

Information Frictions, Monitoring Costs and the Market for CEOs

by

Zhan Li

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved May 2015 by the
Graduate Supervisory Committee:

Hector Chade, Chair
Natalia Kovrijnykh
Alejandro Manelli

ARIZONA STATE UNIVERSITY

August 2015

ABSTRACT

This paper discusses the matching between CEOs of different talent and firms of different size, by considering boards' costly monitoring of CEOs who have private information about firm output. By incorporating a costly state verification model into a matching model, we have a number of novel findings. First, positive assortative matching (PAM) breaks down as larger firms match with less talented CEOs when monitoring is sufficiently costly despite of complementarity in firms' production technology. More importantly, PAM can be the equilibrium sorting pattern for large firms and high talent CEOs even it fails for small firms and low talent CEOs, which implies that empirical applications relying on PAM are more robust by using samples of large firms. Second, under positive assortative matching, CEO compensation can be decomposed into frictionless competitive market pay and information rent. More talented CEOs extract more rent, which makes their wage even higher. Third, firm-level corporate governance depends on aggregate market characteristics such as the scarcity and allocation of CEO talent. Weak corporate governance can be optimal when CEO talent is sufficiently scarce. My analysis yields a number of empirical predictions on equilibrium sorting pattern, CEO compensation, and corporate governance.

TABLE OF CONTENTS

	Page
CHAPTER	
1 INTRODUCTION	1
2 A MODEL of ALLOCATING CEOS	6
Production Stage	7
Matching Stage	15
A Two by Two Case	15
The General Case	17
3 COMPENSATING AND MONITORING CEOS	20
Compensating CEOs	21
Monitoring CEOs	25
Comparative Statics	28
4 EXTENSIONS	29
Two by Two Case	29
Size-dependent Marginal Cost	30
5 IMPLICATIONS AND EMPIRICAL PREDICTIONS	31
Implications	31
Empirical Predictions	32
6 CONCLUSION	35
REFERENCES	

CHAPTER	Page
APPENDIX	
A PROOFS	41

1 INTRODUCTION

How are CEOs of different talent allocated across firms of different size? Answering this question has proven important in a number of applications such as explaining exponential CEO compensation growth, examining how much CEO talent contributes to shareholder wealth, and motivating corporate governance regulation (Gabaix and Landier, 2008; Tervio, 2008; Dicks, 2012). These applications are based on a common view that complementarity between CEO talent and firm size lead to positive assortative matching (PAM), which means there exists a positive monotonic relationship between firm size and CEO talent in equilibrium. However, CEO-firm relationship is subject to incentive problems. Can these incentive problems affect the allocation of CEOs across firms? Furthermore, can these incentive problems, incorporated into the firm-CEO matching market, have other important implications on how CEOs are compensated and monitored?

Particularly, the strategic interaction between CEOs and boards of directors, which is an important and realistic consideration, is missing in most of the related literature on firm-CEO matching market. It's widely known that CEOs have informational advantage over boards of directors, who play an important role in hiring, monitoring, and compensating CEOs. However, how boards serve their duties will be affected by their information disadvantage, thus their decisions on compensating, monitoring and hiring CEOs will likely affect how CEOs match with firms. If so, positive assortative matching, without a careful consideration, might be an artifact from oversimplified models which neglect important and realistic economic forces, and empirical applications relying on PAM can be spurious.

Our paper bridges the gap by integrating information friction and board monitoring in a matching market with CEOs of different talent and firms of different size.

There is a board in each firm hiring, monitoring, and compensating the CEO. There are two stages in our model. In the first stage, the board in each firm simultaneously decides which CEO to hire. Firms' size and CEOs' talent are observable. In the second stage, output will be realized as a product of firm size and CEO talent with some probability and zero otherwise. The realization of firm output is only known to the CEO who reports it to the board and the board audits the CEO's report at a cost. The board maximizes her expected utility by setting auditing probability and CEO compensation to incentivize the CEO for truthful reporting. The board's expected utility is tied to firm profit after paying CEO compensation and the board suffers disutility from monitoring the CEO due to lack of independence. Corporate governance in each firm is defined as the probability of the board's auditing the CEO's report.

