Carnegie Mellon Tepper School of Business

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

**DOCTOR OF PHILOSOPHY
INDUSTRIAL ADMINISTRATION
(ACCOUNTING)**

Titled

"THE MECHANISM OF CONTROL IN ORGANIZATIONS: ESSAYS ON IMPERFECT MEASURES OF MANAGERIAL TALENT"

Presented by

Eunhee Kim

Accepted by

Carlos Corona	4/24/2017
Co-Chair: Prof. Carlos Corona	**Date**
Jonathan Glover	5/5/2017
Co-Chair: Prof. Jonathan Glover	**Date**

Approved by The Dean

Robert M. Dammon	5/8/2017
Dean Robert M. Dammon	**Date**

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)

Isa Hafalir

Pierre Jinghong Liang

Acknowledgments

I am deeply indebted to Carlos Corona and Jonathan Glover for their guidance, advice, and encouragement that immensely improve the dissertation and also shape my research agenda. I am grateful to my committee members, Pierre Liang and Isa Hafalir for their insight and helpful suggestions they bring to every discussion. I also want to thank Jack Stecher, Shyam Sunder, Hao Xue, and Gaoqing Zhang for their generous support.

The interdisciplinary research culture both at the Tepper School of Business and the Department of Philosophy at Carnegie Mellon helps me build my research philosophy and the way I think about academic research. I thank my professors for their thought-provoking seminars and the William Larimer Mellon Fellowship for financial support. I also want to thank Lawrence Rapp for his help and effort. Professors in accounting and my colleagues help me improve my presentation. Especially, I want to thank Hyun Hwang, Steve Karolyi, Ryan Kim, Thomas Rutchi, Timothy Shield, and Austin Sudbury for their time and help.

The second and third chapters of this dissertation are based on my job market paper. I have greatly benefited from the discussions with Chandra Kanodia, Frank Gigler, Paul Ma, Francesco Bova, Gordon Richardson, Baohua Xin, Minlei Ye, Paul Newman, Volker Laux, Neil Brisley, Patricia O'Brien, Itay Goldstein, and workshop participants at Carnegie Mellon University, University of Minnesota, University of Toronto, University of Texas at Austin, City University of Hong Kong, Chinese University of Hong Kong, and Waterloo University.

Last, but not least, I am deeply indebted to my parents and my siblings for their support, love, and sacrifice that help me achieve my dream.

To my family, Yungsul, Insun, Unsil, Hyunho, and Hyunmyung

and in memory of

my grandmother, Kisun

Contents

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 1

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.

1.1 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 between firms and managers. Much is known about two-sided matching with fixed attributes: when the fixed attributes on both sides of the market are productive complements, efficient matching involves sorting by attributes. 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, or whenever managers' performance is available, firms update their perceptions of managers' talents. When the attributes 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, sorting by managers' performance 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, repeated sorting generates career concerns, which influence a manager's actions, thereby creating agency problems. The induced agency problems in turn change firms' preferences for managers, thereby affecting the formation of firms and managers. That is, 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 the manager's 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 performance-based sorting (labor market efficiency) may not guarantee efficient matching (economic efficiency).

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. This implies that the distribution of firm size influences the distribution of market wages for managers with respect to their reputations. Then, depending on the distribution of market wages, each manager has distinct opportunity costs of taking the risky project (i.e., induced preference for risk exposure) since a manager's project choice in a current period leads to his reputation update with different market wages in the next period. 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 which does not require information. In this case, a firm needs to offer extra pay to motivate the manager 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. This is because a distribution of firm size forms a distribution of market wages for managers' reputation. Depending on the size distributions of firms, managers' induced preferences for risk exposure are heterogeneous, and can even be 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 strictly prefers the safe project if firm size increases faster as rank increases: 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, thereby 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 fast at the medium rank but slow at the top rank. Consequently, distributions of firm size influence managers' induced preferences for risk exposure, thus changing firms' preferences for managers' reputation and affecting cross-sectional matching patterns. The results imply that the size distributions of firms not only induce agency problems but also change the formation of firms and managers.

Beyond just showing that managers' career concerns can create distortions in matching, the model also highlights the importance of scrutinizing a distribution of attributes in conducting empirical investigations with endogenous matching of firms and managers. In general, the cross-sectional patterns of matching, incentive contracting, internal control, and investment behaviors are affected by the size distributions of firms in an economy. To illustrate, consider a set of the same attributes of firm-manager matches that belong to different economies (with different size distributions of firms). Then, the model predicts that, even if matches of interest share the same attributes of firms and managers, depending on which economy they belong to, their incentive schemes, investment behaviors, or other policy variables differ. In particular, if firm size increases faster as rank increases, then the managerial career concerns are most severe at the medium reputation, thereby leading to higher pay-performance sensitivity or more option-like incentive schemes for medium reputation managers in order to upset their induced preferences for the safe project. If an economy has different size distributions of firms, then the option-like incentive schemes at the medium reputation will not occur for the same reputation of a manager. In addition to cross-sectional patterns in incentive contracts, the model predicts similar patterns of other control devices including corporate investment policies and management turnover depending on a size distribution of firms. These predictions highlight that considering a size distribution of firms is as important as considering individual firm's and manager's attributes, and that the interplay between the market for managers and agency problems may generate optimal solutions, which differ from those in the standard principal-agent framework.

The model I propose can be applied to various economic problems where one side's talent is traded, including matching of auditor and client, analyst and firm, and lender and borrower. By introducing variations into agency frictions and/or a matching problem depending on a particular conflict or friction of interest, the model provides a framework that enables analyzing the interactions between the market forces and agency problems. More specifically, one can analyze the impact of the labor market on matching patterns (e.g., which auditor matches with which client, which analyst follows which firm, or which lender lends to which borrower) and on aggregate performance (e.g., aggregate audit quality, forecast accuracy, or cost of capital). To provide a direct application, in Section 4, I apply the model to the matching of firms and CEOs and offer a new insight into CEO turnover-performance sensitivity. I also discuss, in Section 4, the potential applications of the baseline model in more detail.

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, Tervio 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 Tervio 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 1.2, I describe the basic model setup. After characterizing the economic ingredients, Section 1.3 analyzes a repeated matching problem with career concerns. In Section 1.4, I discuss potential applications and the possibility of other equilibria. Section 1.5 concludes.

1.2 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 economy consists of a continuum of agents (managers) and a continuum of principals (firms). The economy lasts for three 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 $G(S)$. Let $1 - G(S)$ denote the index of firm S with high index denoting a smaller firm (e.g., the smallest firm S_{min}'s index is largest as $1 - G(S_{min}) = 1$). The index can be interpreted as a firm's rank. Hereafter, I call $1 - G(S)$ firm S's rank. Here, the increase in rank is identical with the decrease in index. That is, if a firm's rank increases from the top 20% to the top 10%, then the firm's size increases from $S = G^{-1}(0.8)$ to $S = G^{-1}(0.9)$.[1] 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 principal 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 one 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 earnings level, m, which can be interpreted as a status quo. The risky project will return either a success (high earnings), $h > m$, or a failure (low earnings), $l < m$. Without loss of generality, assume that the investment

[1] $0.2 = 1 - G(S)$ and $0.1 = 1 - G(S)$.

cost is the same for both projects, which is normalized to zero. The probability of success for the project I_r depends on a project 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 $\upsilon, \hat{\upsilon} \in \{r, s\}$ 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. 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., $\hat{\upsilon} = \upsilon$), but a bad type manager receives $\hat{\upsilon} = \upsilon$ with probability of $\beta \in (\frac{1}{2}, 1)$ and $\hat{\upsilon} \neq \upsilon$ 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

Project \ State	$\upsilon = s$	$\upsilon = r$	
	s_3	s_2	s_1
I_s	m	m	m
I_r	l	l	h

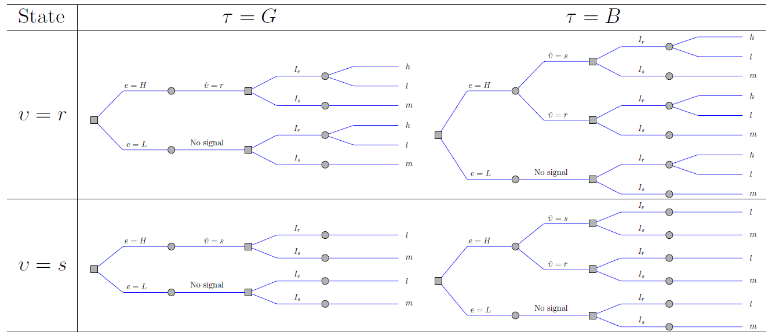

Table 1.1: The Event and Decision Trees of Each Type for Each State

The left table describes the project earnings depending on projects and states. The right table 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 under the good type, and l is more likely under the bad type.

is captured by setting $\alpha = 1/2$.[2] Table 1.1 summarizes the primitives and 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 $\upsilon = r$, then I_r is more profitable, and if $\upsilon = s$, 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, $\gamma_0 \in (0, 1)$, that represents the probability of the agent being a good type.[3] The total measure of managers in the labor market is denoted as $\Gamma \in \mathbb{R}$ such that $\Gamma = 1 + \eta$, where $\eta = Pr(l|\gamma_0)$.[4] After each manager's project outcome is realized, the manager's reputation is

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

[3]To focus on an induced moral hazard problem, I abstract away from an adverse selection problem in which a manager knows his type. Information asymmetry can create signaling behaviors which may differ from the results under the moral hazard in this paper (e.g., Hirshleifer and Thakor (1992), Prendergast and Stole (1996), Sliwka (2007)). The assumption of an initial identical reputation is for the sake of simplicity. The basic idea extends directly to a more general reputation distribution. 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 in those papers.

[4]η can be any positive number, but this is for the sake of simplicity in the benchmark assortative matching pattern in period 2, which results in no matches with poorly performed managers.

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 $t = 1, 2, 3$.

Let $\omega_t(\gamma)$ denote a manager's market value (or outside option) depending on the manager's reputation γ in period t.[5] 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 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. Since all managers have the same reputation in period 1, and the project selection technology, (α, p), that influences a manager's reputation update, is identical, I assume that the period 1 matching is random. After the project outcome is realized at the end of period 1 and 2, every matched manager's reputation is revised, and the demand for rematching of managers and firms arises. From Table 1.1, the probability of each project earnings differs

[5]The 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. It can also be interpreted as a splitting rule of the match surplus between a firm and a manager.

depending on γ. Let $Pr(X|\gamma)$ denote the probability of project earnings 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,

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

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 1.1. Assume that acquiring information about the realized state is valuable enough that principals want to induce high effort from their matched managers and to choose the project according to their signals.

Assumption 1. *For* $\forall\ \gamma \in \{\gamma\}_t$, $t = 1, 2, 3$,

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

Under the first and the second inequalities, a manager's information is valuable enough that

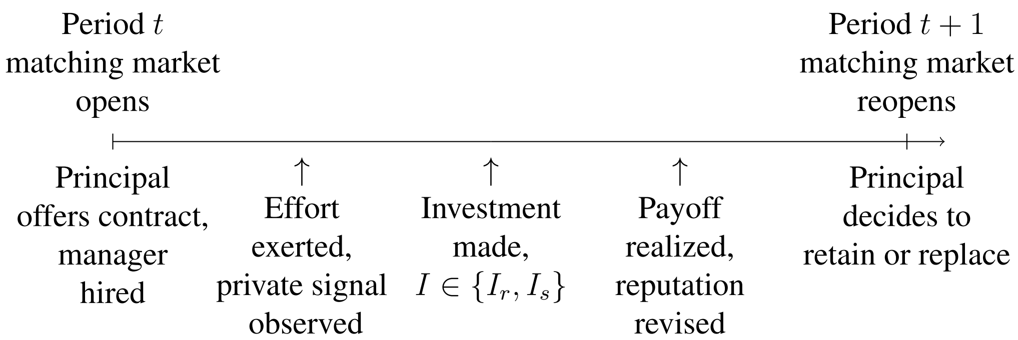

Figure 1.1: **Timeline**

The sequence of events is identical in period $t = 1, 2$. Period 3 has the same sequence of events but there is no rematching at the end.

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 *and a matching function of* $\mu_t : [S_{min}, \infty) \to \{\gamma\}_t$ *such that 1) a contract for manager* γ *induces effort and manager* γ *takes the best offer available or chooses not to participate, 2) each firm chooses its manager optimally, and the market clears.*

For simplicity, I assume that if a matching function in period $t+1$ gives the same reputation as the incumbent manager's updated reputation at the end of period t, 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.[6] I focus on an efficient equilibrium that

[6]Thus, in equilibrium, a firm will just pay the minimum required payment that is determined by the matching

maximizes the aggregate payoffs of the two sides. In the next section, I first solve for the market value for managers, and the corresponding contracts, and then solve for equilibrium matching patterns.

1.3 Analysis

1.3.1 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 confirms the 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. Since there are more managers, the efficiency indicates that 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

market to hire a particular manager.

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,

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

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

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

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 K, P representing participation, NP denoting sitting out the matching market. Observe that a manager's expected payoff in period $t = 1, 2$ includes the expected rematching in the future. 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 a 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

[7] This is because a firm can always hire a manager in the market in each period at the market value. Thus, considering the future matched manager does not affect the current match.

with reputations greater than or equal to γ. Let the set of managers be characterized as N tiers $1 > \gamma_1 > \gamma_2 > \gamma_3 > \cdots > \gamma_N > 0$: 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 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 $S \in \Big[S[i], S[i-1]\Big)$, $i = 2, ..., N$ where $S[i] = G^{-1}(1 - M(\gamma_i))$. Thus, $M(\gamma_i)$ has a direct interpretation of rank with respect to firm size and $S[i]$ represents the top $100 \times M(\gamma_i)\%$ firm size. 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* $1 > \gamma_1 > \gamma_2 > \cdots > \gamma_N > 0$*. Without agency frictions, a reputation based compensation for* γ_i *and* γ_{i+1}*,* $i = 1, 2, \cdots, N-1$*, 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 $F = \frac{1}{2}(1-\beta)\big(h + l(\frac{1}{2} - p)\big),\ \omega_t(\gamma_N) = w_0.$

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. Notice that the extra pay is not only driven by the manager's (incremental) reputation, 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.

[8]An 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.

[9]In 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)).

