$$

where η^i, ξ^i denotes the probability of getting outcome i , and the probability of getting slack upon outcome i respectively.

4.4.2 Organizational Efficiency

Firm Performance: Internal Promotion vs External Hiring

Lemma 15. *Suppose $f(\cdot)$ is convex. Total compensation cost of internal promotion is given by*

$$E^{IP}[w_S] = \eta^S \max \left\{ \frac{d}{f_L^U - f_L}, \frac{2c}{\xi^S(p-q)} \right\}$$

and the total cost of external hiring is given by

$$E^{EH}[w_S + v] = \eta_H^S LL_H + \eta_L^S LL_L + d + \eta^S \frac{c}{\Delta}$$

Proposition 11. *Suppose that $f(\cdot)$ is convex and spillover effect holds (i.e. management slack is sufficient to create investment incentive). Then, external hiring strictly dominates internal promotion if,*

$$\frac{\eta_{HH}^S + \eta_{HL}^S}{\eta^S} (w^{IP} - w_H^{EH}) > \frac{2c}{\lambda(p-q)} \quad (4.5)$$

where $\lambda = \frac{2\Delta}{p-q}$. On the other hand, if spillback effect holds, then if

$$w^{IP} - \left(\frac{\eta_{HH}^S + \eta_{HL}^S}{\eta^S} w_H^{EH} + \frac{\eta_{LL}^S}{\eta^S} w_L^{EH} \right) > \frac{2c}{\lambda(p-q)}$$

then external hiring dominates internal promotion.

Comparative Statics

This subsection investigates when the demand for external hiring increases, why external hire seems to be paid more, and how external hiring firm's performance seems to suffer in spite of the higher pay.

Demand for External Hiring: Recall that p measures difficulty in acquiring firm-specific skills. Likewise, $f(1 + \theta) - f(\theta)$ measures management impact for a fixed skill θ , and $f(H) - f(L)$ measures skill impact. One can naturally examine when the demand for external hiring changes depending on these parameters. The next proposition identifies such conditions under which external hiring tends to be preferred as these parameters change.

Proposition 12. *Conditional on that $f(\cdot)$ is convex, as (i) firm-specific skill tends to be easy to acquire (p high), (ii) the convexity of the production technology becomes greater, then external hiring is more preferred to internal promotion.*

The intuition is as follows. Broadly, if spillover effect works, then incentive aggregation generally dominates incentive separation. However, somewhat counter-intuitively, as p increases, external hiring is preferred. The reason is that, as it becomes easy to acquire firm-specific skill, the precommitment to generate management slack tends to be too expensive relative to its benefit. Hence, in this case, the principal finds it optimal to separate the incentive problems, and rely on the externality made by management allocation to create incentives for subordinates. This inefficiency is strengthened further when the spillback effect works, since the required promotion bonus is already expensive. The Figure 4.2 depicts this finding.

Pay Premium for External Hire: Notice that simple comparison of an external CEO's pay to an internal CEO's pay might not be a fair comparison as the external CEO implements management allocation only, while the internal CEO should do both management effort allocation and investment: i.e the external CEO's total effort cost is d , while the internal CEO's total cost is

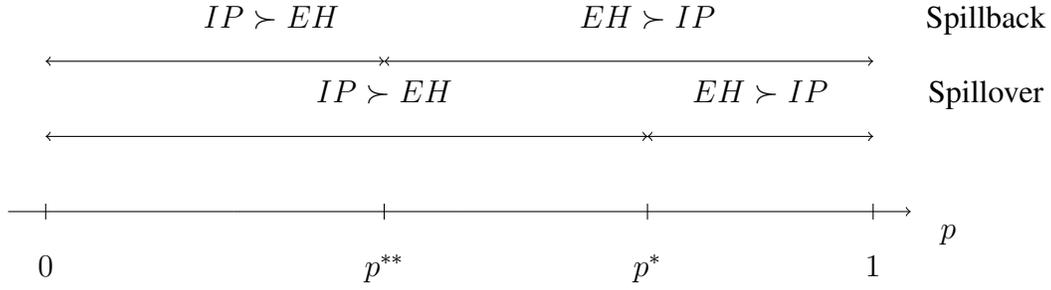


Figure 4.2: **Dominance of CEO Appointment Choice Contingent on p**

Figure 4.2 identifies when internal promotion dominates external hiring or vice versa depending on the difficulty of firm-specific skill acquisition conditional on that production technology is convex. p^*, p^{**} are defined in the appendix.

$c + d$. To make a fair comparison, define the following:

$$v \equiv \text{Pay premium} = \frac{\text{expected total pay}}{\text{total effort costs}} = \frac{E^{IP}[w] \mathbf{1}_{IP} + E^{EH}[w]}{c \cdot \mathbf{1}_{IP} + d}$$

where $\mathbf{1}_{IP}$ denotes an indicator variable that has 1 if IP and 0 otherwise. The basis for the above premium measure comes from the fact that the natural logarithm is often used in the CEO compensation literature (see Murphy and Zabojnik (2007)).¹³ Then, the pay premium for the external hire, v^{EH} , is,

$$v^{EH} = d \frac{\eta_{\theta}^S}{f(1+\theta) - f(\theta)} \frac{1}{d} = \frac{\eta_{\theta}^S}{f(1+\theta) - f(\theta)}$$

Similarly,

$$v^{IP} = d \frac{\eta^S}{f(1+L) - f(L)} \frac{1}{c+d} \quad (\text{if spillover})$$

$$= \frac{\eta^S}{\xi^S} \frac{2c}{p-q} \frac{1}{c+d} \quad (\text{if spillback})$$

¹³I assume that $\beta \ln Sales$ in Murphy and Zabojnik (2007) is a proxy for required effort costs.

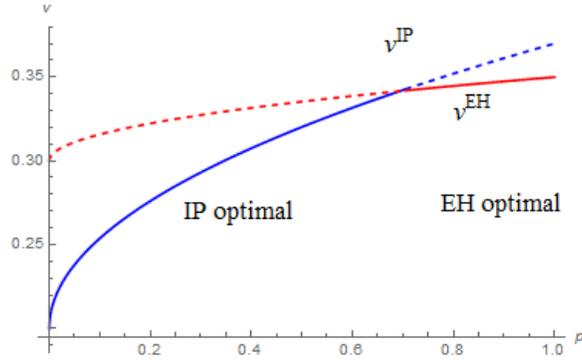


Figure 4.3: Overpayment Illusion

This figure depicts v^{IP} , v^{EH} , the pay premium for each CEO, conditional on that a production function is convex and that spillover effect works. The solid line denotes a pay premium in equilibrium and the dashed line denotes a counterfactual pay premium which is unobservable in equilibrium.

That is, if $v^{EH} > v^{IP}$ is observed, then this will be interpreted as the external CEO gets paid more relative to internal CEO. Combined with the Proposition 11, the following result shows that the pay which is apparently determined optimally might indicate the outsider pay premium that has been observed in the literature.

Proposition 13. (*Implications for outsider pay premium*): *Suppose that the convex production technology is the same across firms and that the spillover effect works. Then the external CEO appears to get paid more relative to the internal CEO. This over-payment illusion is greatest if (L, L) and increases as the convexity of the production technology increases.*

The intuition is as follows. Given that IP dominates EH for small p , and vice versa for large p . If the spillover holds, the principal enjoys the benefit from incentive aggregation, thus decreasing rents to the internal CEO. Since this benefit is only applicable for small p where $v^{EH} > v^{IP}$ is more likely, this might lead to conclusion that the outsider gets paid more, which is not actual

the case. Apparently this is less likely if the spillback case holds since IP already requires high promotion bonus, thus making over-payment illusion for outsider less likely.

Proposition 14. *(Implications for firm performance): Suppose that the production technology is convex. As convexity increases, the principal prefers EH to IP. Then, the performance of an external hiring firm appears to suffer relative to the performance of an internal promoting firm.*

The intuition is again related to the Proposition 11. On the technical side, as the impact of firm-specific skills on the production technology increases, the first term of the threshold (4.5) increases and the second term decreases. Since IP dominates EH for small p , the second term effect is greater than the first term, thus overall effect of the skill is negative. On the intuition side, as skill impact increases, thus the final outcome succeeds more likely, the benefit of incentive aggregation increases, thus making IP dominant. This simultaneously leads to X_S more achievable, thus the performance of the internal promoting firm appears to be more superior than the performance of the external hiring firm. i.e. if success is more likely upon $\theta = H$ and $m = 1$, then IP is more preferred. Therefore, the puzzling evidence on pay and firm performance of external CEO relative to internal CEO can be understood as different firms' different optimal response to their CEO appointment choice.

4.5 Empirical Implications

The main findings of the model in this paper implies that the observed, and often puzzling, practice inherent in CEO appointments, pay and firm performance can be explained by the optimal contracting approach for multiagent: Given that a firm optimally chooses its CEO appointment

choice to overcome contract incompleteness, the pay level and corresponding firm performance are endogenously interlinked with organizational strategy which is also endogenously connected with firm-specific skill parameters.

In reality, however, the only observable variables are the hiring choice, pay and firm performance, but firms' real productivity, thereby we might end up concluding that the pay premium for outsider does not lead to better performance. Furthermore, even if we examine firms in similar industries, which suggests that the issue of omitted firm productivity variable is partly resolved, the direct comparison of those similar industry firms to investigate the association between CEO appointments, pay and performance still remains unsatisfactory since firm-specific skill parameters are hardly observed, thus facing a measurement error problem. To test the predictions of the model, therefore, it is crucial to tackle these econometric challenges in order to incorporate incentive aspects of different hiring structures, which is necessary to better understand the association among CEO appointments, pay and performance.

As a way to overcome the measurement issue of firm-specific skills, the main predictions of the model could be tested using short panel data if the unobservable productivity is almost perfectly estimated by the trend of historical firm performance. To explain more specifically, assume that a researcher collects the data set of one group of internal promoting firms and one group of external hiring firms. Then, according to the model, the probability of restructuring that captures the focused management resource allocation should be higher in firms exploiting external hiring. Many empirical research have found that a newly hired CEO is likely to conduct reshaping an organization including divestiture (See Weisbach (1995), Denis and Denis (1995), Custódio et al. (2013) for example). Moreover, the convexity of the productivity is less likely in

firms exploiting internal promotion.