In sharp contrast to previous literature on how CEOs match with firms, we show that despite of the complementarity between firm size and CEO talent, positive assortative matching can break down when marginal cost of monitoring is high enough. The intuition for the failure of PAM is as follows. Every board prefers a more talented CEO. The board in a larger firm has both advantage and disadvantage to compete against the board in a smaller firm for valuable CEO talent. A larger firm generates higher marginal product from the CEO's talent due to complementary production technology, thus it tends to outbid the smaller firm for the high talent CEO. However, the board in a larger firm suffers from a higher disutility from the need to monitor the CEO more intensively because there are more resources at a larger firm for the CEO to appropriate. Monitoring is costly, thus the board at a larger firm underbids because of a larger monitoring cost. The disadvantage increases with marginal cost of monitoring. When marginal cost of monitoring is high enough, the

disadvantage dominates advantage, positive assortative matching will fail and a more talented CEO match with a smaller firm.

The above result is important from at least two aspects. First, it provides us a robustness check on the validity of using PAM as the equilibrium sorting pattern, which is widely used in empirical applications. When modeling the CEO labor market, it's not suitable to take as granted that there is PAM between firm size and CEO talent.¹ Second, our result explains why CEO turnover can be affected by industry shock, which is a seemingly puzzling empirical finding contradicting with a common belief that market-level shock should be filtered out in the decision of retaining and firing CEOs. This is explained by our finding that, besides marginal cost of monitoring, sorting pattern can also be affected by market-level profitability. And a change in sorting pattern can be considered as CEO turnover and a change in market-level profitability can be considered as industry shock, our result thus explains why CEO turnover is connected to industry shock.

More importantly, we show that as PAM fails, it will fail first for smaller firms and lower talent CEOs. It can be true that there is PAM in top firms and CEOs, but PAM fails for bottom firms and CEOs. This implies that empirical applications relying on PAM are more robust with a subsample of large firms compared to a subsample of small firms. Thus a full sample will not necessarily be more robust than the sample excluding smaller firms.

We have two novel findings on CEO compensation. First, CEO compensation is always increasing in CEO talent irrespective the size of firms that CEOs match with. Thus our result shows that the fundamental determinant of CEO compensation is CEO talent instead of firm size they match with. Empirical results on the significant

¹This is understandable because firms are heterogeneous in many dimensions, then heterogeneity other than size might distort PAM between size and talent. However, our paper shows that even firms and CEOs are only different in a single dimension (i.e., size and talent), PAM can still fail.

positive correlation between firm size and CEO compensation can be reconciled if there is significant positive correlation between firm size and CEO talent. Second, under positive assortative matching, CEO compensation can be decomposed into two components. The first component is explained by competitive market pay, and the second component is explained by rent extraction because CEOs extract information rent from their informational advantage over boards of directors. The information rent increases with CEO talent, thus, in our model, wage spread between different CEOs is larger than wage spread obtained from models without information frictions. When marginal cost of monitoring is lower, CEOs' information rent will be lower which results in lower CEO compensation as well as stronger corporate governance.

We discuss firm-level corporate governance by assuming a specific sorting pattern between firm size and CEO talent. In contrast to most related literature, our discussion applies to any sorting pattern besides positive assortative matching. The main result is that firm-level corporate governance is affected by aggregate market characteristics such as the scarcity of CEO talent and how CEOs match with firms. Cross-sectionally, larger firms have weaker governance than smaller firms when high-talent CEOs are scarce enough or larger firms match with sufficiently talented CEOs. The intuition is as follows. When talented CEOs are sufficiently scarce or larger firms try to match with sufficiently talented CEOs, competition between firms will be more intensive. Thus it's not enough for larger firms to merely use their size advantage to attract more talented CEOs even though larger size gives a CEO higher marginal product. A firm with larger size has to set a lower corporate governance to attract valuable CEO talent. The result on corporate governance shows that weak corporate governance can arise from the competition for CEO talent instead of board's failure in monitoring the CEO.

By connecting marginal cost of monitoring to board characteristics such as boards' independence, ownership, and reputational concerns, our model suggests a number of empirical predictions. First, pay-performance sensitivity, CEO pay and corporate governance are substitutes, thus higher corporate governance will result in a lower pay-performance sensitivity and CEO pay. Second, PAM between firm size and CEO talent is more likely to fail (1.) for an industry with lower boards' independence, boards' ownership or reputational concerns (2.) in economic downturn, (3.) for smaller firms and lower talent CEOs. Third, (1.) monotonicity between CEO compensation and CEO talent is stronger than that between CEO compensation and firm size, (2.) under PAM, wage spread increases with firm size, and CEO compensation is higher for an industry with lower boards' independence, boards' ownership or reputational concerns. Forth, under PAM, (1.) when there is a scarcity of CEO talent (large firms), larger firms have weaker (stronger) corporate governance than smaller firms, (2.) corporate governance is weaker at economic downturns.