1.3.2 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 3 contract when there is no reputational incentive left. The required constraints are as follows.

$$-c + \sum_{X \in \{h,l,m\}} Pr(X|\gamma)_\gamma w_3^X \geq \omega_3(\gamma) \tag{IR}$$

$$-c + \sum_{X \in \{h,l,m\}} Pr(X|\gamma) w_3^X \geq max_{x \in [0,1]} \ \ x w_3^m + (1-x)(Pr(h) w_3^h + Pr(l) w_3^l) \tag{IC}$$

where $w_3^X \geq 0$ for $\forall \ X \in \{h,l,m\}$. The (IR) constraint stipulates that the manager with γ reputation must be paid at least his outside option, $\omega_3(\gamma)$, 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. The mixing incentives exist only if the payoff of the safe choice is the same as the risky one. Thus, for simplicity, I consider the safe or risky choice without effort instead of mixing any combination of the two projects (i.e., x is 1 or 0). Considering the (IR) and (IC) constraints with non-negative payments, the principal finds a contract to maximize,

$$\sum_{X \in \{h,l,m\}} Pr(X|\gamma)(S \times X - w_3^X)$$

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) An optimal contract in period 3 is characterized as follows. If $\omega_3(\gamma)$ is small, then the incentive compatibility constraint binds, the individual rationality constraint does not bind and the optimal wage is uniquely characterized by*

$$\frac{w_3^h - w_3^m}{w_3^m - w_3^l} = \frac{Pr(l)}{Pr(h)} = \frac{1 - \alpha p}{\alpha p}$$

If $\omega_3(\gamma)$ is large, then the individual rationality constraint binds and the incentive compatibility constraint does not bind.

The above feature captures the incremental pay for each performance ($w_3^h - w_3^m$ and $w_3^m - w_3^l$), which I call pay performance sensitivity.[10] When there is no career concern, it is clear that the pay performance sensitivity defined above is independent of reputation γ, rather it only depends on the characteristics of projects. The lemma suggests that given the binding (IC) constraint, providing incentives to induce effort for information acquisition and to select the right project is not influenced by a manager's current reputation. The lemma also suggests that without career concerns, the agency cost exists when a manager's outside option is small.

Now, I consider the period 2 contract which will depend on a manager's implicit incentives. Let γ^X denote the updated reputation from project outcome X. Then, the (IR) and (IC) con-

[10]The literature on CEO compensation defines PPS as the change in CEO pay for the change in shareholder wealth (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 $\frac{\lambda\big(f(h)-f(m)\big)}{(1-\lambda)\big(f(m)-f(l)\big)} > 1$ for $\lambda = \frac{1}{2}$.

straints are characterized as follows.

$$-c+\sum_{X\in\{h,m,l\}} Pr(X|\gamma)\big(w_2^X+\omega_3(\gamma^X)\big)\geq \omega_2(\gamma)+\sum_{X\in\{h,m,l\}}\omega_3(\gamma^X) \tag{IR}$$

$$\begin{aligned}&-c+\sum_{X\in\{h,m,l\}} Pr(X|\gamma)\big(w_2^X+\omega_3(\gamma^X)\big)\\ &\quad\geq \Big\{w_2^m+\omega_3(\gamma^m), Pr(h)(w_2^h+\omega_3(\gamma^h))+(1-Pr(h))(w_2^l+\omega_3(\gamma^l))\Big\}\end{aligned} \tag{IC}$$

Lemma 4 finds the characteristic of an optimal contract in period 2.

Lemma 4. *(With Career Concerns) An optimal contract in period 2 is characterized as follows if the (IC) constraint binds.*

$$\frac{w_2^h-w_2^m}{w_2^m-w_2^l}=\frac{\frac{cPr(l)}{Pr(h|\gamma)-Pr(h)(1-Pr(m|\gamma))}-\overbrace{\Big(\omega_3(\gamma^h)-\omega_3(\gamma^m)\Big)}^{\textit{Upside potential}}}{\frac{cPr(h)}{Pr(h|\gamma)-Pr(h)(1-Pr(m|\gamma))}-\underbrace{\Big(\omega_3(\gamma^m)-\omega_3(\gamma^l)\Big)}_{\textit{Downside potential}}}$$

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.

Lemma 5. *Managerial career concerns can mitigate or exacerbate their incentives. Formally,* *career concerns mitigate the moral hazard problem of manager γ if*

$$2Pr(h|\gamma)\Big(\omega_3(\gamma^h) - \omega_3(\gamma^l)\Big) > \omega_3(\gamma^m) - \omega_3(\gamma^l) > 2(Pr(h) - Pr(h|\gamma))\Big(\omega_3(\gamma^h) - \omega_3(\gamma^l)\Big)$$

On the other hand, career concerns exacerbate the moral hazard if one or both of the inequalities are reversed.

The first inequality mitigates the manager γ's incentive of choosing I_s without effort ($x = 1$), and the second inequality mitigates the manager's incentive of choosing I_r without effort ($x = 0$). Given the downside potential ($\omega_3(\gamma^m) - \omega_3(\gamma^l)$), if the probabilistic upside potential ($\omega_3(\gamma^h) - \omega_3(\gamma^l)$) is not large enough, then the first inequality is likely to be reversed. Meanwhile, compared to the downside potential, if the upside is big enough, then the second inequality is likely to be reversed. Whether a manager's career concerns mitigate or exacerbate the incentive problems depend on market values for managers' reputations, which will be discussed in the next section.

1.3.3 Repeated Matching Equilibrium

In this section, I find an optimal one-to-one matching between firms and managers.[11] The goal here is to analyze the impact of managers' career concerns on matching patterns. Thus, the main analysis is a rematching equilibrium in period 2 since every manager is identical in period 1 and no manager has career concerns in period 3. I first describe period 1 matching and reputation updating by Bayes' rule. Then I provide the main analysis of period 2 and 3 in tandem.

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

Initial Match in Period 1:

Since the primitive parameters, α, p, are independent and identical across firms, reputation change upon project outcomes is independent of firm size. Recall that 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 is matched in equilibrium, and the measure $\Gamma - 1$ of managers remain unmatched in period 1. Thus, the market wage $\omega_1(\gamma_0)$ in an initial match is determined by

$$E[w^X|\omega_1(\gamma_0)] - c + \Pi_1^P(\gamma_0) = w_0 + \Pi_1^{NP}(\gamma_0)$$

With $\omega_1(\gamma_0)$, the compensation contract in period 1 is determined as in Section 1.3.2.[12]

Reputation Update Depending on Project Earnings

Due to an optimal contract found in Section 1.3.2, every matched manager in equilibrium exerts effort to acquire a signal before they select a project. Given the assumption that $\alpha = 1/2$ and the realized project earnings in period 1, the principal and the market will update the manager's

[12]The incentive contract in period 1 does not have to be the same as long as it is incentive compatible, individually rational, and the expected payoff from the contract is the same across managers.

reputation as follows.[13]

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

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

$$= \frac{1}{1 + \frac{1-\gamma}{\gamma}\frac{1-p\beta}{1-p}}$$

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

Performance h always helps manager γ improve his reputation since $\beta < 1$. Since $\alpha = 1/2$, performance m maintains a manager's reputation. How l changes the reputation depends on the parameter values. The natural tendency is that l tarnishes a manager's 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 ms. This is summarized in the following assumption.

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

Since every manager is identical in period 1, and in equilibrium, each matched manager exerts effort to choose a project according to their signal, the distribution of reputation at the beginning of period 2 and its support $\{\gamma\}_{t=2}$, is deterministic: $\gamma^h > \gamma^m = \gamma_0 > \gamma^l$. However, at the beginning of period 3, depending on the matching outcome in period 2, $\{\gamma\}_{t=3}$ differs. For instance, if all γ^h and γ^m, γ_0 managers are matched and exert effort to select a project according to their signals, then there will be five tiers with 9 histories: $\gamma^{hh} > \gamma^{hm} = \gamma^{mh} = \gamma^h > \gamma^{mm} =$

[13]It 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.

$\gamma^m > \gamma^{hl} > \gamma^{ml} = \gamma^l$. 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, γ_0), then there will be six tiers with 11 histories: $\gamma^{hh} > \gamma^{hm} = \gamma^{mh} > \gamma^{mm} = \gamma^m = \gamma_0 > \gamma^{hl} = \gamma^{lh} > \gamma^{ml} = \gamma^l > \gamma^{ll}$. Depending on the distribution of managers' reputations at the beginning of period 3, the market value for managers differ, which will also affect managers' career concerns in period 2. As a benchmark matching pattern, I first find a rematching equilibrium in period 3 (i.e., when there is no career concern).

Period 3 matching: Matching Patterns Without Career Concerns

To analyze the interaction between matching efficiency and managerial career concerns, I first discuss period 3 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 every firm hires a manager in equilibrium and all the bargaining power is given to firms, the (PC-firm) does not bind even for firm S_{min}. Moreover, the (IR) constraint is not critical in determining matching because as long as the expected payoff is at least greater than or equal to his market value (i.e., a manager's outside option), managers are indifferent from switching the current match to the other if the expected payoffs of offered contracts are the same. Lastly, a manager's payoff is only determined by his wage at the end of period 3 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.

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.

Since an equilibrium contract induces the signal acquisition effort, the (IC) constraint-whether binding or not-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.

Recall that, without career concerns, agency costs exist when a manager's outside option is small (Lemma 3). In this case, a firm's preference for a manager is straightforward: a high reputation manager, albeit expensive to match, is superior in both match efficiency and agency cost. Thus, an efficient rematching equilibrium, as in the standard two-sided matching with productive complementarity, maximizes the total match surplus $\int Y_3(S, \mu(S))dG(S)$ subject to the (SC) constraints that determine $\omega_3(\gamma)$ with w_0 for the lowest matched managers.[14]

$$Y_3(S, \gamma) - E[w_3^X|\omega_3(\gamma)] \geq Y_3(S, \gamma^{'}) - E[w_3^X|\omega_3(\gamma^{'})] \;\; \forall S, \gamma, \gamma^{'} \in \{\gamma\}_{t=3} \qquad \text{(SC(S,}\gamma\text{))}$$

where $w_3^X, X \in \{h, m, l\}$ 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. The managers' market value in period 3 for each reputation tier is characterized by Lemma 2, which is summarized in Table 1.2. In an efficient equilibrium, the period 3 rematching pattern would look as if it takes as follows. All manager γ^{hh} are assigned to the largest firms $S \in \Big[S[hh], \infty\Big)$, where $S[hh]$ denotes the top $100 \times M(\gamma^{hh})\%$ firm

[14]In period 3, firm S's and manager γ's payoffs are $Y_3(S, \gamma) - E[w_3^X|\omega_3(\gamma)]$ and $E[w_3^X|\omega_3(\gamma)] - c$ respectively. Thus, an efficient matching maximizes the total surplus, $Y_3(S, \gamma) - c$, and c does not change the solution.

Rank-order: $\gamma^{hh} > \gamma^{hm} = \gamma^{mh} = \gamma^{h} > \gamma^{mm} = \gamma^{m} > \gamma^{hl} > \gamma^{ml} = \gamma^{l}$			
Reputation	Market value, $\omega_3(\gamma)$	Reputation	Market value, $\omega_3(\gamma)$
γ^{hh}	$(\gamma^{hh} - \gamma^{hm}) \cdot F \cdot S[hh] + \omega_3(\gamma^{hm})$	γ^{hl}	$(\gamma^{hl} - \gamma^{m}) \cdot F \cdot S[hl] + \omega_3(\gamma^{ml})$
γ^{hm}	$(\gamma^{hm} - \gamma^{mm}) \cdot F \cdot S[hm] + \omega_3(\gamma^{mm})$	γ^{ml}	w_0
γ^{mm}	$(\gamma^{mm} - \gamma^{hl}) \cdot F \cdot S[mm] + \omega_3(\gamma^{hl})$	γ^{l}	w_0

Table 1.2: Reputation based Market Value of managers

This table summarizes competitively determined market values for managers based on their reputation where $F = \frac{1}{2}(1-\beta)\big(h + l(\frac{1}{2} - p)\big)$ in case that $\forall\ \gamma^h, \gamma^m, \gamma_0$ are matched in period 2.

size. Starting from the right below rank of $M(\gamma^{hh})$, the next largest firms $S \in \Big[S[hm], S[hh]\Big)$ are matched with the second top tier manager γ^{hm} until the measure of $M(\gamma^{hm}) - M(\gamma^{hh})$ firms are matched. The same process continues until every firm gets matched. Lemma 6 summarizes this discussion.

Lemma 6. *The rematching equilibrium in period 3 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.*

Period 2 matching: Match Efficiency and Career Concerns

As in period 3, the two constraints are not important in determining matching in period 2: the (PC-firm) is not critical given that the market clears; neither is the (IR) as it simply guarantees the expected wage to be greater than or equal to a manager's outside option. The (SC) is still the key in finding threshold firms. Contrary to period 3, however, the (PC-manager) can be as important as the (IC) due to period 3 matching: by sitting out, it allows a manager to maintain his reputation. Notice that a manager can also maintain his reputation by participating but taking I_s without effort. The option value of sitting-out also implies that the (IC)$-$particularly, a manager's preference for I_s- may also bind. To see this, recall Lemma 5: if either of two inequalities is reversed, then it exacerbates the moral hazard problem. In particular, relative to the probabilistic upside $(2Pr(h|\gamma)(\omega_3(\gamma^h) - \omega_3(\gamma^l)))$, if the downside potential $(\omega_3(\gamma^m) - \omega_3(\gamma^l))$ is big enough, then a manager is likely to prefer I_s. Meanwhile, if the downside potential is negligible relative to the upside, then a manager is likely to prefer I_r.

First, the lowest reputation manager γ^l has the smallest downside potential across managers, thus the career concerns make him want to take I_r without effort. However, manager γ^l already has the smallest outside option, implying that the (IC) constraint for them anyway binds even without career concerns. Next, consider manager γ^m or γ^h, who can potentially have high outside options. Since they are relatively at the upper rank in the market, these managers face non-negligible downside potential. If the downside is big enough relative to the probabilistic upside potential, then they will likely to prefer I_s to maintain their current reputations. This is when there is positive option value of sitting-out.

To see how career concerns affect the matching (threshold firms and the market values), consider manager γ^i who prefers I_s without effort. This implies that extra rents are necessary for manager γ^i to provide incentive to exert effort for I_r. However, the extra payoff must satisfy the (SC) constraints, i.e., it is incentive compatible for firms to pay extra to their equilibrium matched managers rather than switching their matching partners. Given that the market value for manager γ^i is determined by the indifference condition of the threshold firm, the extra payoff for manager γ^i may no longer be incentive compatible for the top $100 \times M(\gamma^i)\%$ firm. Instead, it shall be shifted toward a larger firm in order 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 either, i.e., those unmatched γ^i, if any, consider participation and sitting out to be equal options. This stability requirement makes the (PC-manager) constraint for γ^i bind, which will not happen in period 3.

To summarize the key mechanism, manager γ^i's career concerns (induced preference for I_s due to period 3 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 $\omega_2^{SC}(\gamma^i, S[i]), \omega_2^{PC}(\gamma^i)$ denote the period 2 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 $\omega_2^*(\gamma^i)$ denote an equilibrium market value. Then, in period 2, for all $\gamma^i \neq \gamma^l$, $\omega_2^*(\gamma^i) = \omega_2^{SC}(\gamma^i, S[i]) \geq \omega_2^{PC}(\gamma^i)$ and the equality holds if $S[i] > G^{-1}(1 - M(\gamma^i))$.*

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 a 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.