4.6 Conclusion

I view this paper as a small but significant first move toward a comprehensive theory of CEO appointments. The essential idea of this paper is to consider a CEO hiring choice as an organizational incentive device for multiple agents to overcome contract incompleteness inherent in unverifiable and nontransferable firm-specific skills. Based on the multiagent incomplete contracting viewpoint, this holistic approach provides a way to fill the missing link between firm performance and the premium involved in external hiring, thereby leading us to a better understanding of an existing and puzzling practice. This paper shows that the seemingly higher pay level for an externally hired CEO might not be a *de facto* premium, and relatively mediocre firm performance is interlinked with the realized firm-specific skills which determine the optimal management decision.

A simplifying, but essential, assumption in this paper is that the principal can fully commit to a hiring structure at the initial contracting stage. A natural extension would be to relax the commitment assumption by allowing the principal to make a hiring choice after period 1 ends. However, inasmuch as the promotion bonus is contractible, the ex ante hiring choice of internal promotion is self-enforcing for the principal. This is because it is also ex post optimal given the productivity condition under which the internal promotion is ex ante optimal. Similarly, the ex ante choice of external hiring remains optimal ex post since the promotion bonus and management lottery will never be optimal for the principal after the skills are realized, i.e. the principal has no incentive to introduce internal promotion after period 1 ends as long as the wage

for subordinates is contractible.

In future work, the current framework could be desirably broadened by endogenizing the number of tasks for each hiring structure. As the number of tasks increases, not only does the efficiency of promotion decrease, but also the discipline effect of management substitution decreases. Therefore, it is not obvious to conclude which hiring choice dominates the other under the setting of more than two tasks. Endogenizing the number of tasks based on the current model could be done by comparing different incentive trade-offs of each hiring structure depending on the firm-size. This extension could provide another interesting implication on hiring choices across different firm sizes.

Appendices

Appendix A

Appendix for Chapter 1

A.1 Heterogeneous Reputation Endowment

This section will program the same optimal assignment problem in the main analysis when the initial distribution of reputation is defined over a continuous support. That is, now agents differ in their initial reputation level indexed by $\gamma \in [\underline{\gamma}, \bar{\gamma}]$ with a well-defined smooth distribution function $G_0(\gamma)$ with a corresponding density $g_0(\gamma)$ at the beginning of period 1, where $S_{min} > 0$, $\underline{\gamma} \geq 0$, and $1 < M(\bar{\gamma} - \underline{\gamma})$, implying that there are more CEO candidates than firms. Let $g_t(\gamma)$ denote the distribution of reputation at the end of period t after the update of every CEO in the market.

To program this continuous case, I first formulate the transition of the reputation distribution, and then program the three constraints for the equilibrium assignment outcome in each period.

A.1.1 Transition of Reputation Distribution

Recall that upon a firm performance outcome i , γ^i is fully computed ex ante. However, the rank is not fully known as the ranking order requires a computation of others' performance. Let $f(\gamma) = Pr(\tau = G|h) = \gamma^h$, $g(\gamma) = Pr(\tau = G|l) = \gamma^l$, and upon m , the reputation stays the same. For each γ , since f, g are uniquely defined, we can find f^{-1}, g^{-1} . Then, for a given distribution function $G : \Gamma_{t-1} \rightarrow [0, 1]$, the posterior distribution becomes,

$$G_t(\gamma) = \int_{\underline{\gamma}}^{\gamma} \sum_{x \in \{f^{-1}(s), g^{-1}(s), s\}} tr(s|x) g_{t=1}(x) ds \quad (\text{A.1})$$

where $tr(s|x)$ denotes a transition function from x to s . i.e. after period 1 and period 2 firm performance outcomes, we have $G_1(\gamma)$ and $G_2(\gamma)$ characterized as (A.1).

A.1.2 An Optimal (Re)Assignment Problem

This section provides a formulation of an optimal assignment problem as in the baseline model. Since every model ingredient is the same but the continuous initial density, the constraints that determine the equilibrium remain the same. Observe that the period 2 equilibrium defining constraints is exactly the same as the baseline model.

$$Y_2(S, \gamma) - \omega_2(\gamma) \geq Y_2(S, \gamma') - \omega_2(\gamma') \quad \forall S, \gamma \quad (\text{SC}(S, \gamma))$$

$$Y_2(S, \gamma) - \omega_2(\gamma) \geq y_0 \quad \forall S \quad (\text{PC-firm})$$

$$\omega_2(\gamma) \geq w_0 \quad \forall \gamma \quad (\text{PC-CEO})$$

where y_0, w_0 denotes the firm's and the CEO's reservation utility respectively. Then as in Terviö (2008), the market value for a given γ -CEO is determined as follows.

$$\omega_2(\gamma) = w_0 + \int_{\underline{\gamma}}^{\gamma} \frac{\partial}{\partial z} Y_2(S, z) g_2(z) dz$$

Then, given this market value, the contract for those hired CEOs is characterized as Lemma 3. Consider a γ -CEO who has been hired, and is certain to be hired in period 2 even with l . Taking account for period 2 market value, his total expected payoff in period 1 upon putting forth effort is,

$$\begin{aligned} & \omega_1(\gamma) + \sum_{X \in \{h, m, l\}} Pr(X|\gamma) E[\tilde{\omega}_2(\gamma^X)] \\ &= \omega_1(\gamma) + w_0 + \sum_{X \in \{h, m, l\}} Pr(X|\gamma) E \left[\int_{\underline{\gamma}}^{\gamma^X} \frac{\partial}{\partial x} Y_1(S, x) g_2(z) dz \right] \\ &= \omega_1(\gamma) + w_0 + \sum_{X \in \{h, m, l\}} Pr(X|\gamma) E \left[\int_{\underline{\gamma}}^{\gamma^X} \frac{\partial}{\partial x} Y_t(S, x) \left(\frac{d}{dz} \int_{\underline{\gamma}}^z \sum_{x \in \{f^{-1}(s), g^{-1}(s), s\}} tr(s|x) g_{t=1}(x) ds \right) dz \right] \end{aligned}$$

where the last equality is due to A.1. Taking into account the above expected payoff with (IC) constraint, the market value for reputation in period 1 and the assignment equilibrium are characterized by

$$Y_1(S, \gamma) - \omega_1(\gamma) \geq Y_1(S, \gamma') - \omega_1(\gamma') \quad \forall S, \gamma \quad (\text{SC}(S, \gamma))$$

$$Y_1(S, \gamma) - \omega_1(\gamma) \geq y_0 \quad \forall S \quad (\text{PC-firm})$$

$$\omega_1(\gamma) + \sum_{X \in \{h, m, l\}} \Pi(\gamma^X) \geq w_0 + \Pi(\gamma) \quad \forall \gamma \quad (\text{PC-CEO})$$

where $\Pi(\gamma^X) = w_0 + E \left[\int_{\underline{\gamma}}^{\gamma^X} \frac{\partial}{\partial x} Y_1(S, x) \left(\frac{d}{dz} \int_{\underline{\gamma}}^z \sum_{x \in \{f^{-1}(s), g^{-1}(s), s\}} tr(s|x) g_{t=1}(x) ds \right) dz \right]$. In the case of both (PC-CEO), (IC) binding, the market value determined by (SC) might be replaced by the market value determined by (PC-CEO).

A.2 Proofs

A.2.1 The Derivation of Event Probability

Proof. Notice that upon getting $\hat{v} = r$, the best response is to take I_r , and upon getting $\hat{v} = s$, the best response is to take I_s .

$$\begin{aligned} Pr(h|\tau = B) &= Pr(v = r) \times Pr(\hat{v} = r, (I_r, h)|v = r, \tau = B) = Pr(h|I_r, v = r) \times Pr(v = r)Pr(\hat{v} = r, |v = r) \\ &= p\alpha\beta \end{aligned}$$

$$\begin{aligned} Pr(l|\tau = B) &= Pr(v = r) \times Pr(\hat{v} = r, (I_r, l)|v = r, \tau = B) + Pr(v = s) \times Pr(\hat{v} = r, (I_r, l)|v = s, \tau = B) \\ &= Pr(l|v, I_r) \times Pr(v = r)Pr(\hat{v} = r, |v = r, \tau = B) \\ &\quad + Pr(l|v = s, I_r) \times Pr(v = s)Pr(\hat{v} = r|v = s, \tau = B) \\ &= (1 - p)\alpha\beta + 1 \times (1 - \alpha)(1 - \beta) \end{aligned}$$

$$\begin{aligned} Pr(m|\tau = B) &= Pr(v = s) \times Pr(\hat{v} = s, I_s|v = s, \tau = B) + Pr(v = r) \times Pr(\hat{v} = s, I_s|v = r, \tau = B) \\ &= (1 - \alpha)\beta + \alpha(1 - \beta) \end{aligned}$$

□

A.2.2 Parameter Spaces for Assumption

Proof. The first and the second inequalities are rearranged as follows.

$$Pr(v = r|\hat{v} = r, \gamma)p(h - l) > m - l$$

$$(1 - Pr(v = s|\hat{v} = s, \gamma))p(h - l) < m - l$$

Rearranging the terms with respect to $\frac{m-l}{p(h-l)}$, then,

$$Pr(v = r|\hat{v} = s, \gamma) < \frac{m-l}{p(h-l)} < Pr(v = r|\hat{v} = r, \gamma)$$

Or, equivalently,

$$1 - \frac{1}{1 + \frac{(1-\gamma)(1-\beta)}{\gamma+(1-\gamma)\beta} \frac{\alpha}{1-\alpha}} < \frac{m-l}{p(h-l)} < \frac{1}{1 + \frac{(1-\gamma)(1-\beta)}{\gamma+(1-\gamma)\beta} \frac{1-\alpha}{\alpha}}$$

The third inequality is identical with

$$\frac{(S \times m - w_t^m) - (S \times l - w_t^l)}{(S \times h - w_t^h) - (S \times l - w_t^l)} > \frac{(S \times m - w_t^m) - (S \times l - w_t^l) + Pr(h)(x - (\gamma + (1-\gamma)\beta))}{Pr(\hat{v} = s|\gamma) + x}$$

The fourth inequality is identical with

$$\frac{(1-\gamma)\alpha(1-\beta)}{\gamma(1-\alpha) + (1-\gamma)((1-\alpha)\beta + \alpha(1-\beta))} < \frac{S \times m - w_t^m}{p(S \times h - w_t^h) + (1-p)(S \times l - w_t^l)}$$

□

A.2.3 The Derivation of Reputation Threshold

Proof. Compare the expected market value when the CEO exerts effort to the expected market value when the CEO takes I_s without effort. For simplicity, I omit the cost of effort.