To our best knowledge, Edmans and Gabaix (2011) is the only paper discussing the distortion of talent allocation and its implication on economic efficiency. Firms in Edmans and Gabaix (2011) differ in multidimensional types (size, riskiness and CEO's cost of effort), which are cleverly transformed into a single type of "effective firm size". They find PAM between firm size and CEO talent can fail after taking moral hazard into consideration and there is efficiency loss associated with such failure. However, in their model, PAM is always the equilibrium sorting pattern between effective firm size and CEO talent. Compared with our paper where the failure of PAM is driven by the absolute level of model parameters, sorting pattern in Edmans and Gabaix (2011) is driven by their magnitude compared with other firms instead of the absolute level. Acharya et al. (2012) examine how corporate governance and competition will affect firms' ability to attract better CEOs. They find firms with weaker corporate

governance will match with better CEOs both theoretically and empirically. In their model, firms have the same size ex ante and the role of board of directors is not considered.

Our paper proceeds as follows. Section 2 solves the model and presents equilibrium sorting condition. Section 3 discusses CEO compensation and monitoring of CEOs. Section 4 considers extensions when firms have different marginal cost of monitoring. Section 5 discusses implications and empirical predictions. Section 6 concludes.

2 A MODEL of ALLOCATING CEOS

There are a measure 1 of firms and CEOs in the market and each firm has a board of directors. Denote the distribution functions of firm size as F and CEO talent as G . A firm with size s is indexed by $i = F(s)$. Denote s_i as the size of the firm which has index i and naturally $i = F(s_i)$. Similarly, we can define a_j as the CEO's talent who has index j among all CEOs. We may refer the firm with index i as firm i , the board at firm i as board i , and the CEO with index j as CEO j . We will not discuss the incentive problems between firms and boards, thus board can be considered as firm in our model, and we will use board and firm interchangeably. We assume both distribution functions are atomless, thus s_i and a_j are both strictly increasing in indexes $i \in [0, 1]$ and $j \in [0, 1]$ respectively, thus we will call indexes i, j as ranks. We also assume s_i and a_j are continuously differentiable.

There are two stages in our model. In the first stage, boards hire CEOs simultaneously from the CEO labor market. In the second stage, each firm's output will be realized, which will be divided between the board and the CEO within the firm, thus here the board fully aligns her interest with shareholders. We call the first stage as the matching stage and the second stage as the production stage. Formally, a matching is a one to one mapping $M : [0, 1] \rightarrow [0, 1]$, where firm i matches with the CEO of

rank $M(i)$ and CEO j matches with the firm of rank $M^{-1}(j)$. When M mapped i to $j = i$, the same ranked firms and CEOs match with each other, we call this sorting pattern as positive assortative matching (PAM), which will play an important role in our discussion.

To define equilibrium, we first define stable matching. Intuitively, if no board and CEO will be strictly better off by breaking current match and forming a new pair or prefer to stay unmatched, then the matching is stable. Note here the notion of stability requires both sides to have no incentives to deviate. Clearly, whether the matching is stable is related to boards' utility and CEOs' wage from such matching. We assume boards and CEOs receive zero payoff if unmatched. In equilibrium, board's utility only depends on firm rank and CEO wage only depends on CEO rank, given distributions of firm size and CEO talent. Denote board utility function as $u : [0, 1] \rightarrow R_+$ and CEO wage as $v : [0, 1] \rightarrow R_+$. An equilibrium consists of a matching $M : [0, 1] \rightarrow [0, 1]$ and payoffs $u : [0, 1] \rightarrow R_+$ and $v : [0, 1] \rightarrow R_+$ which satisfy:

- (i) $\forall j \in [0, 1], i = M(j), (u_i, v_j)$ is feasible
- (ii) the matching is stable

We will start with the production stage and solve the utility frontier of a matched pair of board and CEO by solving a standard costly state verification problem. We will use the utility frontier to solve the equilibrium sorting pattern in the first stage and show how the change in sorting pattern is driven by the the optimal contract in the production stage.