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. Without agency frictions, 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 losses of any firms are strictly greater than agency costs, then an equilibrium matching pattern shall be positive assortative: in period 2, all γ^h managers are assigned to $\forall S \in \Big[S[h], \infty\Big)$ where $S[h] = G^{-1}(1 - M(\gamma^h))$, and all managers γ^m, γ_0 to the rest of the firms, $\forall S \in \Big[S_{min}, S[h]\Big)$. The following proposition demonstrates this.

Proposition 2. *(Assortative Matching) The matching pattern is positive assotative if match efficiency losses are greater than agency costs for all firms: 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., in the large economy, there is no matched manager whose reputation is less than any of the unmatched manager.*

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 become larger. One 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 3. 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 wages, 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 from matching 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 $U(m) = \omega_3(\gamma^{mh}) - \omega_3(\gamma^{mm}), D(m) = \omega_3(\gamma^{mm}) - \omega_3(\gamma^{ml})$ denote an upside and a downside potential, and $S[m]$ denote an equilibrium threshold firm for γ^m. In period 2, manager γ^m's career concerns lead to $S[m] > S_{min} = G^{-1}(1-M(\gamma^m))$ if $Pr(h|\gamma^m) \times U(m) + \omega_2^{SC}(\gamma^m, S_{min}) < Pr(l|\gamma^m) \times D(m)$.*

Note that $U(m) = S[mh] \times (\gamma^{mh} - \gamma^{mm})F, D(m) = \Big(S[mm] \times (\gamma^{mm} - \gamma^{hl}) + S[hl] \times (\gamma^{hl} - \gamma^{ml})\Big)F$. 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 $U(h) = S[hh] \times (\gamma^{hh} - \gamma^{hm})F, D(h) = \Big(S[hm](\gamma^{hm} - \gamma^{mm}) + S[mm](\gamma^{mm} - \gamma^{hl})\Big)F$.

Proposition 3. *(Failure of Assortativity at the Middle) There exists a stable rematching equilibrium in period 2 where some γ^m get unmatched if $Pr(h|\gamma^h) \times U(h) + \omega_2(\gamma^h) \geq Pr(l|\gamma^h) \times D(h)$ and Lemma 7 holds: every manager γ^h is matched with the large firms $S \in \Big[S[h], \infty\Big)$, some manager γ^m is matched with $S \in \Big[S[m], S[h]\Big)$ where $S[m] > S_{min}$, all remaining firms $S \in \Big[S_{min}, S[m]\Big)$ are matched with γ^l. The unmatched managers are indifferent from participation and sitting out.*

Implications for distributions of firm size The characteristic of the sufficient condition is that 1) the size of a threshold firm (i.e., $S[hh]$, the top $100 \times M(\gamma^{hh})\%$ firm) that determines the market value for γ^{hh} is sufficiently large, 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, then the manager γ^h would be less willing to take the I_r. Moreover, a large enough firm $S[hm]$ does not generate the above result because manager γ^m may have incentive to take the I_r. The following proposition characterizes a sufficient condition with respect to firm size that forms induced risk-aversion at the middle (γ^m) while induced risk-seeking at the top (γ^h).

Proposition 4. *Let $G^{-1}(\cdot)$ denote an inverse function of the firm size distribution. Then, manager γ^m faces induced risk-aversion, but manager γ^h does not if*

$$\frac{G^{-1}(1-M(\gamma^{mh}))}{G^{-1}(1-M(\gamma^{hl}))} < K < \frac{G^{-1}(1-M(\gamma^{hh}))}{G^{-1}(1-M(\gamma^{hm}))}$$

where $K = \left((\frac{\alpha}{Pr(h|\gamma)} - 1) \frac{Pr(h)Pr(l|B)}{Pr(l)Pr(h|B)} \right)$.

Since $G^{-1}(1-M(\gamma^{ij}))$ is a period 3 threshold firm at the rank of $M(\gamma^{ij})$ that determines the period 3 market value for manager γ^{ij}, $\frac{G^{-1}(1-M(\gamma^{ij_1}))}{G^{-1}(1-M(\gamma^{ij_2}))}$ measures the slope between the rank of $M(\gamma^{ij_1})$ and $M(\gamma^{ij_2})$. This suggests that the rate of change of threshold firm size that increases as rank increases supports the above equilibrium.

Proposition 5. *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 rank increases.*

Some γ^m replaces γ^h: Distortion at the top The same logic applies to 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, 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 $U(h) = \omega_3(\gamma^{hh}) - \omega_3(\gamma^{hm}), D(h) = \omega_3(\gamma^{hm}) - \omega_3(\gamma^{hl})$ denote an upside and a downside potential respectively for manager γ^h. In period 2, manager γ^h's career concerns lead to $S[h] > G^{-1}(1 - M(\gamma^h))$ if $Pr(h|\gamma^h) \times U(h) + \omega_2(\gamma^h) < Pr(l|\gamma^h)D(h)$.*

The following proposition characterizes an equilibrium that exhibits a hole at the reputation γ^h.

Proposition 6. *(Failure of Assortativity at the Top) There exists a stable rematching equilibrium in period 2 where some γ^h get unmatched if $Pr(h|\gamma^m) \times U(m) + \omega_2^{SC}(\gamma^m, S_{min}) \geq Pr(l|\gamma^m) \times D(m)$ 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. Between participation and sitting out, the unmatched γ^h is indifferent.*

The characteristic of the sufficient condition is that 1) the threshold firm ($S[hh]$) that determines the market value for γ^{hh} 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 rate of change of threshold firm size that increases sharply at the middle supports the above equilibrium.

Lemma 9. *The hole at the top equilibrium is supported by distribution of firm size that indicates that firm size increases fast at the middle but slow at the top. More formally, manager γ^h faces induced risk-aversion, but manager γ^m does not if*

$$\frac{G^{-1}(1-M(\gamma^{mh}))}{G^{-1}(1-M(\gamma^{hl}))} > K > \frac{G^{-1}(1-M(\gamma^{hh}))}{G^{-1}(1-M(\gamma^{hm}))}$$

where $K = \left((\frac{\alpha}{Pr(h|\gamma)} - 1) \frac{Pr(h)Pr(l|B)}{Pr(l)Pr(h|B)} \right)$.

To summarize, the intuition for these findings 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 complementarity 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 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.

1.4 Applications

The baseline model that I proposed in this paper can be applied to various economic problems with two-sided matching including auditor and client, analyst and firm, lender and borrower, 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 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. 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 of firms). 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.

By a similar logic, the baseline model can also be applied to analysts and their coverage firms. The productive complemenarity between an analyst's talent 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 records that form their reputation, and better reputation analysts perform better(Stickel (1992)). Given this one-side concerning about their reputation and productive complementarity, one can analyze the formation of analysts and coverage firms, the interplay between the labor market for analysts and agency conflicts within a match between analyst and coverage firms. Based on such a formation, one can also investigate how aggregate forecast accuracy or the quality of analyst recommendations changes. Other potential applications include matching of lenders and borrowers, and a board of directors and firms.

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 a direct application of the baseline model and its trade-off in this paper. 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.

1.4.1 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)).[15] 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)).[16] However, this empirical regularity can be explained by

[15]The 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).

[16]Both 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

the distortions of matching patterns in the baseline model.

To see this, consider the baseline setup. In the model after period 1 ends, we have managers of $\gamma^h, \gamma^m, \gamma^l$. The period 2 rematching game in the baseline setup is interpreted as a firm' 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).

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 are replaced and remain unmatched in the market will be identified as involuntary turnover. This feature of identification is also consistent with how CEO succession events are classified in the literature. In the event of CEO departure, a commonly used algorithm classifies only those departing CEOs who do not have a next job as involuntary CEO turnover events (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 involuntary turnover, one needs to expand the career horizon of departing CEOs or to collect more information regarding the turnover events.[17]

Even with this caution, the findings in the baseline setup provide implications for CEO

indices, do not change turnover-performance sensitivity.

[17] Kaplan 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.

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 always leads to turnover, which generates strong impact of performance in predicting turnover. Since this strong association is not empirically observed, to test this proposition, 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 relative to other economies.

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^ls. Moreover, the association between performance m and turnover increases by the measure of manager γ^ms sitting out the matching market (recall, in Proposition 2, the relation between performance 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 evidence. Interestingly, the sufficient conditions that derive the matching pattern in Proposition 3 indicate a characteristic where firm size increases faster as the rank increases, which is satisfied by a power law with a shape parameter greater than 1.[18] 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

[18]According to Axtell (2001), the empirically documented shape parameter for the size distribution of firms in the U.S. is 1.059.

law distribution. However, at least, the model confirms that, once the empirically well supported distribution is considered, it will generate an empirically well documented turnover performance relation.

Lastly, in Proposition 6, 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 6. However, what the baseline model predicts is that if the economy exhibits that firm size increases at the middle ranks than the top ranks, then the cross-sectional turnover-performance sensitivity will be close to U-shaped, and there will be more frequent turnover at the top performance levels than the medium.

Overall, depending on the distributions of firm size, the model predicts that the relation between CEO turnover and performance differs, either strong, weak, or U-shaped. However, across all these properties, 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 by industries or by 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.

1.4.2 Discussion

This section discusses some interpretations and implications 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 a 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 the 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.[19]

[19] Another interpretation with respect to the main result is that it is at least assortative among those matched managers and firms. However, the key point here is that a firm may not be willing to match with an available better reputation manager because of the manager's career concerns, and prefers to match with a strictly worse reputation

Another notable implication for the main result is the possibility of many different types of equilibria depending on a size distribution of firms, including an equilibrium where all managers share the same induced risk preferences, or where the hole in the matching appears at the top, middle, and the bottom. My goal in this paper is to offer some sufficient conditions that derive non-assortative matching patterns to prove that performance-based sorting may not guarantee an efficient matching outcome. To this end, I focus only on the induced risk aversion for a middle reputation manager or a high reputation manager. While there can be other types of equilibria, the analysis in this paper can be similarly applied to such different types of equilibria.

The purpose of introducing an agency friction into the labor market for managers is to explore how the formation of firms and managers is influenced by the friction 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 managerial accounting and control. Which contract is offered and how internal control mechanisms are designed would look like an outcome of an individual firm's and manager's problem. However, the formation of firms and managers is closely tied to the labor market for managers, thus the endogeniety 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.

manager instead.

1.5 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 the 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 2

Career Concerns and Project Selection: The Role of Reputation Insurance

ABSTRACT

Motivated by 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 CEO's 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.

2.1 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 governance practices can serve as insurance for a manager to resolve potential agency frictions arising from a manager's career concerns, but each governance mechanism has different implications for matching patterns. 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 insurance for reputation. The aspect of insurance reduces potential frictions due to career concerns, thus arising as part of optimal contracting devices for individual firms. However, adapting such a mechanism as insurance for individual firms collectively creates externality in the labor market for CEOs, thus affecting firm-CEO matches.

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 reduces the reputation damage 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. Although the board's monitoring weakens performance-based sorting, it creates positive externality in the market as it increases both match and incentive efficiency. 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 CEO's 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). Contrary to the board's monitoring, individual firms' disclosure policy for the sake of insurance collectively creates negative externality in the labor market sorting. In principle, non-disclosure at a CEO departure event requires one to speculate root causes of turnover. This speculation, by its nature, is subject to classification error, which not only interrupts the matching in the following period but also 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. The expected future severance as insurance in turn incentivizes the CEO to take the potentially profitable risky project. Thus, the direct insurance payment through severance agreement mitigates managerial career concerns. Similar to the performance non-disclosure, the ex ante severance agreement can weaken the turnover-performance sensitivity due to performance tolerance. This is because 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. However, contrary to the non-disclosure, the negative externality arising from performance tolerance is borne by the incumbent firm: the incumbent firm would have been better off by rematching with a better CEO if the firm did not offer an ex ante severance agreement.

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 2.2 briefly provides the baseline setup. Then, Section 2.3 develops and discusses the board's monitoring as insurance. Section 2.4 considers the disclosure policy as insurance. Section 2.5 introduces a severance agreement as insurance. Section 2.7 concludes. Since CEO turnover has been extensively studied in the literature, I will discuss more about the existing related works in Appendix.

2.2 Baseline Setup

Consider a firm (principal) and a CEO (agent). The game lasts for three 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, he sits out the market, 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 $\upsilon, \hat{\upsilon} \in \{r, s\}$ 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., $\hat{\upsilon} = \upsilon$), but a bad type manager receives $\hat{\upsilon} = \upsilon$ with probability of $\beta \in (0, 1)$ and $\hat{\upsilon} \neq \upsilon$ 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,

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

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 $t = 1, 2, 3$. Then, the firm, taking the market value of each manager as given, chooses the optimal compensation 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. Since a governance choice is costly, the principal considers incurring such a cost only if its matched CEO is career concerned which potentially leads to the principal's inefficient replacement decision. In period 1 when every CEO is identical, assume that no principal wants to incur the cost. Thus, the following models of governance start from period 2 when some CEOs are

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

career concerned.

2.3 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

[2] A 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.

$E[\omega_3(\gamma)|e = L, I_r, b] < E[\omega_3(\gamma)|e = L, I_s, b]$ 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 10. *Suppose that $b > \frac{1}{2}$. Let $\gamma^l_{b+}, \gamma^l_{b-}$ 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.*

$$\begin{aligned} E[\gamma^l|b] = \Big(b\big(Pr(l, \tau = G) - Pr(l, \tau = B)\big) + Pr(l, \tau = B)\Big)\gamma^l_{b+} \\ + \Big(-b\big(Pr(l, \tau = G) - Pr(l, \tau = B)\big) + Pr(l, \tau = G)\Big)\gamma^l_{b-} \end{aligned}$$

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

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

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

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

2 is,

$$Y_2(S,\gamma) - E[w_2^X|\omega_2(\gamma)] - C(b) \geq max_{\gamma'}\left\{Y_2(S,\gamma') - E[w_2^X|\omega_2(\gamma')]\right\} \tag{2.1}$$

The left-hand side of (2.1) 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 11 shows the existence of b^* and its characteristics.

Lemma 11. *(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 7 finds when the market prefers monitor to firm performance. In this case, non-assortativity in matching pattern arises ex post: a firm that is supposed to match with a better CEO in the market find it profitable to retain its poor performing CEO when the board receives a favorable signal.

Proposition 7. *(Monitor as Substitute and Non-assortativity) 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^l_{b+} if an alternative CEO is γ.*

Proposition 7 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]

2.4 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

[3]Since $E[\omega_3(\gamma)|e = L, I_r, b] < E[\omega_3(\gamma)|e = L, I_s, b]$ 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.

not explicitly disclosed.[4] To see how performance disclosure plays as reputation insurance for 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 insurance effect, 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 12 finds a condition in which performance non-disclosure can play as insurance.