$$\gamma(\alpha(U - D) + \omega_3(\gamma^m)) + (1-\gamma)\left(\beta(\alpha(U - D) + \omega_3(\gamma^m)) + (1-\beta)(\alpha D + \omega_3(\gamma^l))\right) > \omega_3(\gamma^m)$$

where $U = p(\omega_3(\gamma^h) - \omega_3(\gamma^l))$, $D = \omega_3(\gamma^m) - \omega_3(\gamma^l)$. If the above inequality is held, then CEO γ prefers to take the risky project ex ante. Rearranging the terms with respect to γ yields that,

$$\gamma > \frac{1}{1-\beta} \frac{(1-\alpha)D}{\alpha U + (1-2\alpha)D} - \frac{\beta}{1-\beta} \equiv R^*$$

which decreases in U and increases in D . Therefore, as upside potential increases, R^* decreases, and as downside potential increases, R^* decreases. □

A.2.4 Proof of Lemma 1

Proof. Recall that for a given γ , the CEO is either $\tau = G$ with probability γ , or $\tau = B$ with probability $1 - \gamma$. If it is $\tau = G$ and the CEO is incentivized to exert effort, then the expected revenue for a given S -size firm is,

$$S \times (Pr(h|\tau = G)h + Pr(l|\tau = G)l + Pr(m|\tau = G)m) \equiv E[R|S]$$

Here, I omit the time index t . Meanwhile, if it is $\tau = B$, then

$$S \times (\alpha Pr(h|\tau = G)h + \alpha(1 - Pr(h|\tau = G))l + Pr(m|\tau = G)m) = E[R|S] - Pr(h|\tau = G)Pr(m|\tau = G)(h - l)S$$

Thus, upon hiring γ -CEO, the expected revenue is,

$$Y_t(S, \gamma) \equiv E[R|S] - (1 - \gamma)Pr(h|\tau = G)Pr(m|\tau = G)(h - l)S \quad (\text{A.2})$$

It is straightforward to see that $\frac{\partial^2 Y}{\partial \gamma \partial S} > 0$. □

A.2.5 Proof of Lemma 2

Proof. Without loss of generality, I will use $\omega_t(\gamma)$ to denote the price to hire CEO γ . If there is rent to induce $e = H$, then the expected total compensation cost will be $\omega_t(\gamma) + \text{rent}$, which will be the case for low $\omega_t(\gamma)$. Consider the smallest firm S that indiffers between γ_1 and γ_2 where $\gamma_1 > \gamma_2$. i.e., $S = S^M(\gamma_1)$.

$$Y_t(S, \gamma_1) - \omega_t(\gamma_1) = Y_t(S, \gamma_2) - \omega_t(\gamma_2)$$

which yields $\omega_t(\gamma_1) = Y_t(S, \gamma_1) - Y_t(S, \gamma_2) + \omega_t(\gamma_2) = (\gamma_1 - \gamma_2)Pr(h|\tau = G)Pr(m|\tau = G)(h - l)S + \omega_t(\gamma_2)$. This way, the reputation based market value for γ_i is characterized as,

$$\omega_t(\gamma_i) = (\gamma_i - \gamma_{i+1})Pr(h|\tau = G)Pr(m|\tau = G)(h - l)S^{M(\gamma_i)} + \omega_t(\gamma_{i+1})$$

If the matching clears at the tier of γ_n , then the market value for the most reputable CEOs is,

$$\omega_t(\gamma_1) = \sum_{i=1}^{n-1} (\gamma_i - \gamma_{i+1})Pr(h|\tau = G)Pr(m|\tau = G)(h - l)S^{M(\gamma_i)} + \omega_t(\gamma_n)$$

□

A.2.6 Proof of No-randomization between a risky and a safe project

Proof. Suppose that in equilibrium, the CEO exerts effort and chooses an investment project contingent on the observed signal (This will be verified in the lemmas in Section 2.3.2). If a γ -reputable G -type CEO considers a deviation by mixing over I_r, I_s by $(\epsilon, 1 - \epsilon)$ without observing a profitability signal, then, the agent solves the following. For simplicity, assume the reputation incentive is the only concern. The monetary incentive can be easily checked based on the same logic.

$$\epsilon(Pr(h|\gamma)\gamma^h + (1 - Pr(h|\gamma))\gamma^l) + (1 - \epsilon)\gamma$$

Observe that since in equilibrium no manager mixes, the updated reputation stays the same, implying that the above reputation change is linear in ϵ . Since, this is linear in ϵ , if $Pr(h|\gamma)\gamma^h + (1 - Pr(h|\gamma))\gamma^l - \gamma > 0$, then the agent will choose $\epsilon = 1$, if $Pr(h|\gamma)\gamma^h + (1 - Pr(h|\gamma))\gamma^l - \gamma < 0$, then $\epsilon = 0$. Only if $Pr(h|\gamma)\gamma^h + (1 - Pr(h|\gamma))\gamma^l - \gamma = 0$, the γ -reputable CEO indiffers.

Since both γ^h, γ^l are one-to-one mapping of γ , such γ can be uniquely found by solving,

$$Pr(h|\gamma) \frac{\gamma}{\gamma + (1 - \gamma)\eta} + (1 - Pr(h|\gamma)) \frac{\gamma}{\gamma + (1 - \gamma)\eta} = \gamma$$

Thus, $\gamma^{mixing} = \frac{-m + \sqrt{m^2 - 4nl}}{2n}$, where $m = 1 + \xi - 2\eta\xi - Pr(h|\gamma)(\xi - \eta)$, $n = (1 - \eta)(1 - \xi)$, $l = \eta(1 + \xi) + Pr(h|\gamma)(\xi - \eta)$ □

A.2.7 Proof of Lemma 3

Proof. I omit time subscript $t = 3$. Rearrainging (IC) constraint is the following.

$$\begin{aligned} Pr(h|\gamma)(w^h - w^l) - \alpha(w^m - w^l) &\geq c, \\ (\alpha p - Pr(h|\gamma))(w^h - w^l) + (1 - \alpha)(w^m - w^l) &\geq c \end{aligned}$$

Rearrange the terms, then

$$\frac{Pr(h|\gamma)}{Pr(l|\gamma)}(w^h - w^m) - \frac{c}{Pr(l|\gamma)} \geq w^m - w^l \geq \frac{Pr(h) - Pr(h|\gamma)}{Pr(m|\gamma)}(w^h - w^l) + \frac{c}{Pr(m|\gamma)}$$

This implies that w^m is bounded (both below and above). Knowing that at least one of these conditions are binding, we have, either $w^h - w^l = \frac{c}{\gamma\alpha p(1 - \alpha)}$ and $w^m - w^l \geq \frac{c}{\gamma(1 - \alpha)}$, or $w^h - w^l \geq \frac{c}{\gamma\alpha p(1 - \alpha)}$ and $w^m - w^l = \frac{c}{\gamma(1 - \alpha)}$. Plug one of these into (IR) and rearrange. Then,

$$\begin{aligned} Pr(h|\gamma)(w^h - w^l) + w^m - \alpha(w^m - w^l) - c &\geq \omega(\gamma) \\ \Leftrightarrow \alpha p(\gamma + (1 - \gamma)\alpha) \frac{c}{\gamma\alpha p(1 - \alpha)} - \alpha \frac{c}{\gamma(1 - \alpha)} + w^m &\geq \omega(\gamma) + c \\ \Leftrightarrow w^m \geq \omega(\gamma) \text{ if } w^m - w^l \geq \frac{c}{\gamma(1 - \alpha)}, w^m \leq \omega(\gamma) &\text{ otherwise.} \end{aligned}$$

Plug w^m into binding (IC), then,

$$w^l \leq \omega(\gamma) - \frac{1}{\gamma} \frac{c}{1-\alpha}, w^h = \omega(\gamma) + \frac{1}{\gamma} \frac{c}{1-\alpha} \left[\frac{1}{\alpha p} - 1 \right] \text{ if } w^m - w^l \geq \frac{c}{\gamma(1-\alpha)}$$

$$w^l = \omega(\gamma) - \frac{1}{\gamma} \frac{c}{1-\alpha}, w^h \geq \omega(\gamma) + \frac{1}{\gamma} \frac{c}{1-\alpha} \left[\frac{1}{\alpha p} - 1 \right] \text{ otherwise.}$$

Since agents are risk neutral, there can be multiple solutions that satisfy the above conditions.

For simplicity, I focus on a solution that binds (IC), then

$$\frac{w^h - w^m}{w^m - w^l} = \frac{1 - \alpha p}{\alpha p}$$

□

A.2.8 Proof of Lemma 4

Proof. This is similar with Lemma 3 except that now each term has dynamic incentive. Observe that although we have extra terms, still we are interested in finding three variables (w^h, w^m, w^l) with three constraints given the expectation value of future reputation, $\omega_3(\gamma^h), \omega_3(\gamma^m), \omega_3(\gamma^l)$.

Again, focusing on a solution that binds (IC), we have,

$$w_2^h = \omega_2(\gamma) + \frac{c(1-\alpha p)}{(1-\alpha)\alpha p \gamma} - (1 - Pr(h|\gamma))\omega_3(\gamma^h) - \omega_3(\gamma^l) + (1-\alpha)\omega_3(\gamma^m) - \omega_3(\gamma^l)$$

$$w_2^m = \omega_2(\gamma) + Pr(h|\gamma)\omega_3(\gamma^h) - \omega_3(\gamma^l) - \alpha\omega_3(\gamma^m) - \omega_3(\gamma^l)$$

$$w_2^l = \omega_2(\gamma) - \frac{c}{\gamma(1-\alpha)} + Pr(h|\gamma)\omega_3(\gamma^h) - \omega_3(\gamma^l) + (1-\alpha)\omega_3(\gamma^m) - \omega_3(\gamma^l)$$

thus, confirming the results. □

A.2.9 Proof of Lemma 5

Proof. Due to supermodularity of expected revenue, the efficiency requires positive assortative assignment of CEOs and firms when there is no dynamic concern. High reputation is formed by the track record of positive performance outcome, implying that the final period assignment pattern exhibits monotone performance induced succession. \square

A.2.10 Proof of Lemma 6

Proof. Consider manager γ . Suppose the (IC) for manager γ binds, but the (PC-manager) not. Within a group of manager γ , the market is competitive. This implies that for $\omega^*(\gamma)$ that binds the (IC), $\omega^*(\gamma) > \omega^{PC}(\gamma)$. Then, another manager γ that is willing to work at $\omega^*(\gamma) - \epsilon$ where $\epsilon > 0$ small, approaches the potential matching partner firms, thus reducing the left hand side of the (PC-manager), which is contradiction to the assumption that the (IC) binds. \square

A.2.11 Proof of Proposition 1

Proof. For convenience, I omit the time subscript $t = 1$. Let $S[i]$ denote a threshold firm in equilibrium. I first show that $\omega^*(\gamma^i) = \omega^{SC}(\gamma^i, S[i])$. Suppose $\omega^*(\gamma^i) > \omega^{SC}(\gamma^i, S[i])$. Then, $S[i]$ will deviate from matching with manager γ^i to matching with next best manager $\gamma' < \gamma^i$. Thus, $\omega^*(\gamma^i) \leq \omega^{SC}(\gamma^i, S[i])$. Now, suppose $\omega^*(\gamma^i) < \omega^{SC}(\gamma^i, S[i])$. Since $\gamma^i > \gamma^l$, then firm $S[i] - \epsilon$, where $\epsilon > 0$ small, will deviate by approaching manager γ^i with an offer of $\omega^*(\gamma^i) + \eta$ where $\eta > 0$ small. Thus $\omega^*(\gamma^i) = \omega^{SC}(\gamma^i, S[i])$.