2.1 Production Stage

We describe the production stage in which a board and a CEO engage in upon matching. Technically, we solve a costly state verification model to characterize the

optimal contract and Pareto frontier of a matched pair, which will be used to solve the first stage equilibrium sorting pattern.

The firm's size is s and the CEO's talent is a , both s and a are observable. There are two states of the world, denoted as $\{H, L\}$. With probability p , state H realizes and the output is sa ; with probability $1-p$, state L realizes and the output is 0. Thus p measures the profitability of firm production. The expected output is thus psa , which is complementary in firm size and CEO talent by noting that the marginal product of CEO talent is increasing in firm size. The complementarity in firm's technology is a natural assumption, considering that in a larger firm, the impact of CEO talent will be larger and thus marginal product of CEO talent will be higher.

The true output is only observable to the CEO, though the distribution is known to both the board and the CEO. Upon observing the output, the CEO submits a report $r \in \{H, L\}$ to the board who then decides a probability g_r to audit the CEO's report, which is called the probability of monitoring. We assume once the board decides to audit the report, she will know the output for sure by paying total monitoring cost kg_r , where $k \geq 0$ is the marginal cost of monitoring.

Throughout the paper, we assume $k < \frac{ps_0a_0}{1-p}$, where s_0 and a_0 are size and talent for the firm and CEO with rank 0 (i.e., the smallest firm and the CEO with the lowest talent). We will show that this assumption is sufficient to guarantee all boards' participation constraint can be satisfied. We also assume all boards in different firms have the same k . There are two reasons for this assumption. The first reason is, in this paper, we focus on how magnitude of k instead of heterogeneity of k would affect equilibrium sorting and equilibrium payoffs. Adding heterogeneity in k will add another factor to affect sorting pattern which is not the focus of this paper and complicates economic intuition. The second reason is purely technical. Firms will have two-dimensional types (s, k) if boards have different k s and firms have

different size. Since our model cannot be simplified by combining (s, k) into a single dimensional type (e.g., Edmans and Gabaix, 2011), we will run into the complicated problem of solving multidimensional matching, on which little is known.

Monitoring the CEO can be costly to the board for various reasons. The cost can be considered as the disutility from effort aversion because the board needs to put effort into auditing the report. In the case of hiring an external auditor to audit the report, the cost can be considered as auditing fee paid to the auditor. The cost can also be considered as board's distaste of monitoring the CEO because of the board's lack of independence. As suggested by Hermalin and Weisbach (1998), a board's careers are tied to the CEO and the board suffers from disutility when directors oppose the CEO. In this paper, we will focus on the third explanation and consider how the marginal cost of monitoring is affected by various board characteristics including board independence.

The CEO is paid with a wage of w_r when there is no auditing and w_{or} when there is auditing, where $o \in \{H, L\}$ refers to true state of the world. When there is no auditing, wage will be paid according to the CEO's report r and when there is auditing, true output will be known to the board, thus wage will be paid according to the CEO's report r and true output o . When $o = r$, the CEO reports truthfully and when $o \neq r$, the CEO falsifies a report. According to revelation principle, we will look for truth-telling equilibrium in which the CEO reports truthfully. Apparently, the contract consists of $\{w_r, w_{or}, g_r\}, r \in \{H, L\}$. From now on, denote

$$C = \{w_L, w_{LL}, w_{LH}, w_H, w_{HH}, w_{HL}, g_L, g_H\}$$

as the optimal contract.

In truth-telling equilibrium, the CEO's incentive compatibility has to be satisfied. In state H , if the CEO falsifies a report of L , she gets $w_L + sa - 0$ with probability $1 - g_L$ and w_{HL} with probability g_L ; instead, if the CEO reports truthfully, she gets w_H with probability $1 - g_H$ and w_{HH} with probability g_H . When the CEO falsifies a report, she pockets the difference between actual output and reported output. The incentive compatibility under state H requires the CEO prefers reporting H to falsifying a report of L . By the same token, incentive compatibility requires that the CEO reports L other than H under state L . The CEO's incentive compatibility conditions are:

$$(1 - g_H)w_H + g_H w_{HH} \geq (1 - g_L)(w_L + sa - 0) + g_L w_{HL}$$

$$(1 - g_L)w_L + g_L w_{LL} \geq (1 - g_H)(w_H + 0 - sa) + g_H w_{LH}$$

Under truth-telling equilibrium, denote expected CEO compensation as EW , expected total monitoring cost as EC , and expected firm output as EO , we have