Lemma 12. *Performance non-disclosure can be used as insurance in period 2 only if firms are sufficiently large,* $S \in \Big[S[hl], S[h]\Big)$.

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 13 finds the expected mismatch costs in period 3 due to performance non-disclosure.

[4] See, 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 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).

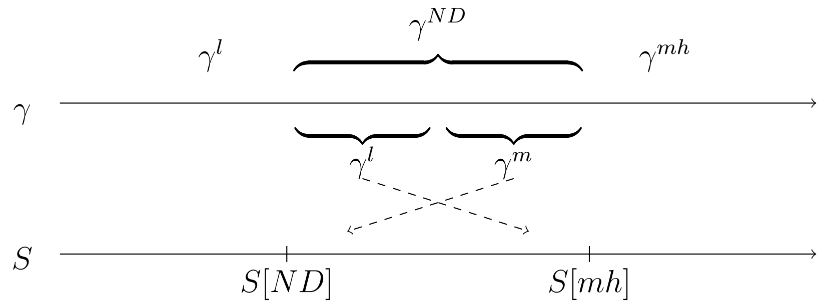

Figure 2.1: **Potential mis-assignment of CEOs and Firms due to non-disclosure in period 2.**

Lemma 13. *By relying on performance non-disclosure, the expected rematch distortion in period 3 is*

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

where $H_1 = \Big[S[l], S[hl]\Big), H_2 = \Big[S[hl], S[m]\Big)$. *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 8 highlights this result.

Proposition 8. *(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,* $S \in \Big[S[ND], S[m]\Big)$.

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 actual performance will not be revealed.

2.5 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

[5] In 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 14 summarizes the result for this investigation.

Lemma 14. *(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 9. *(Performance Tolerance and Endogenous Ex Post Rematching Frictions) 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,* $\forall\ S \in \Big[S[m], S[m] + f(EAS, \alpha_3)\Big)$ *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 9 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.

2.6 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 independently 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 creates 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, 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 3

Generalists versus Specialsts: 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.

3.1 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 Zabojnik (2004); Murphy and Zabojnik (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 Zabojnik (2004); Murphy and Zabojnik (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 organization 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 subordinates 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] Therefore, if a firm-specific skill is necessary, an organization needs to find ways to incentivize its subordinates.

[1]This argument implicitly considers firm-specific skills as assets. Hence, the control right over the firm-specific skill does not change the outside option. Another implicit assumption in this argument is that a subordinate 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.

This paper argues that, although internal promotion and external hiring create subordinates' incentive to develop firm-specific skills, the distinct mechanism in each hiring decision faces different trade-offs. To illustrate this idea, I build an optimal contracting model based on a sequential team production setting with two hiring choices. More specifically, a firm-specific skill developed by subordinates first affects production technology of each task permanently. The skill itself does not create any value yet, so a superior will finalize the productivity by allocating resources into tasks. In case that a subordinate has acquired a firm-specific skill and becomes a superior, the superior can save the cost of managing his own skill (task) due to his expertise (Demski (1998)). While internal promotion provides incentives better for subordinates, the expertise of an internal CEO (specialist) should be viewed as a two-edged sword: the expertise allows the superior to manage his own skill effectively, however the management efficiency may create bias.

By considering the trade-off between management allocational efficiency and incentive efficiency, this paper compares the performance of two distinctive mechanisms contingent on the benefit of expertise from firm-specific skills on management. In general, internal promotion succeeds in providing firm-specific skill development incentives and adequate management allocation incentives by aggregating two incentive problems using promotion bonus. 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 organizational slack, which is defined as extra resource allocation that benefits the promoted superior, is necessary to create bonus since winning the tournament itself may not be enough if the reward is not sufficient. With the expertise benefit and the cost of required organizational slack, the internal promotion that aggregates the two incentive problems is an efficient mechanism if such an expertise benefit creates enough slack for the superior. Since enough organizational slack incentivizes the subordinates effectively, to the extent that the slack is big enough relative to the superior's biased allocation to create slack, the incentive aggregation remain as optimal for the multiagent incomplete contracting problem.

However, external hiring becomes optimal if such organizational slack incurs too much distortions in management. If the slack is not enough, a firm needs to induce more bias to create slack, thereby distorting the specialist superior's management allocation. In this case, it is optimal for the principal to use external hiring in which the externally hired agent (generalist) 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 organizational slack required.

This dominance between internal promotion and external hiring is characterized by the efficiency of firm-specific skills on management. If the impact of the expertise on managing the superior's own skill is high, then a small bias toward his own skill 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. In an area where an optimal mechanism is external hiring, if a firm relied on internal promotion, then a firm's payoff would have been much lower than the payoff under the external hiring due to the biased allocation for the sake of organizational 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 return 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 3.2 describes the model and develops a numerical example. Section 3.3 analyzes the model and finds an optimal form of CEO appointments. Section 3.4 provides empirical implications. Section 3.5 relates this paper to the existing literature. Conclusion is presented in Section 3.6. Related literature and all proofs are in the appendix.

3.2 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, 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

[2]The 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.

[3]In 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.

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.

3.2.1 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 problem[4], contributed to the unverifiability of actions and skills.

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 $p > q$,

[4]The 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.

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 enforeceability problems. 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.

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 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

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

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, $C(m_1, m_2)$, convex increasing in both m_1 and m_2. In case that the subordinate with high skill is promoted, then he can save the cost of managing his own skill. Without loss of generality, call the internal CEO's task t_1. Then, his cost of management is $C(m_1 - d\theta_1, m_2)$, where $d \in (0, 1)$ measures the expertise benefit from his own skill. i.e., a specialist CEO (internal candidate) has comparative advantage in managing his own expertise relative to a generalist CEO (external hire). Assume that d is such that $m_1 - d\theta_1 \geq 0$ and, 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 to be additive. That is,

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

Assume that $f(\theta + m)$ is differentiable and concave-increasing in $(\theta + m)$.

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

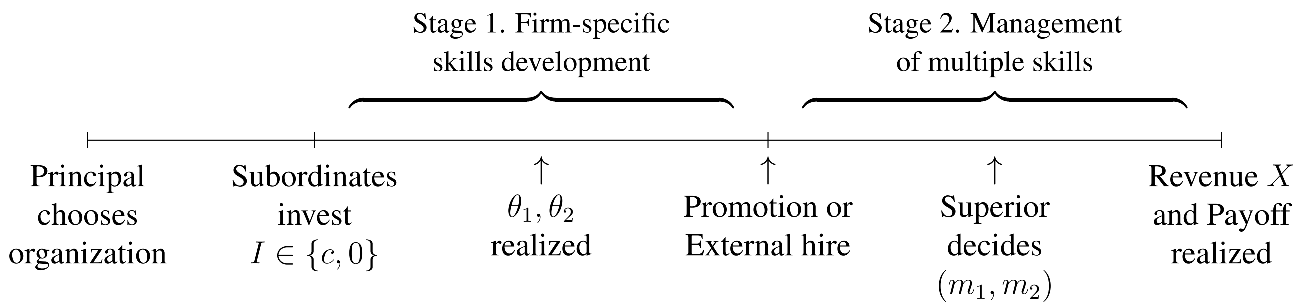

Figure 3.1: **Timeline.**

the agents' investment choice and management effort allocation are not verifiable, the revenue $X_{(t_1,t_2)}$, 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. $X_{(S,S)} = X_{SS} > 2X_S$.[6] To summarize, the principal designs a contract including a 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 3.1 depicts the sequence of events.

3.2.2 Internal Promotion versus External Hiring

This section provides the key mechanism of each hiring structure to understand different incentive effects.

Internal Promotion: If the principal decides to use internal promotion, she designates one of subordinates as a superior depending on the rank order of θs: 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

[6] If $X_{(S,S)} \leq 2X_S$, it is better to split the two tasks as opposed to putting them together as one firm. This assumption can also be interpreted as that there is synergy between the two tasks.

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]

To create incentives for subordinates, they expect to receive a promotion bonus, $\overline{U} > 0$ if promoted, otherwise, they will get v.[8] The definition of the promotion bonus will be discussed in the next subsection.[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. The newly hired agent determines the management allocation decision depending on the subordinates' realized skills. Since the subordinate's incentive in external hiring is provided through management effort externality, assume that the subordinates get paid contingent on the outcome of their own task, which is determined by the allocated management decision. In essence, external hiring deals with the two incentive problems separately: inducing desirable effort choices by tying the superior and subordinates to the final outcome.

3.2.3 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 sepa-

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

[8]It is also possible to compensate the unpromoted agent contingent on the final outcome of the task. But that makes the internal promotion less attractive as the required promotion bonus $\overline{U} > 0$ increases as the unpromoted agent's payoff increases.

[9]Although, the bonus scheme may differ depending on firms and a CEO's contract terms, the promotion bonus is quite common in practice for internally promoted CEOs (Equilar (2013)).

ration. For expositional ease, I will use the terms internal promotion and IP, similarly external hiring and EH interchangeably.

Internal Promotion: Suppose the principal chooses internal promotion. The principal solves for the compensation contract and management allocation that characterizes promotion bonus $\overline{U}$ and losing prize v to maximize the principal's payoff. Given the promotion bonus, the following condition characterizes the incentive compatibility constraint in stage 1.

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

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

$$\overline{U}\geq v+\frac{2c}{p-q} \tag{3.1}$$

Intuitively, the promotion bonus, $\overline{U}$, increases in the losing prize, v, which is well-known in the 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 $v=0$.

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

Here, I define the promotion bonus as what a promoted CEO can enjoy when he has expertise.

$$\Big(E_{\theta,m}[w] - C(m_1 - d\theta_1, m_2)\Big) - \Big(E_{\theta,m}[w] - C(m_1, m_2)\Big) = C(m_1, m_2) - C(m_1 - d\theta_1, m_2)$$

The interpretation of the promotion bonus is that, all else equal, the promoted superior strictly prefers to have expertise. The saved cost (from his expertise) generates organizational slack in stage 2 depending on the realized skills in stage 1. For the organizational slack to provide investment incentive in stage 1 under IP, the principal solves the following program. In equilibrium, the superior's contract will exhaust all feasible management effort for the two tasks. Thus $m_2 = 1 - m_1$. All the bold symbols denote vectors.

$$max_{m,w} \;\; E[\mathbf{X} - \mathbf{w}] \tag{IP}$$

$$E[\mathbf{w}] - C(m_1 - d\theta_1, m_2) \geq \overline{u} \tag{IR-2}$$

$$\mathbf{m} \in argmax_n \;\; E_n[\mathbf{w}] - C(n_1 - d\theta_1, n_2) \tag{IC-2}$$

$$C(m_1, m_2) - C(m_1 - d\theta_1, m_2) \geq \frac{2c}{p-q} \tag{Slack}$$

Note that the two tasks are independent,

$$E[\mathbf{X} - \mathbf{w}] = \mathbf{Pr}(\mathbf{X_{SS}}|\theta, \mathbf{m})(\mathbf{X_{SS}} - \mathbf{w_{SS}}) + \mathbf{Pr}(\mathbf{X_S}|\theta, \mathbf{m})(\mathbf{X_S} - \mathbf{w_S}) \tag{3.2}$$

where $Pr(X_{SS}|\theta, \mathbf{m}) = f(\theta_1 + m_1)f(\theta_2 + m_2)$, $Pr(X_S|\theta, \mathbf{m}) = f(\theta_1 + m_1)(1 - f(\theta_2 + m_2)) + (1 - f(\theta_1 + m_1))f(\theta_2 + m_2)$. Basically, the promotion bonus refers to potential cost saving for the superior that increases his payoff indirectly.

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. For expositional convenience, assume that, in case of $\theta_1 \neq \theta_2$, it is the task 1 whose H.

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

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

which yields,

$$v \geq \frac{c}{Pr(X_S|I = c) - Pr(X_S|I = 0)} \tag{3.3}$$

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.

Under the EH, the principal solves the following program to maximize her payoff.

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

$$E[\mathbf{w}] - C(m_1, m_2) \geq \overline{u} \tag{IR-2}$$

$$\mathbf{m} \in argmax_n \quad E_n[\mathbf{w}] - C(n_1, n_2) \tag{IC-2}$$

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

Now, the principal's problem is complex because the incentive problem in stage 2 is interlinked 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. To provide intuition for the key trade-off, a numerical example precedes the analysis.

3.2.4 Numerical Example

Let $S = 10, F = 0, f(\theta + m) = \sqrt{\frac{\theta+m}{5}}, C(m_1, m_2) = \sum_i \frac{1}{2}m_i^2, H = 0.5, L = 0, c = 0.01, p = 0.6, q = 0.2, \overline{u} = 0$. Then, the program for external hiring is,

$$\max_{m,w,v} \; E_\theta\Big[\sum_i \sqrt{\frac{\theta_i + m_i}{5}}\big(X - w^\theta - v\big)\Big] \tag{EH}$$

$$\sum_i \sqrt{\frac{\theta_i + m_i}{5}} w^\theta - C(m_1(\theta), m_2(\theta)) \geq \overline{u} \tag{IR-2}$$

$$m \in argmax_n \;\; \sum_i \sqrt{\frac{\theta_i + m_i}{5}} w^\theta - C(n_1, n_2), \;\; \forall \theta_i, i = 1, 2 \tag{IC-2}$$

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

Then, the optimal management allocation is: $\mathbf{m} = (0.5, 0.5)$ if $\theta_1 = \theta_2$, $\mathbf{m} = (0.43, 0.57)$ if $\theta_1 > \theta_2$, and $\mathbf{w} = (w^{HH}, w^{HL}, w^{LL}) = (2.24, 1.89, 1.58)$, $v = 0.22$, and the principal's payoff is 6.96.

Similarly, the program for internal promotion is,

$$\max_{m,w} \; E_\theta\Big[\sum_i \sqrt{\frac{\theta_i + m_i}{5}}\big(X - w^\theta\big)\Big] \tag{IP}$$

$$\sum_i \sqrt{\frac{\theta_i + m_i}{5}} w^\theta - C(m_1(\theta) - d\theta_1, m_2(\theta)) \geq \overline{u} \tag{IR-2}$$

$$m \in argmax_n \;\; \sum_i \sqrt{\frac{\theta_i + m_i}{5}} w^\theta - C(n_1, n_2), \;\; \forall \theta_i, i = 1, 2 \tag{IC-2}$$

$$E_\theta\Big[\frac{1}{2}\big(m_1^2(\theta) - (m_1(\theta) - d\theta_1)^2\big)\Big] \geq \frac{2c}{p - q} \tag{Slack}$$

Then, depending on d, the optimal management allocation, pay, and the principal's payoffs are found and summarized in Table 3.1. As the benefit of expertise becomes more efficient (large d), the small distortion in management creates large organizational slack, thus performing better than external hiring ($7.07 > 6.96$). However, as d decreases, the required slack tends to lead to

	$\boldsymbol{d = 0.2}$	$\boldsymbol{d = 0.18}$	$\boldsymbol{d = 0.16}$
$(m_1(\theta_H), m_2(\theta_H))$	(0.78,0.22)	(0.85,0.15)	(0.95,0.05)
$(m_1(\theta_H), m_2(\theta_L))$	(0.54,0.46)	(0.6,0.4)	(0.66,0.34)
$(m_1(\theta_L), m_2(\theta_L))$	(0.5,0.5)	(0.5,0.5)	(0.5,0.5)
w^{HH}	2.14	2.24	2.42
w^{HL}	1.71	1.76	1.85
w^{LL}	1.58	1.58	1.58
Principal	**7.07**	**6.96**	**6.79**

Table 3.1: Numerical solution for Internal promotion

too much distortion, thereby performing worse than external hiring ($6.79 < 6.96$).