Next, I show that $\omega^{SC}(\gamma^i, S[i]) \geq \omega^{PC}(\gamma^i)$, and equality holds if $S[i] > G^{-1}(1 - M(\gamma^i))$.

Suppose $\omega^{SC}(\gamma^i, S[i]) < \omega^{PC}(\gamma^i)$. Then, manager γ^i will deviate by sitting out. Thus, $\omega^{SC}(\gamma^i, S[i]) \geq \omega^{PC}(\gamma^i)$. If $S[i] > G^{-1}(1 - M(\gamma^i))$, then, the smallest firm size is greater than the size of the firm that has the same rank as manager γ^i . Suppose $\omega^{SC}(\gamma^i, S[i]) > \omega^{PC}(\gamma^i)$. Then, firm $S[i] - \epsilon$, where $\epsilon > 0$, will deviate by offering a contract to manager γ^i , thus contradiction that $S[i]$ is a threshold firm. Thus, $\omega^{SC}(\gamma^i, S[i]) = \omega^{PC}(\gamma^i)$ if $S[i] > G^{-1}(1 - M(\gamma^i))$. \square

A.2.12 Proof of Lemma 7

Proof. Again, I omit the time subscript for period 1 but keep for period 2. When there is no distortion, the threshold firm for manager γ^m is S_{min} . If manager γ^m strictly prefers the safe project to maintain his current reputation, and for this to create a deviation from S_{min} , it shall be the case that $\omega^{SC}(\gamma^m, S_{min}) < \omega^{PC}(\gamma^m)$. From firm S_{min} 's (SC) constraint, we know that,

$$Y(S_{min}, \gamma^m) - \omega(\gamma^m) \geq Y(S_{min}, \gamma^l) - \omega(\gamma^l) \Rightarrow \omega^{SC}(\gamma^m) \leq \omega(\gamma^l) + \Delta Y(S_{min})$$

Similarly, from manager γ^m 's (PC-manager) constraint, we know that

$$\omega(\gamma^m) - c + E[\omega_2(\gamma^{mX})|X] \geq w_0 + \omega_2^*(\gamma^m) \Rightarrow \omega^{PC}(\gamma^m) \geq c + w_0 + \omega_2^*(\gamma^m) - E[\omega_2(\gamma^{mX})|X]$$

Thus,

$$\omega^{SC}(\gamma^m, S_{min}) < \omega^{PC}(\gamma^m) \Rightarrow \omega(\gamma^l) + \Delta Y(S_{min}) < c + w_0 + \omega_2^*(\gamma^m) - E[\omega_2(\gamma^{mX})|X]$$

where $\Delta Y(S[m]) = \left(Y(S[m], \gamma^m) - Y(S[m], \gamma^l) \right)$, manager γ^m 's reputation premium (relative to manager γ^l) or threshold firm $S[m]$'s marginal productivity loss in case of a deviation.

Plug $\omega_2(\gamma^i) = w_0 + c + \sum_{\gamma^j \leq \gamma^i} \Delta Y(S[j])$ into above and rearrange the terms. Then,

$$Pr(h|\gamma^m)\Delta Y(S[mh]) + \Delta Y(S_{min})w_0 + c < Pr(l|\gamma^m)\Delta Y(S[mm])$$

$$\Rightarrow Pr(h|\gamma^m)U(m) + \omega^{SC}(\gamma^m, S_{min}) < Pr(l|\gamma^m)D(m)$$

where $U(m) = \omega_2(\gamma^{mh}) - \omega_2(\gamma^{mm})$, $D(m) = \omega_2(\gamma^{mm}) - \omega_2(\gamma^{ml})$, the upside and downside potential for manager γ^m respectively. □

A.2.13 Proof of Proposition 3

Proof. Lemma 7 is also applied to manager γ^h : $S[h] = G^{-1}(1 - M(\gamma^h))$ implies that $Pr(h|\gamma^h) \times U(h) + \omega_1(\gamma^h) \geq Pr(l|\gamma^h) \times D(h)$. i.e., $\omega^*(\gamma^h) = \omega^{SC}(\gamma^h, S[h]) \geq \omega^{PC}(\gamma^h)$. Thus, all manager γ^h get matched. Due to Lemma 7, some manager γ^m get matched by the measure of $G(S[m]) - M(\gamma^h)$, and rest of manager γ^m remain unmatched; and all firm $S \in [S_{min}, S[m]]$.

I now show that this is indeed stable. Notice that,

$$Y(S, \gamma^m) - \omega^*(\gamma^m) > Y(S, \gamma^l) - \omega^*(\gamma^l) \quad \forall S \in (S[m], S[h]),$$

$$Y(S, \gamma^m) - \omega^*(\gamma^m) < Y(S, \gamma^l) - \omega^*(\gamma^l) \quad \forall S \in [S_{min}, S[m]]$$

Thus, there is no profitable deviation by firms. Also, for γ^m , if he/she rejects the offer $\omega(\gamma^m)$ and asks instead $\omega^*(\gamma^m) + \epsilon$ for $\epsilon > 0$, then he/she will not be hired as $\forall S \in [S[m], S[h]]$ can find others at $\omega^*(\gamma^m)$. On the other hand, if he/she accepts $\omega^*(\gamma^m) - \epsilon$, then since their participation constraint binds, sitting out makes them strictly better off.

Let $u(S) = Y(S, \mu(S)) - \omega(\mu(S))$ denote a firm S 's equilibrium payoff in this equilibrium.

Observe that this also guarantees that there is no block that Pareto-improves because for $\forall S, \gamma$,

$$\omega(\gamma) + u(S) = Y(S, \gamma) \text{ if } \mu(S) = \gamma,$$

$$\omega(\gamma) + u(S) > Y(S, \gamma) \text{ if } \mu(S) \neq \gamma$$

Therefore, there is no block that can Pareto-improve. Thus, $\omega^*(\gamma)$ and

$$\mu(S) = \gamma^l \text{ for } S \in [S_{min}, S[m]), \mu(S) = \gamma^m \text{ for } S \in [S[m], S[h]),$$

$$\mu(S) = \gamma^h \text{ for } S \in [S[m], S[h])$$

is an equilibrium in period 1. □

A.2.14 Proof of Lemma 8

Proof. Straightforward from the proof of Lemma 7. □

Appendix B

Appendix for Chapter 2

15 Proof of Lemma 9

Proof. Notice that regardless of the outcome, there is always type 1 and type 2 error. Also, the CEO himself doesn't know his type correctly, thus he must consider both $\tau = G$ and $\tau = B$ cases. For the sake of simplicity, I omit the subscript for time index.

$$E[\gamma|l] = \left(\gamma\alpha(1-p)b + (1-\gamma)\alpha(1-\alpha)(1-b) \right) \gamma_{b+}^l + \left(\gamma\alpha(1-p)(1-b) + (1-\gamma)\alpha(1-\alpha)b \right) \gamma_{b-}^l$$

Rearranging yields the result in the lemma. □

16 Proof of Lemma 10

Proof. To economize on notation, I omit time subscript $t = 3$. To find the unique and effective level $b^* \in (\frac{1}{2}, 1)$, it has to satisfy the following.

$$\left[\gamma b + (1-\gamma)(1-b) \right] (\gamma_{b+}^l - \gamma^l) S^{M(\gamma^l)} = \frac{c - (\omega(\gamma) - w^m) + \alpha\omega(\gamma^m) - Pr(h|\gamma)\omega(\gamma^h)}{FS^{M(\gamma^l)}} \equiv H(\alpha, \gamma)$$

(*)

Rewrite (*), we have the following second degree of polynomial with respect to b .

$$b^2 f(\alpha, \gamma, w_0) - bg(\alpha, \gamma, w_0, H) + h(\alpha, \gamma, w_0, H) = 0$$

where,

$$f(\alpha_1, \gamma, w_0) = (1 - 2\gamma) \left((3\gamma(1 - p) + (1 - \gamma)(1 - \alpha p)) w_0 - 2S^{M(\gamma^l)} \frac{\gamma(1 - p)(1 - \gamma)(1 - \alpha p)}{\gamma(1 - p) + (1 - \gamma)(1 - \alpha p)} \right)$$

$$g(\alpha_1, \gamma, w_0, H) = \frac{1}{\gamma(1 - p) + (1 - \gamma)(1 - \alpha p)} \times \left(H(3\gamma(1 - p) + (1 - \gamma)(1 - \alpha p)(\gamma(1 - p) + (1 - \gamma)(1 - \alpha p)) \right. \\ \left. - (1 - \gamma)((1 - p)(1 - \alpha p)\gamma(-3 + 4\gamma)S^{M(\gamma^l)} + (\gamma(1 - \alpha) + (1 - \gamma)(1 - \alpha p)(7\gamma(1 - p) + (4 - 3\gamma) \right.$$

$$h(\alpha_1, \gamma, w_0, H) = -\frac{1}{\gamma(1 - p) + (1 - \gamma)(1 - \alpha p)} \times (1 - \alpha p)(1 - \gamma)(H(\gamma(1 - p) + (1 - \gamma)(1 - \alpha p) - (1 - \gamma) \\ \left. + (1 - \gamma)(1 - \alpha p))w_0 - \gamma(1 - p)S^{M(\gamma^l)})$$

Solving the above polynomial yields two real number values, one less than 0 and another one is greater than 0. Since $b < 1$, the unique optimal solution is characterized as $b = \frac{g + \sqrt{g^2 - 4fh}}{2f}$, here arguments for each functional form of coefficient are omitted.