$$EW = p((1 - g_H)w_H + g_H w_{HH}) + (1 - p)((1 - g_L)w_L + g_L w_{LL})$$

$$EC = pkg_H + (1 - p)kg_L$$

$$EO = psa$$

The board solves the following problem

$$\begin{aligned}
u(s, a, v) &= \underset{C}{Max} EO - EW - EC \\
s.t. & p((1 - g_H)w_H + g_H w_{HH}) + (1 - p)((1 - g_L)w_L + g_L w_{LL}) \geq v, (PK) \\
& (1 - g_H)w_H + g_H w_{HH} \geq (1 - g_L)(w_L + sa - 0) + g_L w_{HL}, (ICH) \\
& (1 - g_L)w_L + g_L w_{LL} \geq (1 - g_H)(w_H + 0 - sa) + g_H w_{LH}, (ICL) \\
& w_L, w_{LL}, w_{LH}, w_H, w_{HH}, w_{HL} \geq 0, (LL) \\
& w_H, w_{HH} \leq sa, w_L, w_{LL} \leq 0, (FC) \\
& g_H, g_L \in [0, 1]
\end{aligned} \tag{1}$$

$u(s, a, v)$ is the board's value function. PK is promise keeping constraint, and v is the value promised to the CEO by the board in the first stage when the board matches with the CEO. The CEO is protected by limited liability (LL) which means the CEO cannot be paid with a strictly negative wage. FC is feasibility constraint, which means the CEO cannot be paid with a wage higher than the output.

From LL and FC, $w_L = w_{LL} = 0$. We can also have $w_{HL} = w_{LH} = 0$ as in standard costly verification model (Border and Sobel, 1987). This is because the board wants to maximize penalty to reduce total monitoring cost, however, the board cannot penalize the CEO with a negative wage because of the limited liability constraint, thus $w_{HL} = w_{LH} = 0$. By LL, the left hand side of ICL is non-negative and by FC, the right hand side of ICL is smaller or equal to zero, thus ICL can always be satisfied from LL and FC, which implies the CEO will not falsify a report if the true state is L , thus a report of high output must be true. Because monitoring is costly, the board will not audit the CEO's report if she reports H, thus $g_H = 0$. By substituting

$w_L = w_{LL} = 0$, $w_{HL} = w_{LH} = 0$, and $g_H = 0$, the board's problem can now be simplified as:

$$\begin{aligned}
u(s, a, v) &= \underset{g_L \in [0,1], w_H}{Max} \quad psa - pw_H - (1-p)kg_L \\
& \quad s.t. \quad pw_H \geq v, (PK) \\
& \quad \quad \quad w_H \geq (1-g_L)sa, (ICH)
\end{aligned}$$

ICH must be binding, if not, the board can reduce g_L and increase her expected utility, thus $g_L = 1 - \frac{w_H}{sa}$. By substituting $g_L = 1 - \frac{w_H}{sa}$ into the above problem, we have

$$\begin{aligned}
u(s, a, v) &= \underset{g_L \in [0,1]}{Max} \quad pg_Lsa - (1-p)kg_L \\
& \quad s.t. \quad p(1-g_L)sa \geq v, (PK)
\end{aligned}$$

Note that we have assumed $k < \frac{ps_0a_0}{1-p}$, because $s \geq s_0$ and $a \geq a_0$ for all s and a , then we have $k < \frac{psa}{1-p}, \forall s, a$, which implies the board's expected utility is increasing in g_L . The board will want to set g_L as large as possible, thus the promise keeping constraint must be binding. We will also show that in equilibrium $v \leq psa$. The optimal solution is

$$g_L = 1 - \frac{v}{psa}, w_H = \frac{v}{p}$$

which shows that expected CEO wage pw_H equals to the promised value v and the probability of monitoring $g_L = \frac{psa-v}{psa}$ equals to the share of total output obtained by

the board of directors excluding CEO compensation. Thus monitoring determines the division of match surplus between the board and the CEO.