3.3 Analysis

3.3.1 Management Effort Choice and Compensation

Management Effort Allocation Problem

In this section, I characterize management effort choices 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. With θ as given, the revenue of a firm is,

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

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,

$$m_1 = \frac{1-\theta_1+\theta_2}{2}, \quad m_2 = \frac{1+\theta_1-\theta_2}{2} \tag{3.4}$$

For instance, if $\theta_1 = \theta_2$ then $m_1 = m_2$,or if $(\theta_1, \theta_2) = (H, L)$, then $m_1 < m_2$. This will generate equal likelihood of getting success by implementing balanced management allocation. This is the first best when there is no agency friction at the superior level. To compare the benefit and the cost of expertise, I first consider external hire's compensation where expertise is not exploited at the management stage.

External Hiring

This subsection finds an optimal compensation to implement management allocations that are determined by production technology and the compensation for subordinates under the external hiring regime. One can easily see that since the external hire has no preference for any task, it is optimal to implement balanced management allocation. Then, the optimal allocation is: $\frac{1}{2}$ if $\theta_1 = \theta_2$ and $\left(\frac{1-\theta_1+\theta_2}{2}, \frac{1+\theta_1-\theta_2}{2}\right)$. From the first order condition of the external hire, this is easy to implement as,

$$\begin{aligned} m_1 \text{ such that } & \big(f'(\theta_1+m_1)f(\theta_2+1-m_1) - f(\theta_1+m_1)f'(\theta_2+1-m_1)\big)w^{\theta}_{SS} \\ & + \big(f'(\theta_1+m_1) + f'(\theta_2+1-m_1)\big)w^{\theta}_{S} = 0 \end{aligned}$$

which is true if allocated resources are the same across tasks. Thus, for the external hire, an optimal pay is determined by his (IR) constraint.

$$f^2\big(\frac{1+\theta_1-\theta_2}{2}\big)\big(w^{\theta}_{SS}-2w^{\theta}_{S}\big)+2f\big(\frac{1+\theta_1-\theta_2}{2}\big)w^{\theta}_{S}-C\big(\frac{1-\theta_1-\theta_2}{2},\frac{1+\theta_1-\theta_2}{2}\big)\geq\overline{u}$$

which will be binding.

To find the pay for subordinates, v, recall (3.3). Let $Pr(X_S|I=c)-Pr(X_S|I=0)\equiv\Delta$. Then, $v=\frac{c}{\Delta}$. Given that the balanced management will be implemented, Δ can be written as follows.

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

The required pay for subordinates in external hiring regime consists of two sources of incentives. One is a subordinate's intrinsic incentive to increase the success likelihood through their skill, and the other is a free-riding incentive. Since the balanced allocation will be implemented, a subordinate knows that he will get extra resources ($\frac{H-L}{2}$) in case the other subordinate is H. Given $p>\frac{1}{2}$, v increases as free-riding incentive increases. The following proposition summarizes the discussion.

Proposition 10. *Under the external hiring regime, an optimal management allocation is always balanced. For* θ_1,θ_2, $(m_1^*,m_2^*)=(\frac{1-\theta_1+\theta_2}{2},\frac{1+\theta_1-\theta_2}{2})$*. The externally hired CEO's (IR) is binding* *with expected compensation cost of* $\overline{u}+C(m_1^*, m_2^*)$. *The subordinates' wage is*

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

Internal Promotion

This subsection finds an optimal compensation to implement management allocations that are determined by production technology and the compensation for subordinates under the internal promotion regime. Now, whether the balanced management allocation still remains efficient is unclear. This is because of organizational slack that is required to incentivize subordinates at the firm-specific skill acquisition stage.

I consider two cases: 1) the (Slack) constraint does not bind, 2) the (Slack) constraint binds. First, if the (Slack) constraint does not bind, then the principal can implement the same management allocation as in external hiring but with the cheaper cost of $\overline{u}+C(m_1-\theta_1, m_2)$. This is true if d is sufficiently high. In this case, internal promotion strictly dominates external hiring. The following lemma summarizes this discussion.

Lemma 15. *If* d *is sufficiently large that the (Slack) does not bind, then internal promotion always dominates external hiring.*

Next if the (Slack) constraint binds, then, this implies that the balanced allocation may not be optimal. This is clear from the first order condition:

$$\begin{aligned}
&\big(f'(\theta_1+m_1)f(\theta_2+1-m_1)-f(\theta_1+m_1)f'(\theta_2+1-m_1)\big)(X_{SS}-2X_S)\\
&+\big(f'(\theta_1+m_1)+f'(\theta_2+1-m_1)\big)X_S+\lambda C'(m_1-d\theta_1, 1-m_1)=0
\end{aligned}$$

where $\lambda > 0$ denotes a Lagrange multiplier for the (Slack) constraint and C' denotes derivative of C with respect to m_1. Since $f(\cdot)$ is concave, $f'(x) > f'(y)$ if $x < y$, and $C(\cdot)$ is convex increasing, thus $C'(\cdot) > 0$. Thus, for any θ, $m_1 \geq 1/2$. Observe that the above first order condition depends on θ and d. First, if d is large, implying that the management cost reduction from the firm-specific skill is big, then $C'(m_1 - d\theta_1, 1 - m_1)$ is small, thus the distortion to satisfy the first order condition becomes small. However, as d decreases, $C'(m_1 - d\theta_1, 1 - m_1)$ increases, thereby making the distortion to satisfy the first order condition becomes larger. Next, consider firm-specific skills. If $\theta_1 = \theta_2 = L$, then, there is no distortion since there will be no cost reduction at management stage, thus it is optimal to allocate resources evenly. However, the distortion in management occurs if there is at least one H. Interestingly, the distortion becomes larger (i.e., management resource is allocated toward task 1 more) if both tasks are H.

Lemma 16. *The management distortion becomes smaller if d is large, and the distortion becomes larger if $\theta_2 = H$: $\frac{\partial m_1^*}{\partial d} < 0$, $\frac{\partial m_1^*}{\partial \theta_2} > 0$.*

This seems first counterintuitive as internal promoting firm creates more distortion when the other task looks profitable ($\theta_2 = H$). However, the intuition is the following. If task 2 also has H, then the marginal cost of shifting resources from task 2 to task 1 becomes cheaper than the case where task 2 has $\theta_2 = L$ as $f(\cdot)$ is concave. The cost of shifting management effort from task 2 to task 1 becomes larger if task 2 has $\theta_2 = L$.

Proposition 11. *Under the internal promotion regime, an optimal management allocation is balanced if the (Slack) constraint does not bind. If the (Slack) constraint binds, then, $m_1^* > \frac{1}{2} > m_2^*$ if $\theta_1 = H \geq \theta_2$, and $m_1^* = m_2^* = \frac{1}{2}$ if $\theta_1 = \theta_2 = L$. Given that $\theta_1 = H$, $m_1^*(\theta_2 = H) > m_1^*(\theta_2 = L)$. The total compensation cost is $\overline{u} + C(\frac{1-\theta_1+\theta_2}{2} - d\theta_1, \frac{1+\theta_1-\theta_2}{2})$*

3.3.2 Organizational Efficiency

Firm Performance: Internal Promotion vs External Hiring

I first compare the optimal compensation for each hiring regime. Since the external hiring regime implements a balanced allocation, I use m_i to denote the allocation for task i. But internal promotion implements an unbalanced allocation, which depends on the realized θ. For internal promotion, I use m_i^θ.

Lemma 17. *The optimal compensation for external CEO is*

$$w_1^{HH} = w_2^{HH} = \frac{C'(1/2, 1/2)}{f'(H+1/2)}, \quad w_1^{HL} = w_2^{HL} = \frac{C'(m_1, 1-m_1)}{f'(H+m_1)} = \frac{C'(m_1, 1-m_1)}{f'(L+1-m_1)}$$
$$w_1^{LL} = w_2^{LL} = \frac{C'(1/2, 1/2)}{f'(L+1/2)}$$

The optimal compensation for internal CEO is

$$w_1^{HH} = \frac{C'(m_1^{HH} - dH, 1-m_1^{HH})}{f'(H+m_1^{HH})} > w_2^{HH} = \frac{C'(m_1^{HH} - dH, 1-m_1^{HH})}{f'(H+1-m_1^{HH})},$$
$$w_1^{HL} = \frac{C'(m_1^{HL} - dH, 1-m_1^{HL})}{f'(H+m_1^{HL})}, w_2^{HL} = \frac{C'(m_1^{HL} - dH, 1-m_1^{HL})}{f'(L+1-m_1^{HL})}, \quad w_1^{LL} = w_2^{LL} = \frac{C'(1/2, 1/2)}{f'(L+1/2)}$$

Since the external CEO implements the balanced allocation, he is compensated equally regardless of the firm-specific skills of the two tasks. For instance, even if one task is strictly less profitable than the other (say $\theta_1 = H, \theta_2 = L$), still, the external CEO gets the same level of pay from the success of both tasks. For internal promotion, since the internal CEO puts more resources into his own task, the pay sensitivity of his own task is much higher. Technically, this is due to $f(\cdot)$ being concave and $C(\cdot)$ being convex. Intuitively, this is because the marginal productivity of extra resources that the internal CEO puts into his own task is small.

Given that the expected compensation tends to be lower in the internal promotion regime than in the external hiring regime, whether internal promotion dominates the external hiring depends on the production consequence of the unbalanced allocation. Since low expertise (small d) leads to large distortion in the management allocation, this can reduce the expected firm performance (therefore the principal's expected payoff). To highlight the production consequence of management allocation, hereafter I use $\mathbf{m} \in \{\mathbf{m^B}, \mathbf{m^{UB}}(d)\}$ to denote the optimal allocation for external hiring and internal promotion respectively. Let $\Pi(\mathbf{m})$ denote a firm's expected payoff depending on $\mathbf{m}$.

$$\Pi(\mathbf{m^B}) = E[\mathbf{X}|\mathbf{m^B}] - C(m_1^B, 1 - m_1^B) - 2f(\theta_1 + m_1)v$$
$$\Pi(\mathbf{m^{UB}}(d)) = E[\mathbf{X}|\mathbf{m^{UB}}] - C(m_1^{UB} - d\theta_1, 1 - m^{UB})$$

Then, the external hiring strictly dominates internal promotion if $\Pi(\mathbf{m^B}) > \Pi(\mathbf{m^{UB}}(d))$, otherwise, internal promotion dominates.

Proposition 12. *As (i) firm-specific skill becomes less effective in management (d low), (ii) the synergy between the two tasks becomes greater ($X_{SS} - 2X_S$ large), then the external hiring performs better than the internal promotion.*

Since the internal CEO enjoys the expertise benefit, which will decrease his management cost after being promoted, the expected total compensation cost is always lower than the compensation cost for external hire. However, this compensation cost saving is not free: to create enough incentives for subordinates, the required organizational slack through the unbalanced allocation in the management stage might incur production inefficiency. That is, the key tension of internal promotion is between management allocation (organizational slack) and incentive efficiency (aggregation through tournament).

Comparative Statics

This subsection investigates when the demand for external hiring increases, why external hire seems to be paid more, and how the principal's payoff in external hiring seems to be lower than the payoff in internal promotion. To highlight the trade off between allocational efficiency and incentive efficiency, I focus on $\mathbf{X}$ and d for this section.

Demand for External Hiring: Note that d measures efficiency of firm-specific skills in management and that $X_{SS} - 2X_S > 0$ measures synergy or interconnectedness between the two tasks. One can examine when the demand for external hiring or internal promotion changes depending on these parameters. The next proposition identifies such conditions under which external hiring tends to be preferred as these parameters change.

The intuition is as follows. Broadly, if d is high, then incentive aggregation generally dominates incentive separation. Although incentive aggregation may incur allocational inefficiency, as long as incentive efficiency is large enough, internal promotion remains as optimal. As d becomes small however, the incentive aggregation leads to more severe allocational inefficiency through the biased management. Since the biased management reduces the chance of getting synergy $(X_{SS} - 2X_S)$, as interrelation between the two tasks increases, the allocational efficiency dominates, thereby preferring external hiring. The Figure 3.2 depicts this finding.

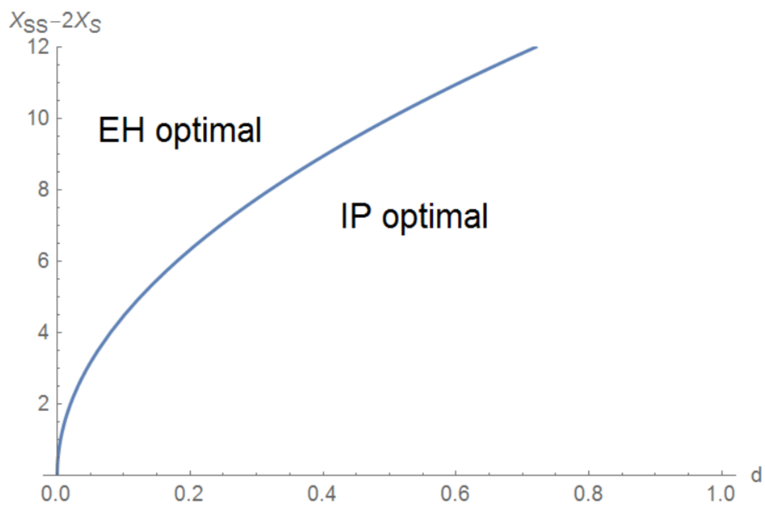

Figure 3.2: **Dominance of CEO Appointment Choice**

Figure 3.2 depicts when internal promotion dominates external hiring depending on the management efficiency of firm-specific skills (d) and synergistic relations between the two tasks ($X_{SS} - 2X_S$). If the skill expertise is sufficiently efficient (d high), then internal promotion dominates external hiring. The efficiency of internal promotion decreases as d decrease and $X_{SS} - 2X_S$ increases.