Due to the competitive equilibrium pay (Lemma 2), as the equilibrium pay for those with low reputation increases (decreases), the pay level of high reputable CEOs increases (decreases). Introducing extra insurance from board' monitoring intensity weakens the required premium for those who have reputation maintaining incentives, thus reducing the pay for CEOs in above that level. \square

17 Proof of Proposition 6

Proof. To highlight the effect from market condition, let $b(\alpha, \gamma^m)$ denote the optimal level of board competency as a function of α and the current reputation of the incumbent CEO. Recall

that upon a poor performance but with a board's signal of s_G , the incumbent CEO's reputation will be $\gamma_{B+}^{ml} = \left(1 + \frac{1-\gamma^m}{\gamma^m} \left(\beta + \frac{1-\alpha}{\alpha} \frac{1-\beta}{1-p}\right) \frac{1-b(\alpha_1, \gamma)}{b(\alpha_1, \gamma)}\right)^{-1}$. For γ_{B+}^{ml} to be preferred to γ^m , it should be the case that,

$$\left(\beta + \frac{1-\alpha}{\alpha} \frac{1-\beta}{1-p}\right) \frac{1-b(\alpha_1, \gamma)}{b(\alpha_1, \gamma)} \leq 1$$

Here the effect of m is canceled out as both γ_{B+}^{ml} and γ^m share that history. The above is equivalent to $b(\alpha, \gamma) \geq \frac{\beta + \frac{1-\alpha}{\alpha} \frac{1-\beta}{1-p}}{\left(\beta + \frac{1-\alpha}{\alpha} \frac{1-\beta}{1-p} + 1\right)} = \frac{\frac{Pr(l|\tau=B)}{Pr(l|\tau=G)}}{\frac{Pr(l|\tau=B)}{Pr(l|\tau=G)} + 1}$. □

18 Proof of Lemma 11

Proof. Recall that the demand for insurance occurs only for those $S \in [S^{M(\gamma^m)}, S^{M(\gamma^h)}]$, which corresponds to rank of $M(\gamma^h)$ to $M(\gamma^m)$. Now, after the period 2 performance is realized, the order of ranks of CEOs is as follows.

$$\gamma^{hh} > \gamma^{hm} > \gamma^{0h} > \gamma^{hl} > \gamma^{mm} > \gamma^{0m} > \gamma^{ml} > \gamma^{0l} > \gamma^{ll}$$

That is, if a firm size S ranks above $M(\gamma^{lh})$ (i.e., $S \geq G^{-1}(M(\gamma^{lh}))$), implying that a newly rematched CEO has reputation greater than γ^{mm} . Then, upon replacement without disclosing performance outcome, it can be either m or l , thus preventing the exact outcome from being inferred. □

19 Proof of Lemma 12

Proof. Observe that the mismatch can happen only for those firms within $[S^{M(\gamma^{ND})}, S^{M(\gamma^m)}]$. If performance outcome were observable, then this group is divided into large ($S \in [S^{M(\gamma^{mm})}, S^{M(\gamma^m)}]$)

and small ($S \in [S^{M(\gamma^l)}, S^{M(\gamma^{mm})}]$) respectively. The mismatch is defined as when a large firm matches with γ^{ml} (or equivalently, when a firm matches with γ^{mm}). Thus, whenever γ^{mm} is assigned to a small firm S rather than a large firm S' , the marginal loss is,

$$Y_3(S, \gamma^{mm}) - Y_3(S', \gamma^{mm}) = (S - S') \times \left(E[R] - (1 - \gamma)F^{\gamma^{mm}} \right) < 0$$

This is because $Y_3(S, \gamma) = S \times \left(E[R] - (1 - \gamma)F^\gamma \right)$. Let \mathcal{M} denote a mismatch likelihood.

Observe that

$$\begin{aligned} \mathcal{M} &= Pr(\text{mismatch}) = Pr(\gamma^{mm} \text{ is selected})Pr(\text{assigned to small } S|\gamma^{mm}) \\ &= Pr(\gamma^{ml} \text{ is selected})Pr(\text{assigned to large } S|\gamma^{ml}) \\ &= \pi_m(1 - \pi_m) = \pi_l(1 - \pi_l) \end{aligned}$$

where $\pi_m = \frac{Pr(m|\gamma)}{Pr(m|\gamma) + Pr(h|\gamma)}$. The last equality is due to $\pi_l = 1 - \pi_m$. The likelihood of being misassigned is determined by π_m, π_l . This is because, within a group of $[S^{M(\gamma^l)}, S^{M(\gamma^m)}]$, the measure of π_m firms (from the top within this group) is considered large, and the rest of them as small. Since this allocation process is independent of which CEO is being selected, the likelihood of γ^{mm} being misassigned to a small firm is exactly the likelihood of choosing the small firm within this group, which occurs with probability of $1 - \pi_m$. The same logic applies to the case when γ^{ml} is selected. Hence, $\mathcal{M} = \pi_m\pi_l$. Therefore, the expected mismatch distortion is,

$$\pi_m\pi_l \left(E \left[S | S^{M(\gamma^{ml})}, S^{M(\gamma^{mm})} \right] - E \left[S | S^{M(\gamma^{mm})}, S^{M(\gamma^m)} \right] \right) Y_S(S, \gamma^{mm}) < 0$$

The inequality is due to the average size of large firm is strictly greater than the average size of small firm.

□

20 Proof of Proposition 7

Proof. The existence of assignment distortion is shown in Lemma 12, and the distortion arises only for $\forall S \in [S^{M(\gamma^{ND})}, S^{M(\gamma^m)}]$, and this is shown in Lemma 11. \square

21 Proof of Lemma 13

Proof. Suppose that firm of size S is currently matched with the incumbent CEO γ who is entrenched in reputation maintaining incentive. Assume that the expected matched level of reputation in period 2 is γ^m . Recall that for $EAS > 0$ to create incentive, it shall satisfy the following two conditions.

$$EAS \leq E[\tilde{Y}(S, \gamma^m) - \tilde{\omega}_3(\gamma^m) - \tilde{Y}(S, \gamma^l) + \tilde{\omega}_3(\gamma^l)] \equiv \mathcal{F}(S, \gamma)$$

$$EAS \geq \frac{Extra}{Pr(l|\gamma)}$$

where $Extra = c - (\omega_2(\gamma) - w^m) - Pr(h|\gamma)(E[\tilde{\omega}_3(\gamma^h)] - E[\tilde{\omega}_3(\gamma^l)]) + (1 - Pr(m|\gamma))(E[\tilde{\omega}_3(\gamma^m)] - E[\tilde{\omega}_3(\gamma^l)])$. The first condition requires that ex ante severance takes place in expectation. The second condition is to create incentive upon the presence of severance pay agreement when $Extra$ incentive is required due to reputation maintaining incentive. To have $EAS > 0$ in equilibrium, the feasibility condition requires $\mathcal{F}(S, \gamma) \geq \frac{Extra}{Pr(l|\gamma)}$.

Notice that $\mathcal{F}(S, \gamma)$ increases with S .

$$\begin{aligned} \mathcal{F}(S, \gamma) &= E[\Delta\tilde{Y}(S, \gamma) - (\gamma^m - \gamma^l)F_2^m S^{M(\gamma^m)}] \\ &= S \left(E[R] - (1 - \gamma^m)E[F_3^m] - E[R] + (1 - \gamma^l)E[F_3^l] \right) - (\gamma^m - \gamma^l)E[F_3^m]S^{M(\gamma^m)} \\ &= S \left(\gamma^m E[F_3^m] - \gamma^l E[F_3^l] \right) - (\gamma^m - \gamma^l)E[F_3^m]S^{M(\gamma^m)} \end{aligned}$$

where $E[F_3^i] = E[F_3|\gamma^i]$. Since $\gamma^m > \gamma^l$, $E[F_3^m] > E[F_3^l]$. Thus, as firm size $S \in [S^{M(\gamma^m)}, S^{M(\gamma^h)}]$ increases, the likelihood of severance upon a poor performance is more likely. Notice also that the efficient level of EAS is found at a level that makes (IC) binding. That is, as long as the feasibility is satisfied, $EAS^* = \frac{Extra}{Pr(l|\gamma)}$. Notice that,

$$\begin{aligned} & -Pr(h|\gamma)(E[\tilde{\omega}_3(\gamma^h)] - E[\tilde{\omega}_3(\gamma^l)]) + (1 - Pr(m|\gamma))(E[\tilde{\omega}_3(\gamma^m)] - E[\tilde{\omega}_3(\gamma^l)]) \\ & = Pr(l|\gamma)\Delta\gamma^m E[F_3^m]S^{M(\gamma^m)} - Pr(h|\gamma)\left(\sum_{i \in \{h,hl,lh\}} \Delta\gamma^i E[F_3^i]S^{M(\gamma^i)}\right) \end{aligned}$$

where $\Delta\gamma^i$ denotes the difference between γ^i and the right below of γ^i . Thus,

$$EAS^* = \frac{c - (\omega_1(\gamma) - w^m)}{Pr(l|\gamma)} + \left(\Delta\gamma^m E[F_3^m]S^{M(\gamma^m)} - \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum_{i \in \{h,hl,lh\}} \Delta\gamma^i E[F_3^i]S^{M(\gamma^i)}\right)$$

For EAS^* to be feasible,

$$\begin{aligned} \mathcal{F}(S, \gamma) \geq EAS^* & \Leftrightarrow S\left(\gamma^m E[F_3^m] - \gamma^l E[F_3^l]\right) - \Delta\gamma^m E[F_3^m]S^{M(\gamma^m)} \geq EAS^* \\ & \Leftrightarrow S \geq \frac{(Pr(l|\gamma))^{-1}\left(c - (\omega_1(\gamma) - w^m) - Pr(h|\gamma) \sum_{i \in \{h,hl,lh\}} \Delta\gamma^i E[F_3^i]S^{M(\gamma^i)}\right) + 2\Delta\gamma^m E[F_3^m]S^{M(\gamma^m)}}{\gamma^m E[F_3^m] - \gamma^l E[F_3^l]} \equiv S \end{aligned}$$

Therefore, those sufficiently large firms $S \in \left[\max\left\{S^{EAS}, S^{M(\gamma^m)}\right\}, S^{M(\gamma^h)}\right)$ can credibly use the severance pay as insurance to create incentive in period 1.

Now, I proceed the market volatility condition to claim that if the market volatility is too high, then the feasibility cannot be satisfied. Recall $E[F_3^i]$.

$$E[F_3^i] = Pr(l|\gamma^i)Pr(m|\gamma^i) - \sigma^2 p\left(2(\gamma^i(1 - \alpha_2) + \alpha_2) + \alpha_2(1 - \gamma^i) - 1\right)$$

To see if $\mathcal{F}(S, \gamma) < EAS^*$ can occur given that $EAS^* > 0$, rewrite EAS^* with respect to σ^2

$$\begin{aligned} EAS^* & = -\sigma^2 p\left(\Delta\gamma^m S^{M(\gamma^m)} u(\gamma^m) - \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum \Delta\gamma^i S^{M(\gamma^i)} u(\gamma^i)\right) + \frac{c - (\omega_2(\gamma) - w^m)}{Pr(l|\gamma)} + \Delta\gamma^m S^{M(\gamma^m)} Pr(l|\gamma) \\ & \quad - \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum \Delta\gamma^i S^{M(\gamma^i)} Pr(l|\gamma^i) Pr(m|\gamma^i) \end{aligned}$$

where $u(\gamma^i) = 2(\gamma^i(1 - \alpha_2) + \alpha_2) + \alpha_2(1 - \gamma^i) - 1$. Observe that,

$$EAS^* > 0 \Leftrightarrow \sigma^2 < \sigma^{EAS}$$

where

$$\begin{aligned} \sigma^{EAS} &= \left[p \left(\Delta \gamma^m S^{M(\gamma^m)} u(\gamma^m) - \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum \Delta \gamma^i S^{M(\gamma^i)} u(\gamma^i) \right) \right]^{-1} \\ &\times \left(\frac{c - (\omega_2(\gamma) - w^m)}{Pr(l|\gamma)} + \Delta \gamma^m S^{M(\gamma^m)} Pr(h|\gamma) Pr(m|\gamma) - \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum \Delta \gamma^i S^{M(\gamma^i)} Pr(h|\gamma^i) Pr(m|\gamma^i) \right) \end{aligned}$$

Rewrite $\mathcal{F}(S, \gamma)$ with respect to σ^2 ,

$$\begin{aligned} \mathcal{F}(S, \gamma) &= -\sigma^2 p \left(\gamma^m u(\gamma^m) S - \gamma^l u(\gamma^l) S - \Delta \gamma^m u(\gamma^m) S^{M(\gamma^m)} \right) + S \left(\gamma^m Pr(h|\gamma) Pr(m|\gamma) - \gamma^l Pr(h|\gamma^l) Pr(m|\gamma^l) \right. \\ &\quad \left. - \Delta \gamma^m S^{M(\gamma^m)} Pr(h|\gamma) Pr(m|\gamma) \right) \end{aligned}$$

Similarly, observe that,

$$EAS^* > \mathcal{F}(S, \gamma) \Leftrightarrow \sigma^2 > \sigma^{FEAS}$$

where

$$\begin{aligned} \sigma^{FEAS} &= \left[p \left(\gamma^m u(\gamma^m) S - \gamma^l u(\gamma^l) S - 2\Delta \gamma^m S^{M(\gamma^m)} u(\gamma^m) + \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum \Delta \gamma^i S^{M(\gamma^i)} u(\gamma^i) \right) \right]^{-1} \\ &\times \left(S \left[\gamma^m Pr(h|\gamma) Pr(m|\gamma) - \gamma^l Pr(h|\gamma) Pr(m|\gamma) \right] - 2\Delta \gamma^m S^{M(\gamma^m)} Pr(h|\gamma) Pr(m|\gamma) - \frac{c - (\omega_2(\gamma) - w^m)}{Pr(l|\gamma)} + \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum \right) \end{aligned}$$

Therefore, as long as $\sigma^{FEAS} < \sigma^{EAS}$, EAS^* cannot be used when the market volatility is,

$$\sigma^{FEAS} < \sigma^2 < \sigma^{EAS}.$$

To finalize this argument, the parameter space such that $\sigma^{FEAS} < \sigma^{EAS}$ should exist. This is

indeed true if,

$$\begin{aligned} &\frac{S \left[\gamma^m S^{M(\gamma^m)} Pr(h|\gamma) Pr(m|\gamma) - \gamma^l S^{M(\gamma^l)} Pr(h|\gamma^l) Pr(m|\gamma^l) \right] - \Delta \gamma^m S^{M(\gamma^m)} Pr(h|\gamma) Pr(m|\gamma) - D}{D} \\ &< \frac{S \left[\gamma^m u(\gamma^m) - \gamma^l u(\gamma^l) \right] - \Delta \gamma^m u(\gamma^m) - T}{T} \end{aligned}$$

where $D = \frac{c - (\omega_2(\gamma) - w^m)}{Pr(l|\gamma)} + \Delta \gamma^m S^{M(\gamma^m)} Pr(h|\gamma) Pr(m|\gamma) - \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum \Delta \gamma^i S^{M(\gamma^i)} Pr(h|\gamma^i) Pr(m|\gamma^i)$,
 $T = \Delta \gamma^m u(\gamma^m) S^{M(\gamma^m)} - \frac{Pr(h|\gamma)}{Pr(l|\gamma)} \sum \Delta \gamma^i S^{M(\gamma^i)} u(\gamma^i)$. \square

22 Proof of Proposition 8

Proof. Recall that firms will retain the CEO upon the poor outcome instead of replacing him and paying severance pay, if

$$\begin{aligned} Y_3(s, \gamma^l) - \omega_3(\gamma^l) &> Y_3(s, \gamma^m) - \omega_3(\gamma^m) - EAS \\ \Leftrightarrow \omega_3(\gamma^m) - \omega_3(\gamma^l) &> Y_3(s, \gamma^m) - Y_3(s, \gamma^l) - EAS \\ \Leftrightarrow (\gamma^m F_3^m - \gamma^l F_3^l) S^{M(\gamma^m)} &> (\gamma^m F_3^m - \gamma^l F_3^l) S - EAS \end{aligned}$$

Recall that $F_3^\gamma = Pr(h|\gamma) Pr(m|\gamma) = \alpha_3 p(\gamma(1 - \alpha_3) + \alpha_3)(1 - \alpha_3)$ and observe that,

$$\gamma^m F_3^m - \gamma^l F_3^l = \alpha_3 p(1 - \alpha_3)(\gamma^m - \gamma^l)(\alpha_3 + (1 - \alpha_3)(\gamma^m + \gamma^l))$$

Thus, the distortion occurs if,

$$S^D(\alpha_3) \equiv S^{M(\gamma^m)} + \frac{EAS}{\alpha_3 p(1 - \alpha_3)(\gamma^m - \gamma^l)(\alpha_3 + (1 - \alpha_3)(\gamma^m + \gamma^l))} > S$$

This is more likely as α_3 goes to extreme, and the small firms are more sensitive with respect to this. More formally, given that $\gamma^m + \gamma^l < 1$

$$\begin{aligned} &\frac{\partial}{\partial \alpha_3} \left(\alpha_3 p(1 - \alpha_3)(\gamma^m - \gamma^l)(\alpha_3 + (1 - \alpha_3)(\gamma^m + \gamma^l)) \right) \\ &= -3\alpha_3^2 p(\gamma^m - \gamma^l)(1 - \gamma^m - \gamma^l) - 2\alpha_3 p(\gamma^m - \gamma^l)(1 - 2\gamma^m - 2\gamma^l) + p(\gamma^m - \gamma^l)(\gamma^m + \gamma^l) \\ &= -p(\gamma^m - \gamma^l) \left(3\alpha_3^2(1 - \gamma^m - \gamma^l) + 2\alpha_3(1 - 2\gamma^m - 2\gamma^l) - (\gamma^m + \gamma^l) \right) \\ &= -p(\gamma^m - \gamma^l) \left(\alpha_3(3\alpha_3 + 2) - (\gamma^m + \gamma^l) \left\{ (\alpha_3 + 1)(3\alpha_3 + 1) \right\} \right) \\ &= -p(\gamma^m - \gamma^l) H(\alpha_3, \gamma^m + \gamma^l) \end{aligned}$$

Observe that $H(\alpha_3, \gamma^m + \gamma^l)$ is monotone increasing in α_3 . i.e.

$$H(\alpha_3, \gamma^m + \gamma^l) > 0 \text{ if } \alpha_3 > \frac{1}{3} \sqrt{1 + \frac{1}{(1 - \gamma^m - \gamma^l)^2} - \frac{1}{1 - \gamma^m - \gamma^l} - \frac{1 - 2\gamma^m - 2\gamma^l}{3(1 - \gamma^m - \gamma^l)}} \equiv \xi > 0$$

$H(\alpha_3, \gamma^m + \gamma^l) \leq 0$ otherwise.

Thus, $\frac{\partial}{\partial \alpha_3} \left(\alpha_3 p(1 - \alpha_3)(\gamma^m - \gamma^l)(\alpha_3 + (1 - \alpha_3)(\gamma^m + \gamma^l)) \right)$ starts from positive up until $\alpha_3 = \xi$, then turns to negative onward. i.e. $\left(\alpha_3 p(1 - \alpha_3)(\gamma^m - \gamma^l)(\alpha_3 + (1 - \alpha_3)(\gamma^m + \gamma^l)) \right)$ is single peaked and goes to 0 at the extreme, thus confirming that $S^D(\alpha_3)$ blows up at the boundary.

Now, the presence of distortion is clear from the proof of Lemma 13 and the proof above. \square

Appendix C

Appendix for Chapter 3

C.1 Proofs

A.1 Likelihoods

case 1. $f(\cdot)$ is concave:

$$\eta^{SS} = p^2(f_H^B)^2 + 2p(1-p)f_H^B f_L^B + (1-p)^2(f_L^B)^2$$

$$\begin{aligned}\xi^{SS} &= Pr(X_{SS}, (\theta_H, \theta_H), \text{promoted}) + Pr(X_{SS}, (\theta_H, \theta_L), \text{promoted}) - Pr(X_{SS}, (\theta_L, \theta_L), \text{promoted}) \\ &= \frac{p^2}{2}((f_H^B)^2 - (f_L^B)^2) + p(1-p)(f_H^B f_L^B - (f_L^B)^2)\end{aligned}$$

case 2. $f(\cdot)$ is convex:

$$\eta^S = p^2(f(1+H) + f(H)) + 2p(1-p)(f(1+H) + f(L)) + (1-p)^2(f(1+L) + f(L))$$

$$= f(1+H) + p^2(f(H) - f(L)) - (1-p)^2(f(1+H) - f(1+L)) + f(L)$$

$$\eta_H^S = p^2(f(1+H) + f(H)) + 2p(1-p)(f(1+H) + f(L))$$

$$\eta_L^S = (1-p)^2(f(1+L) + f(L))$$

$$\xi^S = \frac{p(2-p)}{2}(f(1+H) - f(1+L))$$

$$\lambda = (2-p)(f(1+H) - f(H)) + (1+p)(f(H) - f(L))$$

$$= 2(f(1+H) - f(H)) + (f(H) - f(L)) - p(f(1+H) - 2f(H) + f(L))$$

The last inequality of λ is due to the assumption of $f(1+L) = f(H)$.