This comparative statics is consistent with empirical observations. According to a recent study on CEO succession (Custódio et al. (2013), Strategy& (2016)), the percentage of internal promotion is higher in industries where a promoted CEO can effectively manage using his expertise, including IT and financials. However, the dominance of internal promotion shrinks as interconnectedness and/or synergies across business units become larger relative to the management efficiency. For instance, the percentage of outside CEOs is higher in telecommunication services and utilities where each unit's operation is closely related or generates high synergy between units.

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 does both management allocation and firm-specific skill investment: i.e the external CEO's total effort cost is $C(m_1, m_2)$, while the internal CEO's total cost is $c + C(m_1 - d\theta_1, m_2)$. To make a fair comparison, define the following:

$$v \equiv \text{Pay premium} = \frac{\text{expected total pay}}{\text{total effort costs}} = \frac{E^{IP}[w]\mathbf{1}_{IP} + E^{EH}[w](1 - \mathbf{1}_{IP})}{(c + C(m_1 - d\theta_1, m_2)) \cdot \mathbf{1}_{IP} + C(m_1, m_2)(1 - \mathbf{1}_{IP})}$$

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)).[11] Again for notational convenience, I use m^B, m^{UB} to denote each CEO's different allocation. The pay premium for the external hire, v^{EH}, is,

$$v^{EH} = \frac{C'(m_1^B, 1 - m_1^B)}{C(m_1^B, 1 - m_1^B)} \left(\frac{1}{f'(\theta_1 + m_1^B)} + \frac{1}{f'(\theta_2 + 1 - m_1^B)} \right) = \frac{C'(m_1^B, 1 - m_1^B)}{C(m_1^B, 1 - m_1^B)} \frac{2}{f'(\theta_1 + m_1^B)}$$

Similarly,

$$v^{IP} = \frac{C'(m_1^{UB}, 1 - m_1^{UB})}{c + C(m_1^{UB}, 1 - m_1^{UB})} \left(\frac{1}{f'(\theta_1 + m_1^{UB})} + \frac{1}{f'(\theta_2 + 1 - m_1^{UB})} \right)$$

That is, if $v^{EH} > v^{IP}$ is observed, then this will be interpreted as the external CEO gets paid more relative to the internal CEO. Combined with the Proposition 12, 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): The external CEO appears to get paid more relative to the internal CEO. This over-payment illusion is particularly large when*

[11] I assume that $\beta lnSales$ in Murphy and Zabojnik (2007) is a proxy for required effort costs.

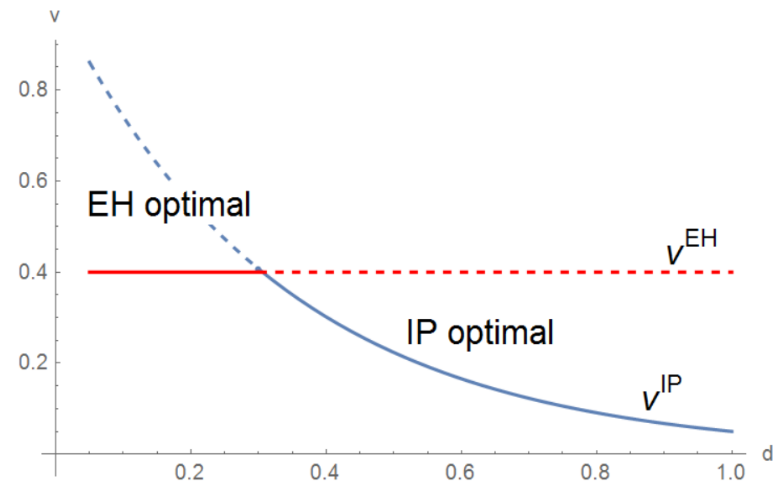

Figure 3.3: Overpayment Illusion
This figure depicts v^{IP}, v^{EH}, the pay premium for each CEO. The solid line denotes a pay premium in equilibrium and the dashed line denotes a counterfactual pay premium which is unobservable in equilibrium.

$(\theta_1, \theta_2) = (L, L)$.

The intuition is as follows. Given that IP dominates EH for large d, the principal enjoys the benefit from incentive aggregation, and the total effort cost of the internal CEO already lowers his pay premium (in particular, in case of (L, L)). Interestingly, the impact of the expertise (d) on the pay premium is ambiguous. While the expertise benefit reduces the specialist CEO's management cost on his task, it increases the total management cost as the specialist CEO now puts more resources into the other task. Moreover, as the expertise benefit increases, there will be less biased allocation. While the less biased allocation increases the marginal productivity of his own task (since $f(\cdot)$ concave), it also reduces the marginal productivity of the other task, thereby making the impact of d on pay premium unclear. Overall, due to the difference from total effort costs, which makes $v^{EH} > v^{IP}$ more likely, one might conclude that the outsider gets paid more. As the internal CEO's allocation becomes biased, his pay premium increases (since $f'(x)$ increases as x decreases). However, given that the principal optimally chooses the hiring decision depending on its production and d, the case where insider's pay premium is higher than the outsider might not be observed if the biased allocation incurs too much production losses.

Proposition 14. *(Implications for shareholder return): Fix $X_{SS} - 2X_S$ and consider two firms with high d and low d. Then, the principal's payoff under the external hiring (low d) appears to suffer relative to the principal's payoff under the internal promotion (high d).*

The intuition is again related to the Proposition 12. Since IP dominates EH for large d, the management allocation bias is smaller, leading to $E[\mathbf{X}|\mathbf{m}^{\mathbf{UB}}]$ close to $E[\mathbf{X}|\mathbf{m}^{\mathbf{B}}]$. In that case, incentive aggregation strictly lowers the total compensation cost, making the principal's net payoff higher. As d decreases, the required distortion in the internal promotion regime increases, thus making internal promotion worse than the external hiring. Then, low d firm turns to external hiring, which incurs on average higher compensation costs than internal promotion, thus making the external hiring principal's payoff lower than high d principal. However, if the low d principal implemented internal promotion, the consequence of production would have been much worse than external hiring. Therefore, the puzzling evidence on pay and firm performance of external CEO relative to internal CEO can be understood as different firms' different optimal responses to their CEO appointment choice taking into account the trade-off between incentive efficiency and management efficiency.

3.4 Empirical Implication and Discussion

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 incentivize subordinates for firm-specific skill acquisition, 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 (e.g., expertise benefit, d and synergistic relations, $X_{SS} - 2X_S$), 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 the expertise benefit from firm-specific skill is 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.

It is worth noting that the production technology ($f(\cdot)$) in this paper strictly prefers a balanced management allocation. In practice, however, it is not hard to observe that the externally hired CEO pursues restructuring or reshaping an organization including divestiture (See Weisbach (1995), Denis and Denis (1995), Custódio et al. (2013) for example). This can occur if the production technology becomes convex, which makes an unbalanced allocation optimal. The main argument and trade-off between allocational efficiency and incentive efficiency this paper proposes is based on the presumption that the firm owner prefers to implement a balanced management to maintain the structure of the organization.

3.5 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 generalist CEO might not be a *de facto* premium, and relatively mediocre return to firm owners is interlinked with the organizational incentive problems for multiple executives.

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 stage 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 through an unbalanced management allocation will never be optimal for the principal after the skills are realized, i.e. the principal has no incentive to introduce internal promotion after stage 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 incentive efficiency through aggregation decrease, but also the impact of biased allocation 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 Discussion on existing literature

This section reviews three strands of literature related to this paper: 1) managerial project selection 2) reputational concerns and project selection, and 3) applied works on competitive assignment. As most of these areas have been largely studied, I will discuss only the part of each that has direct connection to this paper.

A.1.1 Managerial Project Selection Problem

One of the fundamental agency frictions in this paper is a manager's project choice problem. The main goal of this literature is to find a way to align the investment preference of an agent with that of a principal. Lambert (1986) considers an induced moral hazard model where the agent is supposed to exert effort in order to acquire information about investment projects to adapt an either risky or safe project. Since the effort does not shift the distribution of the outcome in a first order stochastic dominance sense, this particular form of moral hazard problem requires a different optimal contract from the standard moral hazard problem. Holmstrom (1989) and Aghion and Tirole (1994) also consider an agent choosing a project between risky and less-risky (innovative and non-innovative) projects. In Holmstrom, on account of a measurability issue, an optimal incentive becomes less sensitive to performance in order to motivate a risky project. On the other hand, Aghion and Tirole examine an optimal allocation of control changes if a more risky project is preferred. Manso (2011) considers a similar set up but allows the agent to learn about production technology, finding that the adaption of a risky project is induced by a contract that has high tolerance for early failure. Inderst and Klein (2007) also consider an agent exerting effort to acquire information about projects, but the key focus is on the interaction between project choice and capital budgeting. These papers share the similar feature with the present paper, namely, a project selection with different riskiness has a distinct feature from a standard moral hazard problem. However, the present paper examines why a manager ends up having such a preference for risk by introducing career concerns.

A.1.2 Reputational Concerns and Project Selection

Managerial reputational concern, one of the key ingredients in this paper, is ubiquitous. A manager's past track record as an executive forms his current reputation, as well as his potential capability and suitability as a manager. Previous research has shown that this reputational concern can incentivize a manager to make a current effort choice, what is called a reputation building incentive (See, Fama (1980), Holmstrom and Ricart I Costa (1986), Holmstrom (1999), Gibbons and Murphy (1992)). Extending this literature, this paper also explores a latent aspect of managerial reputational incentive. That is, if a manager's track record (i.e., reputation) is highly subject to a negative shock, which can potentially harm his or her reputation, then it can also incentivize the manager to avoid these situations. More specifically, even when the manager himself knows that there is an ex ante profitable risky project, the potential risk might hinder the manager from taking the project to avoid his reputation suffering upon an unsatisfactory outcome, what I shall call a *reputation maintaining incentive*. This feature of reputation and the potential distortion thereof have also been studied in the literature (e.g., Kanodia et al. (1989)). I extend this work by endogenizing reputational concerns and introducing the labor market for managers that provide a foundation of why a manager cares about their reputation.

Career concern literature, starting with Fama (1980) and Holmstrom (1999), assumes away information asymmetry to capture a reputation building incentive. In a large set of follow-up studies to these, Holmstrom and Ricart I Costa (1986) argues that the agent's career concern, in contrast to moral hazard, causes an incongruity in investment preferences. Similarly, Narayanan (1985) shows how the agent is entrenched in the shortermism in an attempt to boost his perceived ability. Scharfstein and Stein (1990) investigates how a manager would show a herding behavior in their project choice in order to signal that he/she learns what other talented managers learn. Among other results, Jeon (1998) proves that, because of future wage concern formed by the updated reputation, the agent entirely ignores his potentially beneficial signal in making investment decision. Among more recent studies, 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, Zwiebel (1995) rationalizes corporate conservatism building on herding incentives (i.e., choosing a project that is most chosen by others) so as to have accurate yard stick. Other studies with information asymmtry 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.[1] 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

[1]The 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.

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 realization 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 managers due to heterogeneous firms and by examining the interplay between such career concerns and market-wide rematching patterns.

A.1.3 Applied Models of Competitive Assignment Framework

By extending Terviö (2008), Gabaix and Landier (2008) derive the functional forms of the CEO talents distribution, and calibrate their model to show that the increase in pay levels are attributed to the increase in aggregate firm size. In the same spirit of a competitive assignment framework without agency problems, but with repeated interactions, Eisfeldt and Kuhnen (2013) examine the effect of industry performance on CEO succession patterns to highlight that the change in the required skill sets of management makes CEO turnover more likely. With the absence of agency problems, both Terviö (2008) and Gabaix and Landier (2008) analyze a static competitive equilibrium model. Eisfeldt and Kuhnen (2013) consider the dynamics of a market equilibrium upon an industry shock without agency problems. In my paper, the attribute of a manager (i.e., a manager's reputation) evolves whenever its matched firm performance is realized, thus influencing the agency problem in a dynamic assignment framework.

In a two-sided matching framework with risk-averse agents, Serfes (2008) shows that depending on how a matching surplus function is built, the equilibrium matching pattern is either positive or negative assortative, or non-assortative. As in Serfes (2008), the focus in this paper is also to analyze equilibrium matching patterns. Contrary to Serfes (2008), this paper focuses on why risk neutral agents form induced risk aversion depending on the size distribution of firms and the associated matching patterns.

A.2 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 $\gamma \in [\underline{\gamma}, \overline{\gamma}]$ 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 $1 < M(\overline{\gamma} - \underline{\gamma})$, 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.2.1 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 $G : \Gamma_{t-1} \to [0, 1]$, the posterior distribution becomes,

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

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.2.2 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.

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

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

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

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 $\gamma-$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 $\gamma-$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,

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

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

$$Y_1(S,\gamma) - \omega_1(\gamma) \geq Y_1(S,\gamma^{'}) - \omega_1(\gamma^{'}) \ \ \forall S,\gamma \tag{SC(S,γ)}$$

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

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

where $\Pi(\gamma^X) = w_0 + E\Big[\int_{\underline{\gamma}}^{\gamma^X} \frac{\partial}{\partial x} Y_1(S,x)\Big(\frac{d}{dz}\int_{\underline{\gamma}}^{z} \sum_{x\in\{f^{-1}(s),g^{-1}(s),s\}} tr(s|x)g_{t=1}(x)ds\Big)dz\Big]$. 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.3 Proofs

A.3.1 The Derivation of Event Probability

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

$$
\begin{aligned}
Pr(h|\tau = B) &= Pr(\upsilon = r) \times Pr(\hat{\upsilon} = r, (I_r, h)|\upsilon = r, \tau = B) \\
&= Pr(h|I_r, \upsilon = r) \times Pr(\upsilon = r)Pr(\hat{\upsilon} = r, |\upsilon = r, \tau = B) \\
&= p\alpha\beta = \frac{1}{2}p\beta \\
Pr(l|\tau = B) &= Pr(\upsilon = r) \times Pr(\hat{\upsilon} = r, (I_r, l)|\upsilon = r, \tau = B) + Pr(\upsilon = s) \times Pr(\hat{\upsilon} = r, (I_r, l)|\upsilon = s, \tau = B) \\
&= Pr(l|\upsilon, I_r) \times Pr(\upsilon = r)Pr(\hat{\upsilon} = r, |\upsilon = r, \tau = B) \\
&+ Pr(l|\upsilon = s, I_r) \times Pr(\upsilon = s)Pr(\hat{\upsilon} = r|\upsilon = s, \tau = B) \\
&= (1-p)\alpha\beta + 1 \times (1-\alpha)(1-\beta) = \frac{1}{2}(1-p\beta) \\
Pr(m|\tau = B) &= Pr(\upsilon = s) \times Pr(\hat{\upsilon} = s, I_S|\upsilon = s, \tau = B) + Pr(\upsilon = r) \times Pr(\hat{\upsilon} = s, I_s|\upsilon = r, \tau = B) \\
&= (1-\alpha)\beta + \alpha(1-\beta) = \frac{1}{2}
\end{aligned}
$$