Limited liability rents:

$$LL_H = \frac{f(H)}{f(1+H) - f(H)}d$$

$$LL_L = \frac{f(L)}{f(1+L) - f(L)}d$$

A.2 Proof of Proposition 9

Proof. Recall the following IR and IC constraints.

$$f(\theta_1 + m^*)f(\theta_2 + 1 - m^*)w^{SS} - d \geq 0 \quad (\text{IR})$$

$$\operatorname{argmax}_n f(\theta_1 + n)f(\theta_2 + 1 - n)w^{SS} - d \quad (\text{IC})$$

Since the optimal allocation m^* makes the first order differentiation of IC constraint zero, the optimal w^{SS} is simply IR binding solution. Thus,

$$w^{SS} = \frac{d}{f(\theta_1 + m^*)f(\theta_2 + 1 - m^*)} = \frac{d}{(f_\theta^B)^2}$$

□

A.3 Proof of Proposition 11

Proof. Suppose the spillover effect holds. External hiring is preferred to internal promotion if

$E^{IP}[w_S] > E^{EH}[w_S + v]$, i.e.

$$\begin{aligned} E^{IP}[w_S] &> E^{EH}[w_S + v] \\ \eta^S \frac{d}{f(1+L) - f(L)} &> \sum_{\theta} \eta_{\theta}^S \frac{d}{f(1+\theta) - f(\theta)} + E^{EH}[v] \\ \eta_H^S \left(\frac{d}{f(1+L) - f(L)} - \frac{d}{f(1+H) - f(H)} \right) &> \eta^S \frac{c}{\Delta} \\ \frac{\eta_H^S}{\eta^S} (w^{IP} - w_H^{EH}) &> \frac{2c}{\lambda(p-q)} \end{aligned} \quad (3.1)$$

where $\eta_H^S = \eta_{HH}^S + \eta_{HL}^S$.

Now, suppose that the spillback effect holds. Recall that $E^{IP}[w_S] = \frac{\eta^S}{\xi^S} \frac{2c}{p-q}$. Then, external hiring is better than internal promotion if,

$$\begin{aligned} E^{IP}[w_S] &> E^{EH}[w_S + v] \\ \frac{\eta^S}{\xi^S} \frac{2c}{p-q} &> \sum_{\theta} \eta_{\theta}^S \frac{d}{f(1+\theta) - f(\theta)} + E^{EH}[v] \\ w^{IP} - \left(\frac{\eta_{HH}^S + \eta_{HL}^S}{\eta^S} w_H^{EH} + \frac{\eta_{LL}^S}{\eta^S} w_L^{EH} \right) &> \frac{2c}{\lambda(p-q)} \end{aligned}$$

□

A.4 Proof of Proposition 12

Proof. Suppose the spillover effect holds. Then, as (3.1) in the proof of Proposition 11, $w^{IP} - w_H^{EH}$ is independent of p , but

$$\frac{\partial}{\partial p} \left(\frac{c}{p-q} \frac{2}{\lambda} \right) = \frac{c}{p-q} \frac{2}{\lambda} \left(\frac{f(1+H) + f(L)}{\lambda} - \frac{1}{p-q} \right)$$

Since the first term in the parenthesis of the right hand side is less than 1, but the second term is greater than 1, this confirms that the right hand side of (3.1) decreases in p . In the meantime, $\frac{\partial}{\partial p} \left(\frac{\eta_H^S}{\eta^S} \right)$ converges to 1 as p converges to 1. Thus, the inequality (3.1) is relaxed p increases.

Similarly, the convexity of the production technology expands (shrinks) the region where EH is preferred. To make a fair comparison, assume that for both IP and EH,

$$f(1+H) + f(H) + f(L) = C < \infty \quad (\text{C.1})$$

Convexity assumption requires,

$$f(1+H) - f(H) \geq f(H) - f(L) \quad (\text{C.2})$$

The only difference between IP and EH is in the convexity. Since there are infinitely many distributions that satisfy C.1 and C.2, fix $f(H)$ for both EH and IP. Then, the production technology for firm i is more convex than firm j if,

$$f_i(1+H) > f_j(1+H)$$

The above condition automatically implies that $f_i(L) < f_j(L)$. It is straightforward to see that as the production technology becomes more convex, λ increases, thus decreasing the right hand side of (3.1). Also, more convexity increases $\frac{\eta_H^S}{\eta^S}$ and increases $w^{IP} - w_H^{EH}$, hence increasing the

left hand side of (3.1). Therefore, EH becomes dominant as the production technology becomes convex. □

A.5 Proof of Proposition 13

Proof. The pay premium based on the realized payment for each CEO is,

$$v^{EH} = \frac{f(1+\theta) + f(\theta)}{f(1+\theta) - f(\theta)}$$

$$v^{IP} = \frac{d}{c+d} \frac{f(1+\theta) + f(\theta)}{f(1+L) - f(L)}$$

Define ρ as a measure that shows the over-payment illusion as $\frac{v^{EH}}{v^{IP}}$. Then,

$$\rho = \frac{c+d}{d} \frac{f(1+L) - f(L)}{f(1+H) - f(H)} \quad (\text{if } (H, H) \text{ or } (H, L))$$

$$= \frac{c+d}{d} \quad (\text{if } (L, L))$$

Since $f(\cdot)$ is convex, $\frac{f(1+L)-f(L)}{f(1+H)-f(H)} < 1$, thus ρ attains its maximum at (L, L) . To see the effect of the convexity, let $\epsilon > 0$ be a parameter such that

$$\frac{f(1+L) - (f(L) - \epsilon)}{(f(1+H) + \epsilon) - f(H)}$$

Notice that,

$$\frac{\partial}{\partial \epsilon} \left(\frac{f(1+L) - (f(L) - \epsilon)}{(f(1+H) + \epsilon) - f(H)} \right) = \frac{f(1+H) + f(L) - 2f(H)}{((f(1+H) + \epsilon) - f(H))^2} > 0$$

The last inequality is because $f(\cdot)$ is convex. Therefore, as the convexity of $f(\cdot)$ increases, ρ increases. □

A.6 Proof of Proposition 14

Proof. Let $\Delta X = X_{SS} - 2X_S > 0$. Recall the ex ante revenue contingent on $p, f(\cdot)$,

$$E[V] = p^2 \left(f(1+H)f(H)\Delta X + (f(1+H) + f(H))X_S \right) \\ + 2p(1-p) \left(f(1+H)f(L)\Delta X + (f(1+H) + f(L))X_S \right) + (1-p)^2 \left(f(1+L)f(L)\Delta X + (f(1+L) + f(L))X_S \right)$$

To see the effect of convexity on $E[V]$, apply ϵ as before: $f(1+H) + \epsilon, f(L) - \epsilon$. Then,

$$\frac{d}{d\epsilon} E[V] = (2p-1)(f(H)\Delta X + X_S) - 2p(1-p)\Delta X(f(1+H) - f(L) + 2\epsilon)$$

Notice that if $p \leq \frac{1}{2}$, then this is always negative, implying that for small p firm, increase in convexity always leads to firm performance suffering. Now, for $p > \frac{1}{2}$, the increase in convexity leads to firm performance suffering if,

$$\frac{2p-1}{2p(1-p)} \left(f(H) + \frac{X_S}{\Delta X} \right) - f(1+H) + f(L) < 2\epsilon$$

Since $\epsilon \in (0, f(L))$, if $\frac{X_S}{\Delta X}$ is large enough, then the increase in convexity does not lead to negative performance. Now, suppose this is not too large, then once the convexity parameter ϵ reaches the threshold, $\frac{1}{2} \left(\frac{2p-1}{2p(1-p)} \left(f(H) + \frac{X_S}{\Delta X} \right) - f(1+H) + f(L) \right)$, increase in the convexity always causes negative impact on firm performance. Now, the last step is to show whether the principal prefers to rely on EH if the convexity increases. This is straightforward from (3.1) in the proof of Proposition 3 as,

$$(3.1) \quad \frac{p^2((f(1+H) + \epsilon) + f(H)) + 2p(1-p)((f(1+H) + \epsilon) + f(L))}{\eta^S} \left(\frac{d}{f(1+L) - f(L)} - \frac{d}{f(1+H) - f(H)} \right) \\ > \frac{2c}{p-q} \frac{1}{(2-p)((f(1+H) + \epsilon) - f(H)) + (1+p)(f(H) - (f(L) - \epsilon))}$$

Clearly, as ϵ increases, the left hand side increase, but the right hand side decrease, thus, EH dominance region expands. □

References

- Anderson, A. and Smith, L. (2010). Dynamic matching and evolving reputations, *The Review of Economic Studies* **77**(1): 3–29.
- Atiase, R. K. (1985). Predisclosure information, firm capitalization, and security price behavior around earnings announcements, *Journal of Accounting Research* pp. 21–36.
- Autrey, R. L., Dikolli, S. S. and Newman, D. P. (2007). Career concerns and mandated disclosure, *Journal of Accounting and Public Policy* **26**(5): 527–554.
- Autrey, R. L., Dikolli, S. S. and Newman, D. P. (2010). Performance measure aggregation, career incentives, and explicit incentives, *Journal of Management Accounting Research* **22**(1): 115–131.
- Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes, *Science* **293**(5536): 1818–1820.
- Bebchuk, L. and Fried, J. (2009). *Pay without performance: The unfulfilled promise of executive compensation*, Harvard University Press.
- Becker, G. S. (1973). A theory of marriage: Part i, *Journal of Political economy* **81**(4): 813–846.
- Bhagat, S. and Bolton, B. (2008). Corporate governance and firm performance, *Journal of Corporate Finance* **14**(3): 257–273.