□

A.3.2 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

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

□

A.3.3 Proof of Lemma 1

Proof. Recall that for a given γ, the manager is either $\tau = G$ with probability γ, or $\tau = B$ with probability $1 - \gamma$. If it is $\tau = G$ and the manager 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

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

Thus, upon hiring manager γ, the expected revenue is,

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

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

A.3.4 Proof of Lemma 2

Proof. Without loss of generality, I will use $\omega_t(\gamma)$ to denote the price to hire manager γ. If there is rent to induce $e = H$, then the expected total compensation cost will be $\omega_t(\gamma)$+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 managers is,

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

$\square$

A.3.5 Proof of Lemma 3

Proof. I omit time subscript $t = 3$. Rearrainging (IC) constraint for each $x = 1$ and 0 case is the following.

$$\begin{aligned} &Pr(h|\gamma)(w^h - w^m) - Pr(l|\gamma)(w^m - w^l) \geq c, \\ &-(Pr(h) - Pr(h|\gamma))(w^h - w^l) + \frac{1}{2}(w^m - w^l) \geq c \end{aligned}$$

Use $Pr(h|\gamma) + Pr(l|\gamma) = 1/2$ and rearrange the terms, then

$$2Pr(h|\gamma)\Big((w^h - w^l) - c\Big) \geq w^m - w^l \geq 2\Big((Pr(h) - Pr(h|\gamma))(w^h - w^l) + c\Big)$$

This implies that $w^m - w^l$ is bounded (both below and above). Graphically, this is described

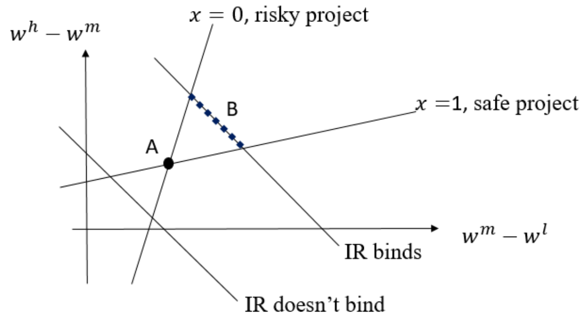

Figure A.1: Graphical Description of Feasible Solutions

in Figure A.1. Thus, the minimum incentive compatible wage is characterized as equalizing both constraints (i.e., point A). Then, there are two cases: the (IR) binds or not. First, consider the case where the (IR) binds. Then, the minimum incentive compatible solution A is feasible. In this case, the optimal solution is uniquely found by A, i.e., $w^h - w^m = \frac{c(2-p)}{2Pr(h|\gamma)-Pr(h)}$, $w^m - w^l = \frac{cp}{2Pr(h|\gamma)-Pr(h)}$. Thus,

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

This occurs when $\omega(\gamma)$ is small. To see this, plug the above solution into a manager's expected payoff.

$$\omega(\gamma) \leq -c + E[w^X|e = H] = \frac{cp}{2Pr(h|\gamma) - Pr(h)}$$

Next, consider the case where the (IR) does not bind, which happens if $\omega(\gamma)$ is large. Thus, the minimum incentive compatible solution A is not feasible. In this case, there can be multiple solutions, and any wage scheme on the line B is both incentive compatible and individually rational, but the incentive compatible constraint does not bind. Formally, if $\omega(\gamma) > \frac{cp}{2Pr(h|\gamma)-Pr(h)}$, then any wage on B is both individually rational and incentive compatible.

□

A.3.6 Proof of Lemma 5

Proof. From Lemma 3, we have

$$
\begin{aligned}
&Pr(h|\gamma)(w^h + \omega_3(\gamma^h) - w^m - \omega_3(\gamma^m)) - Pr(l|\gamma)(w^m + \omega_3(\gamma^m) - w^l - \omega_3(\gamma^l)) \geq c,\\
&-(Pr(h) - Pr(h|\gamma))(w^h + \omega_3(\gamma^h) - w^l - \omega_3(\gamma^l)) + \frac{1}{2}(w^m + \omega_3(\gamma^m) - w^l - \omega_3(\gamma^l)) \geq c
\end{aligned}
$$

Rearrange to distinguish the future payoff. Then,

$$
\begin{aligned}
&\Big(Pr(h|\gamma)(\omega_3(\gamma^h) - \omega_3(\gamma^m)) - Pr(l|\gamma)(\omega_3(\gamma^m) - \omega_3(\gamma^l))\Big) + Pr(h|\gamma)(w^h - w^m) - Pr(l|\gamma)(w^m - w^l) \geq c,\\
&\Big(-(Pr(h) - Pr(h|\gamma))(\omega_3(\gamma^h) - \omega_3(\gamma^l)) + \frac{1}{2}(\omega_3(\gamma^m) - \omega_3(\gamma^l))\Big)\\
&\qquad -(Pr(h) - Pr(h|\gamma))(w^h - w^l) + \frac{1}{2}(w^m - w^l) \geq c
\end{aligned}
$$

If the first term (the future payoff) is positive, then it relaxes the incentive compatibility constraints, and the negative first term increases the required pay to make it incentive compatible. Use $Pr(h|\gamma) + Pr(l|\gamma) = 1/2$, and rearrange the first terms of both inequalities with respect to $\omega_3(\gamma^m) - \omega_3(\gamma^l)$, we have

$$
2Pr(h|\gamma)\Big(\omega_3(\gamma^h) - \omega_3(\gamma^l)\Big) > \omega_3(\gamma^m) - \omega_3(\gamma^l) > 2(Pr(h) - Pr(h|\gamma))\Big(\omega_3(\gamma^h) - \omega_3(\gamma^l)\Big)
$$

□

A.3.7 Proof of Lemma 6

Proof. Due to supermodularity of expected revenue, the efficiency requires positive assortative assignment of managers 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. □

A.3.8 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 matcing 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 $S[i] > G^{-1}(1 - M(\gamma^i))$. □

A.3.9 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 $\Delta Y(S[m]) = \Big(Y(S[m], \gamma^m) - Y(S[m], \gamma^l)\Big)$, 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. $\square$

A.3.10 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 $S \in \Big[S_{min}, S[m]\Big)$. I now show that this is indeed stable. Notice that,

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

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 $\forall\; S \in \Big[S[m], S[h]\Big)$ 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

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

is an equilibrium in period 2. □

A.3.11 Proof of Proposition 4

Proof. Given the equilibrium in period 2, we have the following reputation in period 3.

$$\gamma^{hh} > \gamma^{hm} = \gamma^{mh} > \gamma^{mm} = \gamma^m > \gamma^{hl} = \gamma^{lh} > \gamma^{ml} = \gamma^l > \gamma^{ll}$$

Now, consider the expected market value change for each manager in the beginning of period 2. I will first consider manager γ^m. Define the difference between expected market value change upon $e = H$ and $e = L$ with I_s.

$$\begin{aligned}
R(p, \gamma^m) &\equiv E[\omega_3(\gamma^{mi})|e = H] - E[\omega_3(\gamma^{mm})|e = L, I_s] \\
&= \alpha p(\gamma^m + (1-\gamma^m)\beta)\Big(\omega_3(\gamma^{mh}) - \omega_3(\gamma^{ml})\Big) \\
&- \alpha\Big(\omega_3(\gamma^{mm}) - \omega_3(\gamma^{ml})\Big) \\
&= \alpha(p(\gamma^m + (1-\gamma^m)\beta)U(m) - D(m))
\end{aligned}$$

Plug the market value from the Lemma 2 into above. Since we are interested in the sign of the value function, normalize this by common parameter F and rearrange the terms. Then

$R(p, \gamma^m) < 0$ is equivalent to,

$$p(\gamma^m + (1-\gamma^m)\beta)\frac{\Big((\gamma^{mh} - \gamma^{mm})S[mh] + (\gamma^{mm} - \gamma^{hl})S[mm] + (\gamma^{hl} - \gamma^{ml})S[hl]\Big)}{(\gamma^{mm} - \gamma^{hl})S[mm] + (\gamma^{hl} - \gamma^{ml})S[hl]} < 1$$
$$p(\gamma^m + (1-\gamma^m)\beta)\Big(\frac{(\gamma^{mh} - \gamma^{mm})S[mh]}{(\gamma^{mm} - \gamma^{hl})S[mm] + (\gamma^{hl} - \gamma^{ml})S[hl]} + 1\Big) < 1$$

I show that the upper bound of the left hand side is less than 1. Since $S[mm] > S[hl]$, the left hand side is bounded above by $p(\gamma^m + (1-\gamma^m)\beta)\Big(\frac{(\gamma^{mh}-\gamma^{mm})S[mh]}{(\gamma^{mm}-\gamma^{ml})S[hl]} + 1\Big)$. Also, $\gamma^{mh} = \gamma^m \frac{Pr(h|G)}{Pr(h)}$, $\gamma^{ml} = \gamma^m \frac{Pr(l|G)}{Pr(l)}$. Thus, the upper bound can be written as,

$$\frac{Pr(h|\gamma^m)}{\alpha}\Big(\frac{\gamma^m(\frac{Pr(h|G)}{Pr(h)} - 1)S[mh]}{\gamma^m(1 - \frac{Pr(l|G)}{Pr(l)})S[hl]} + 1\Big) = \frac{Pr(h|\gamma^m)}{\alpha}\Big(\frac{Pr(l)Pr(h|B)}{Pr(h)Pr(l|B)}\frac{S[mh]}{S[hl]} + 1\Big)$$

Thus, a sufficient condition to have $R(p, \gamma^m) < 0$ is,

$$\frac{S[mh]}{S[hl]} < (\frac{\alpha}{Pr(h|\gamma)} - 1)\frac{Pr(h)Pr(l|B)}{Pr(l)Pr(h|B)}$$

Similarly, consider manager γ^h.

$$\begin{aligned}
R(p, \gamma^h) &= E[\omega_3(\gamma^{hi})|e = H] - E[\omega_3(\gamma^{hm})|e = L, I_s] \\
&= \alpha p(\gamma^h + (1-\gamma^h)\beta)\Big(\omega_3(\gamma^{hh}) - \omega_3(\gamma^{hl})\Big) \\
&- \alpha\Big(\omega_3(\gamma^{hm}) - \omega_3(\gamma^{hl})\Big) \\
&= \alpha(p(\gamma^h + (1-\gamma^h)\beta)U(h) - D(h))
\end{aligned}$$

Similarly, normalize by F and rearrange the terms. Then $R(p,\gamma^h) > 0$ is equivalent to,

$$
\begin{aligned}
&p(\gamma^h + (1-\gamma^h)\beta)\frac{(\gamma^{hh}-\gamma^{hm})S[hh] + (\gamma^{hm}-\gamma^{mm})S[hm] + (\gamma^{mm}-\gamma^{hl})S[mm]}{(\gamma^{hm}-\gamma^{mm})S[hm] + (\gamma^{mm}-\gamma^{hl})S[mm]} > 1 \\
&= p(\gamma^h + (1-\gamma^h)\beta)\Big(\frac{(\gamma^{hh}-\gamma^{hm})S[hh]}{(\gamma^{hm}-\gamma^{mm})S[hm] + (\gamma^{mm}-\gamma^{hl})S[mm]} + 1\Big) > 1
\end{aligned}
$$

In this case, I show that the lower bound of the left hand side is greater than 1. Since $S[hm] > S[mm]$, the left hand side is bounded below from $\frac{Pr(h|\gamma^h)}{\alpha}\Big(\frac{(\gamma^{hh}-\gamma^{hm})S[hh]}{\gamma^h(\gamma^{hm}-\gamma^{hl})S[hm]} + 1\Big)$. Also, $\gamma^{hh} = \gamma^h\frac{Pr(h|G)}{Pr(h)}, \gamma^{hl} = \gamma^h\frac{Pr(l|G)}{Pr(l)}$. Thus, the lower bound can be written as

$$
\frac{Pr(h|\gamma^h)}{\alpha}\Big(\frac{\gamma^h(\frac{Pr(h|G)}{Pr(h)} - 1)S[hh]}{\gamma^h(1-\frac{Pr(l|G)}{Pr(l)})S[hm]} + 1\Big) = \frac{Pr(h|\gamma^h)}{\alpha}\Big(\frac{Pr(l)Pr(h|B)}{Pr(h)Pr(l|B)}\frac{S[hh]}{S[hm]} + 1\Big)
$$

Thus, a sufficient condition for $R(p,\gamma^h) > 0$ is that this lower bound is greater than 1. Put it differently,

$$
\frac{S[hh]}{S[hm]} > (\frac{\alpha}{Pr(h|\gamma^h)} - 1)\frac{Pr(h)Pr(l|B)}{Pr(l)Pr(h|B)}
$$

Note $Pr(h|\gamma^h) > pr(h|\gamma)$ and $S[hm] = S[mh]$. Thus, to satisfy both $R(p,\gamma^m) < 0$ and $R(p,\gamma^h) > 0$,

$$
\frac{G^{-1}(1-M(\gamma^{hh}))}{G^{-1}(1-M(\gamma^{hm}))} > (\frac{\alpha}{Pr(h|\gamma)} - 1)\frac{Pr(h)Pr(l|B)}{Pr(l)Pr(h|B)} > \frac{G^{-1}(1-M(\gamma^{hm}))}{G^{-1}(1-M(\gamma^{hl}))}
$$

□

A.3.12 Proof of Proposition 5

Proof. Consider a function $f : R \to R$. Then, $\frac{f(x+b)-f(x-a)}{b+a} \equiv \eta$ is a slope between $x+b$ and $x+a$. Then,

$$\frac{f(x+b)}{f(x-a)} = \frac{\eta(b+a)}{f(x-a)} + 1$$

Similarly, let $\frac{G^{-1}(1-M(\gamma^{hh}))-G^{-1}(1-M(\gamma^{hm}))}{M(\gamma^{hm})-M(\gamma^{hh})} = \eta_h$, $\frac{G^{-1}(1-M(\gamma^{hm}))-G^{-1}(1-M(\gamma^{hl}))}{M(\gamma^{hl})-M(\gamma^{hm})} = \eta_l$. Then,

$$\frac{G^{-1}(1-M(\gamma^{hh}))}{G^{-1}(1-M(\gamma^{hm}))} = \frac{\eta_h(M(\gamma^{hm})-M(\gamma^{hh}))}{G^{-1}(1-M(\gamma^{hm}))} + 1, \quad \frac{G^{-1}(1-M(\gamma^{hm}))}{G^{-1}(1-M(\gamma^{hl}))} = \frac{\eta_l(M(\gamma^{hl})-M(\gamma^{hm}))}{G^{-1}(1-M(\gamma^{hl}))} + 1$$

The condition in Lemma 4 implies that

$$\frac{\eta_h(M(\gamma^{hm})-M(\gamma^{hh}))}{G^{-1}(1-M(\gamma^{hm}))} > \frac{\eta_l(M(\gamma^{hl})-M(\gamma^{hm}))}{G^{-1}(1-M(\gamma^{hl}))}$$

Since $G^{-1}(1-M(\gamma^{hm})) = S[hm] > G^{-1}(1-M(\gamma^{hl})) = S[hl]$, and $(M(\gamma^{hm})-M(\gamma^{hh})) < (M(\gamma^{hl})-M(\gamma^{hm}))$, this implies that $\eta_h > \eta_l$. i.e., the rate of change of threshold firm size increases as rank increases. $\square$

Appendix B

Appendix for Chapter 2

B.1 Discussion on existing literature

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 CEOs' 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 severance 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.