- Boot, A. W. (1992). Why hang on to losers? divestitures and takeovers, *The Journal of Finance* **47**(4): 1401–1423.
- Brickley, J. A. (2003). Empirical research on CEO turnover and firm-performance: A discussion, *Journal of Accounting and Economics* **36**(1): 227–233.
- Casamatta, C. and Guembel, A. (2010). Managerial legacies, entrenchment, and strategic inertia, *The Journal of Finance* **65**(6): 2403–2436.
- Chan, W. (1996). External recruitment versus internal promotion, *Journal of Labor Economics* pp. 555–570.
- Chen, K.-P. (2005). External recruitment as an incentive device, *Journal of Labor Economics* **23**(2): 259–277.
- Coughlan, A. T. and Schmidt, R. M. (1985). Executive compensation, management turnover, and firm performance: An empirical investigation, *Journal of Accounting and Economics* **7**(1): 43–66.
- Crémer, J. (1995). Arm’s length relationships, *The Quarterly Journal of Economics* **110**(2): 275–295.
- Custódio, C., Ferreira, M. A. and Matos, P. (2013). Generalists versus specialists: Lifetime work experience and chief executive officer pay, *Journal of Financial Economics* **108**(2): 471–492.
- Dantzig, G. B. (1963). *Linear programming and its extensions*, Princeton University Press, Princeton, NJ.
- Denis, D. J. and Denis, D. K. (1995). Performance changes following top management dismissals, *The Journal of finance* **50**(4): 1029–1057.

- Dewatripont, M., Jewitt, I. and Tirole, J. (2000). Multitask agency problems: Focus and task clustering, *European Economic Review* **44**(4): 869–877.
- Dewatripont, M. and Tirole, J. (1999). Advocates, *Journal of political economy* **107**(1): 1–39.
- Dikolli, S. S., Mayew, W. J. and Nanda, D. (2014). CEO tenure and the performance-turnover relation, *Review of Accounting Studies* **19**(1): 281–327.
- Dye, R. A. (1984). The trouble with tournaments, *Economic Inquiry* **22**(1): 147–149.
- Eisfeldt, A. L. and Kuhnen, C. M. (2013). CEO turnover in a competitive assignment framework, *Journal of Financial Economics* **109**(2): 351–372.
- Equilar (2013). Paying the new boss: Compensation analysis for newly hired CEOs, *Technical report*, Equilar.
- Eriksson, T. (1999). Executive compensation and tournament theory: Empirical tests on danish data, *Journal of labor Economics* **17**(2): 262–280.
- Fairburn, J. A. and Malcomson, J. M. (2001). Performance, promotion, and the peter principle, *The Review of Economic Studies* **68**(1): 45–66.
- Fisman, R. J., Khurana, R., Rhodes-Kropf, M. and Yim, S. (2013). Governance and CEO turnover: Do something or do the right thing?, *Management Science* **60**(2): 319–337.
- Gabaix, X. (2016). Power laws in economics: An introduction, *Journal of Economic Perspectives* **30**(1): 185–206.
- Garrett, D. F. and Pavan, A. (2012). Managerial turnover in a changing world, *Journal of Political Economy* **120**(5): 879–925.

- Gibbons, R. and Murphy, K. (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence, *Journal of Political Economy* **100**(3): 468–505.
- Green, J. R. and Stokey, N. (1983). A comparison of tournaments and contracts, *Journal of Political Economy* (3): 349–364.
- Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration, *The Journal of Political Economy* pp. 691–719.
- Harris, D. and Helfat, C. (1997). Specificity of CEO human capital and compensation, *Strategic Management Journal* **18**(11): 895–920.
- Hart, O. and Moore, J. (1990). Property rights and the nature of the firm, *Journal of political economy* pp. 1119–1158.
- Hermalin, B. E. and Katz, M. L. (1991). Moral hazard and verifiability: The effects of renegotiation in agency, *Econometrica: Journal of the Econometric Society* pp. 1735–1753.
- Hermalin, B. E. and Weisbach, M. S. (1998). Endogenously chosen boards of directors and their monitoring of the CEO, *American Economic Review* pp. 96–118.
- Hirshleifer, D. (1993). Managerial reputation and corporate investment decisions, *Financial Management* pp. 145–160.
- Hirshleifer, D. and Thakor, A. V. (1992). Managerial conservatism, project choice, and debt, *Review of Financial Studies* **5**(3): 437–470.
- Holmstrom, B. (1999). Managerial incentive problems: A dynamic perspective, *The Review of Economic Studies* **66**(1): 169–182.
- Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts,

- asset ownership, and job design, *Journal of Law, Economics, & Organization* pp. 24–52.
- Huson, M. R., Parrino, R. and Starks, L. T. (2001). Internal monitoring mechanisms and CEO turnover: A long-term perspective, *The Journal of Finance* **56**(6): 2265–2297.
- Ijiri, Y. and Simon, H. A. (1977). *Skew distributions and the sizes of business firms*, Vol. 24, North Holland.
- Jensen, M. C. and Murphy, K. J. (1990). Performance pay and top-management incentives, *Journal of Political Economy* pp. 225–264.
- Jenter, D. and Lewellen, K. (2010). Performance-induced CEO turnover, *Work. Pap., Stanford Univ.*
- Kahn, C. and Huberman, G. (1988). Two-sided uncertainty and “up-or-out” contracts, *Journal of Labor Economics* pp. 423–444.
- Kale, J. R., Reis, E. and Venkateswaran, A. (2009). Rank-order tournaments and incentive alignment: The effect on firm performance, *The Journal of Finance* **64**(3): 1479–1512.
- Kaneko, M. and Yamamoto, Y. (1986). The existence and computation of competitive equilibria in markets with an indivisible commodity, *Journal of Economic Theory* **38**(1): 118–136.
- Kaplan, S. N. (2012). Executive compensation and corporate governance in the us: perceptions, facts and challenges, *Technical report*, National Bureau of Economic Research.
- Kaplan, S. N. and Minton, B. A. (2012). How has CEO turnover changed?, *International Review of Finance* **12**(1): 57–87.
- Karaevli, A. (2007). Performance consequences of new CEO ‘outsiderness’: Moderating effects of pre-and post-succession contexts, *Strategic Management Journal* **28**(7): 681–706.

- Larcker, D. and Tayan, B. (2015). *Corporate governance matters: A closer look at organizational choices and their consequences*, Pearson Education.
- Laux, V. (2008). Board independence and CEO turnover, *Journal of Accounting Research* **46**(1): 137–171.
- Laux, V. (2014). Corporate governance, board oversight, and CEO turnover, *Foundations and Trends in Accounting* **8**(1): 1–73.
- Laux, V. (2015). On the interdependencies of supporting, monitoring, and assessing agents, Working paper.
- Lazear, E. P. and Rosen, S. (1981). Rank-order tournaments as optimum labor contracts, *Journal of Political Economy* pp. 841–864.
- Legros, P. and Newman, A. F. (2007). Beauty is a beast, frog is a prince: assortative matching with nontransferabilities, *Econometrica* **75**(4): 1073–1102.
- MacDonald, G. M. (1982). A market equilibrium theory of job assignment and sequential accumulation of information, *The American Economic Review* **72**(5): 1038–1055.
- MacDonald, G. and Marx, L. M. (2001). Adverse specialization, *Journal of Political Economy* **109**(4): 864–899.
- Main, B. G., O’Reilly III, C. A. and Wade, J. (1993). Top executive pay: Tournament or teamwork?, *Journal of Labor Economics* pp. 606–628.
- Malcomson, J. M. (1984). Work incentives, hierarchy, and internal labor markets, *The Journal of Political Economy* pp. 486–507.
- Meyer, M. A. and Vickers, J. (1997). Performance comparisons and dynamic incentives, *Journal*

- of Political Economy* **105**(3): 547–581.
- Murphy, K. J. (1999). Executive compensation, *Handbook of labor economics* **3**: 2485–2563.
- Murphy, K. J. and Zabojnik, J. (2004). CEO pay and appointments: A market-based explanation for recent trends, *American Economic Review* pp. 192–196.
- Murphy, K. J. and Zabojnik, J. (2007). Managerial capital and the market for CEOs, *Working Paper* .
- Murphy, K. J. and Zimmerman, J. L. (1993). Financial performance surrounding CEO turnover, *Journal of Accounting and Economics* **16**(1): 273–315.
- O’Reilly III, C. A., Main, B. G. and Crystal, G. S. (1988). CEO compensation as tournament and social comparison: A tale of two theories, *Administrative Science Quarterly* pp. 257–274.
- Pan, Y. (2015). The determinants and impact of executive-firm matches, *Management Science* .
- Parrino, R. (1997). CEO turnover and outside succession a cross-sectional analysis, *Journal of Financial Economics* **46**(2): 165–197.
- Prendergast, C. (1993). The role of promotion in inducing specific human capital acquisition, *The Quarterly Journal of Economics* pp. 523–534.
- Prendergast, C. and Stole, L. (1996). Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning, *Journal of Political Economy* pp. 1105–1134.
- Puffer, S. M. and Weintrop, J. B. (1991). Corporate performance and CEO turnover: The role of performance expectations, *Administrative Science Quarterly* pp. 1–19.
- Rau, P. R. and Xu, J. (2013). How do ex ante severance pay contracts fit into optimal executive

- incentive schemes?, *Journal of Accounting Research* **51**(3): 631–671.
- Rees, A. (1992). The tournament as a model for executive compensation, *Journal of Post Keynesian Economics* pp. 567–571.
- Roth, A. E. and Sotomayor, M. (1992). Two-sided matching, *Handbook of game theory with economic applications* **1**: 485–541.
- Sattinger, M. (1993). Assignment models of the distribution of earnings, *Journal of Economic Literature* **31**(2): 831–880.
- Shapley, L. S. and Shubik, M. (1971). The assignment game I: The core, *International Journal of Game Theory* **1**(1): 111–130.
- Shen, W. and 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.
- Sliwka, D. (2007). Managerial turnover and strategic change, *Management Science* **53**(11): 1675–1687.
- Stickel, S. E. (1992). Reputation and performance among security analysts, *The Journal of Finance* **47**(5): 1811–1836.
- Taylor, L. A. (2010). Why are CEOs rarely fired? evidence from structural estimation, *The Journal of Finance* **65**(6): 2051–2087.
- Terviö, M. (2008). The difference that CEOs make: An assignment model approach, *The American Economic Review* pp. 642–668.
- Warner, J. B., Watts, R. L. and Wruck, K. H. (1988). Stock prices and top management changes,

Journal of Financial Economics **20**: 461–492.

Weisbach, M. S. (1995). CEO turnover and the firm's investment decisions, *Journal of Financial Economics* **37**(2): 159–188.

Wilson, R. (1975). Informational economies of scale, *The Bell Journal of Economics* pp. 184–195.

Zajac, E. J. (1990). CEO selection, succession, compensation and firm performance: A theoretical integration and empirical analysis, *Strategic Management Journal* **11**(3): 217–230.