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 asymmtry 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.[1] 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 realization 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).

[1]The 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.

B.2 Proofs

1 Proof of Lemma 10

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.

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

Rearranging yields the result in the lemma. □

2 Proof of Lemma 11

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.

$$\Big[\gamma b + (1-\gamma)(1-b)\Big](\gamma^l_{b+} - \gamma^l)S[l] = \frac{c - (\omega(\gamma) - w^m) + \alpha\omega(\gamma^m) - Pr(h|\gamma)\omega(\gamma^h)]}{FS[l]} \equiv H(\alpha, \gamma) \tag{*}$$

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,

$$f(\alpha_1,\gamma,w_0) = (1-2\gamma)\Big(\big(3\gamma(1-p)+(1-\gamma)(1-\alpha p)\big)w_0 - 2S[l]\frac{\gamma(1-p)(1-\gamma)(1-\alpha p)}{\gamma(1-p)+(1-\gamma)(1-\alpha p)}\Big)$$

$$g(\alpha_1,\gamma,w_0,H) = \frac{1}{\gamma(1-p)+(1-\gamma)(1-\alpha p)}$$

$$\times \Big(H\big(3\gamma(1-p)+(1-\gamma)(1-\alpha p)(\gamma(1-p)+(1-\gamma)(1-\alpha p)\big)$$

$$-(1-\gamma)\big((1-p)(1-\alpha p)\gamma(-3+4\gamma)S[l]$$

$$+(\gamma(1-\alpha)+(1-\gamma)(1-\alpha p)(7\gamma(1-p)+(4-3\gamma)(1-\alpha p))w_0\big)\Big)$$

$$h(\alpha_1,\gamma,w_0,H) = -\frac{1}{\gamma(1-p)+(1-\gamma)(1-\alpha p)}$$

$$\times (1-\alpha p)(1-\gamma)\big(H(\gamma(1-p)+(1-\gamma)(1-\alpha p)-(1-\gamma)(\gamma(1-p)$$

$$+(1-\gamma)(1-\alpha p))w_0-\gamma(1-p)S[l]\big)$$

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. □

3 Proof of Proposition 7

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 $\gamma^{ml}_{B+} = \left(1 + \frac{1-\gamma^m}{\gamma^m}\left(\beta + \frac{1-\alpha}{\alpha}\frac{1-\beta}{1-p}\right)\frac{1-b(\alpha_1,\gamma)}{b(\alpha_1,\gamma)}\right)^{-1}$. For γ^{ml}_{B+} 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 I the effect of m is canceled out as both γ^{ml}_{B+} and γ^m share that history. The above is equivalent to $b(\alpha,\gamma) \geq \frac{\beta + \frac{1-\alpha}{\alpha}\frac{1-\beta}{1-p}}{\left(\beta + \frac{1-\alpha}{\alpha}\frac{1-\beta}{1-p} + 1\right)} = \frac{\frac{Pr(l|\tau=B)}{Pr(l|\tau=G)}}{\frac{Pr(l|\tau=B)}{Pr(l|\tau=G)}+1}$.

□

4 Proof of Lemma 12

Proof. Recall that the demand for insurance occurs only for those $S \in [S[m], S[h])$, 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.

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

That is, if a firm size S ranks above $M(\gamma^{lh})$ (i.e., $S \geq G^{-1}(1 - M(\gamma^{lh}))$, 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. $\square$

5 Proof of Lemma 13

Proof. Observe that the mismatch can happen only for those firms within $[S[ND], S[m])$. 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,

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

This is because $Y_3(S, \gamma) = S \times \Big(E[R] - (1-\gamma)F^{\gamma}\Big)$. 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 $[S[l], S[m])$, 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,

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

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

□

6 Proof of Proposition 8

Proof. The existence of assignment distortion is shown in Lemma 13, and the distortion arises only for $\forall S \in [S[ND], S[m])$, and this is shown in Lemma 12. □

7 Proof of Lemma 14

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.

$$\begin{aligned} &EAS \leq E[\tilde{Y}(S, \gamma^m) - \tilde{\omega}_3(\gamma^m) - \tilde{Y}(S, \gamma^l) + \tilde{\omega}_3(\gamma^l)] \equiv \mathcal{F}(S, \gamma) \\ &EAS \geq \frac{Extra}{Pr(l|\gamma)} \end{aligned}$$

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.

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

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 $S \in [S[m], S[h])$ 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,

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

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

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

For EAS^* to be feasible,

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

Therefore, those sufficiently large firms $S \in \Big[max\Big\{S^{EAS}, S[m]\Big\}, S[h]\Big)$ 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 $E[F_3^i]$.

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

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

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

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

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

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

$$\mathcal{F}(S,\gamma) = -\sigma^2 p\Big(\gamma^m u(\gamma^m)S - \gamma^l u(\gamma^l)S - \Delta\gamma^m u(\gamma^m)S[m]\Big)$$
$$+ S\Big(\gamma^m Pr(h|\gamma)Pr(m|\gamma) - \gamma^l Pr(h|\gamma^l)Pr(m|\gamma^l)\Big)$$
$$- \Delta\gamma^m S[m]Pr(h|\gamma)Pr(m|\gamma)$$

Similarly, observe that,

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

where

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

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,

$$\frac{S\Big[\gamma^m S[m]Pr(h|\gamma)Pr(m|\gamma) - \gamma^l S[l]Pr(h|\gamma^l)Pr(m|\gamma^l)\Big] - \Delta\gamma^m S[m]Pr(h|\gamma)Pr(m|\gamma) - D}{D}$$
$$< \frac{S\Big[\gamma^m u(\gamma^m) - \gamma^l u(\gamma^l)\Big] - \Delta\gamma^m u(\gamma^m) - T}{T}$$

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

8 Proof of Proposition 9

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

$$Y_3(s,\gamma^l) - \omega_3(\gamma^l) > Y_3(s,\gamma^m) - \omega_3(\gamma^m) - EAS$$
$$\leftrightarrow \omega_3(\gamma^m) - \omega_3(\gamma^l) > Y_3(s,\gamma^m) - Y_3(s,\gamma^l) - EAS$$
$$\leftrightarrow (\gamma^m F_3^m - \gamma^l F_3^l)S[m] > (\gamma^m F_3^m - \gamma^l F_3^l)S - EAS$$

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,

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

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 $\gamma^m + \gamma^l < 1$

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

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

$H(\alpha_3, \gamma^m + \gamma^l) > 0$ 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, $\frac{\partial}{\partial \alpha_3}\Big(\alpha_3 p(1-\alpha_3)(\gamma^m - \gamma^l)(\alpha_3 + (1-\alpha_3)(\gamma^m + \gamma^l))\Big)$ starts from positive up until $\alpha_3 = \xi$, then turns to negative onward. i.e. $\Big(\alpha_3 p(1-\alpha_3)(\gamma^m - \gamma^l)(\alpha_3 + (1-\alpha_3)(\gamma^m + \gamma^l))\Big)$ 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 14 and the proof above. $\square$

Appendix C

Appendix for Chapter 3

C.1 Discussion on existing 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 motivating 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 subordintes' 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.

C.2 Proofs

A.1 Proof of Lemma 16

Proof. Since $f'(\cdot) > 0$, the first order condition implies that

$$f'(\theta_1 + m_1)f(\theta_2 + m_2) - f(\theta_1 + m_1)f'(\theta_2 + m_2) < 0 \Rightarrow \theta_1 + m_1 \neq \theta_2 + 1 - m_1$$

Since $f(\cdot)$ is concave increasing, $\theta_1 + m_1 < \theta_2 + 1 - m_1$. Let $Diff = \theta_1 + m_1 - (\theta_2 + 1 - m_1) = 2m_1 + \theta_1 - \theta_2 - 1$. Observe that $\frac{\partial}{\partial d} C'(\cdot) < 0$. Thus, as d increases, $C'(\cdot)$ decreases, implying that the required $Diff$ between the two tasks becomes smaller. i.e., m_1 decreases. Similarly, given the required difference, say ϵ, as θ_2 increases, m_1 must increase to be equal to ϵ. $\square$

A.2 Proof of Proposition 12

Proof. Since $E[X|m] \propto f(\theta_1 + m)f(\theta_2 + 1 - m_1)(X_{SS} - X_S)$. Since $f(\cdot)$ is concave, the probability of $X_{SS} - X_S$ obtains the maximum at m such that $\theta_1 + m = \theta_2 + 1 - m_1$. Due to Lemma 2, as d decreases, the required slack makes m_1^{UB} increase. i.e., $\theta_1 + m_1^{UB} > \theta_2 + 1 - m_1^{UB}$ for sufficiently small d. Then,

$$f(\theta_1 + m_1^{UB}(d))f(\theta_2 + 1 - m_1^{UB}(d))(X_{SS} - X_S) < f(\theta_1 + m_1^B)f(\theta_2 + 1 - m_1^B)(X_{SS} - X_S)$$

Thus, all else equal, as d decreases (thus m_1^{UB} increases), or as $X_{SS} - X_S$ increases, the external hiring performs better than the internal promotion.

$\square$

A.3 Proof of Proposition 13

Proof. Fix $X_{SS} - 2X_S$ so that both external hiring and internal promotion coexist. Then,

$$\begin{aligned} v^{EH} > v^{IP} \Leftrightarrow \frac{C'(m_1^B, 1 - m_1^B)}{C(m_1^B, 1 - m_1^B)} \frac{2}{f'(\theta_1 + m_1^B)} \\ > \frac{C'(m_1^{UB} - d\theta_1, 1 - m_1^{UB})}{C(m_1^{UB} - d\theta_1, 1 - m_1^{UB})} \Big(\frac{1}{f'(\theta_1 + m_1^{UB})} + \frac{1}{f'(\theta_2 + 1 - m_1^{UB})} \Big) \end{aligned}$$

Observe that $\theta_1 + m_1^B + \theta_2 + 1 - m_1^B = \theta_1 + m_1^{UB} + \theta_2 + 1 - m_1^{UB}$. Since $f(\cdot)$ is strictly concave increasing, $f'(\cdot)$ is also concave. Thus,

$$\frac{2}{f'(\theta_1 + m_1^B)} = \frac{1}{f'(\theta_1 + m_1^B)} + \frac{1}{f'(\theta_2 + 1 - m_1^B)} \leq \frac{1}{f'(\theta_1 + m_1^{UB})} + \frac{1}{f'(\theta_2 + 1 - m_1^{UB})} \text{ for } m_1^{UB} \geq m_1^B$$

Equality holds if $m_1^B = m_1^{UB}$. First, consider $\theta_1 = \theta_2 = L$. Then, $m_1^B = m_1^{UB} = 1/2$. This implies that $\frac{2}{f'(\theta_1+m_1^B)} = \frac{1}{f'(\theta_1+m_1^{UB})} + \frac{1}{f'(\theta_2+1-m_1^{UB})}$. Since $C(1/2, 1/2) < c + C(1/2, 1/2)$, this confirms that if $\theta_1 = \theta_2 = L$, $v^{EH} > v^{IP}$.

□

A.4 Proof of Proposition 14

Proof. Due to Proposition 13, there exists d^* such that for $d < d^*$, $v^{EH} < v^{IP}(d)$. Due to continuity, there exists $\underline{d} < d^*$ such that $\Pi(m_1^{UB}(d)) \leq \Pi(m_1^B)$. Thus, for $d \leq \underline{d}$, external hiring dominates internal promotion. Consider $\underline{d}, \overline{d}$ such that $\underline{d} < d^* < \overline{d}$. Due to Lemma 2, $m_1^{UB}(\underline{d}) > m_1^{UB}(\overline{d})$. This implies that $E[X|m_1^{UB}(\underline{d})] < E[X|m_1^{UB}(\overline{d})]$ because as compensation cost is higher for $\underline{d}$ than $\overline{d}$. Therefore $\Pi(m_1^B) \leq \Pi(m_1^{UB}(\overline{d})$, but $v^{EH} > v^{IP}$. □

References

Aghion, P. and Tirole, J. (1994). The management of innovation, *The Quarterly Journal of Economics* pp. 1185–1209.

Anderson, A. and Smith, L. (2010). Dynamic matching and evolving reputations, *The Review of Economic Studies* **77**(1): 3–29.

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.

Demski, J. S. (1998). Performance measure manipulation, *Contemporary Accounting Research* **15**(3): 261–285.

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.

Fama, E. F. (1980). Agency problems and the theory of the firm, *The Journal of Political Economy* pp. 288–307.

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.

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

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. (1989). Agency costs and innovation, *Journal of Economic Behavior & Organization* **12**(3): 305–327.

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.

Holmstrom, B. and Ricart I Costa, J. (1986). Managerial incentives and capital management, *The Quarterly Journal of Economics* pp. 835–860.

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.

Inderst, R. and Klein, M. (2007). Innovation, endogenous overinvestment, and incentive pay, *The RAND Journal of Economics* **38**(4): 881–904.

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* .

Jeon, S. (1998). Reputational concerns and managerial incentives in investment decisions, *European Economic Review* **42**(7): 1203–1219.

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.

Kanodia, C., Bushman, R. and Dickhaut, J. (1989). Escalation errors and the sunk cost effect: An explanation based on reputation and information asymmetries, *Journal of Accounting Research* pp. 59–77.

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.

Lambert, R. A. (1986). Executive effort and selection of risky projects, *The RAND Journal of Economics* pp. 77–88.

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.

Manso, G. (2011). Motivating innovation, *The Journal of Finance* **66**(5): 1823–1860.

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.

Narayanan, M. (1985). Managerial incentives for short-term results, *The Journal of Finance* **40**(5): 1469–1484.

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.

Scharfstein, D. S. and Stein, J. C. (1990). Herd behavior and investment, *The American Economic Review* pp. 465–479.

Serfes, K. (2008). Endogenous matching in a market with heterogeneous principals and agents, *International Journal of Game Theory* **36**(3-4): 587–619.

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.

Strategy& (2016). 2015 CEO success study, *Technical report*, Strategy&.

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.

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

Zwiebel, J. (1995). Corporate conservatism and relative compensation, *Journal of Political Economy* pp. 1–25.