Proposition 1. The optimal contract consists of $g_H = 0, g_L = 1 - \frac{v}{psa}, w_H = \frac{v}{p}, w_L = w_{LL} = w_{LH} = w_{HL} = 0$

From the optimal contract, g_L is increasing in firm size s , which implies the probability of monitoring is larger at a larger firm, and thus total monitoring cost is higher at a larger firm. Intuitively, at a larger firm, there are more opportunities for the CEO to appropriate, thus the board needs to impose higher probability of monitoring to prevent the CEO from misreporting. g_L is a decreasing function of v , and in equilibrium, the CEO's expected wage will exactly equal to v , thus a higher CEO compensation implies a lower probability of monitoring. Intuitively, a higher compensation and a stronger corporate governance can both incentivize the CEO for truthful reporting, thus monitoring and compensation can substitute for each other. However, both are costly to the board, the board will impose a lower probability of monitoring when paying the CEO a higher compensation to save monitoring cost.

Another noticeable feature of the optimal contract is the probability of monitoring is independent of k . Mathematically, this is because optimal contract can be pinned down exactly by the CEO's incentive constraints, which are independent of how costly monitoring is. Intuitively, when monitoring becomes less costly, the board will not increase the probability of monitoring because doing so is not in the interest of the board as long as monitoring is costly. Similarly, the board will not decrease the probability of monitoring when monitoring becomes more costly, this is because the board always prefers to impose a higher probability of monitoring instead of a higher CEO compensation. However, the independence of the probability of monitoring with respect to the marginal cost of monitoring k is only true in the production stage

when we treat v as an exogenously given. In equilibrium, v is a function of k , and the probability of monitoring will be affected by k through CEO compensation.

By plugging the optimal contract to the board's objective function, the board's value function is

$$u(s, a, v) = \underbrace{psa - v}_{\text{first best}} - \underbrace{(1-p)k\left(1 - \frac{v}{psa}\right)}_{\text{total monitoring cost}} \quad (2)$$

which contains two parts: first best part and total monitoring cost. The first best part corresponds to a CEO market without monitoring cost and information friction, i.e., both the board and the CEO know the firm's output and auditing the CEO's report isn't needed. It's easy to show that under the first best case, there will always be positive assortative matching (PAM). However, with information friction, monitoring cost might generate a force to distort positive assortative matching between firm size and CEO talent. Total monitoring cost is increasing in firm size, which implies a larger firm suffers from a larger loss from matching because of a higher probability of monitoring in a larger firm.

Technically, in the first best case, we have a transferable utility matching, which means the exchange rate of payoffs between the board and the CEO is 1:1, and Becker (1973) shows that when utility is transferable, complementarity in production technology is sufficient for positive assortative matching. In our model, the board's utility function can be written as $u(s, a, v) = psa - (1-p)k - \left[1 - \frac{(1-p)k}{psa}\right]v$, thus the exchange rate between u and v is $1 - \frac{(1-p)k}{psa} : 1$, which corresponds to an imperfectly transferable utility (ITU) matching.² For ITU matching, Legros and Newman (2007)

²It's sometimes called non-transferable utility (NTU) by some authors, however, NTU is also used to refer to matching when transfer is completely prohibited. We think imperfectly transferable utility matching is a more suitable name.

shows that complementarity is not sufficient to generate PAM, and thus we will move to the matching stage to examine how firms match with CEOs with the presence of monitoring cost.

2.2 Matching Stage

After solving the production stage optimal contract and the board's value function, we solve the first stage matching problem. In the first stage, firms match with CEOs by anticipating the second stage payoffs. We start from a simple two by two case, which serves the purpose of conveying intuition on the driving forces of how firms match with CEOs.

2.2.1 A Two by Two Case

We first consider a simple case with two firms and two CEOs. We denote two firms' size as s_0 and s_1 with $s_0 < s_1$ and CEOs' talent as a_0 and a_1 with $a_0 < a_1$. Clearly, there are only two sorting patterns: "firm 0 matches with CEO 0, firm 1 matches with CEO 1" and "firm 0 matches with CEO 1, firm 1 matches with CEO 0". We refer to the first one as PAM and the second one as negative assortative matching (NAM). Denote the utility for board $i \in \{0, 1\}$ to match with CEO $j \in \{0, 1\}$ as

$$u(s_i, a_j, v_j) = ps_i a_j - (1 - p)k - \left[1 - \frac{(1 - p)k}{ps_i a_j}\right]v_j \quad (3)$$

If the equilibrium sorting pattern is PAM, then the board in firm 1 prefers to match with CEO 1 to CEO 0 and the board in firm 0 prefers to match with CEO 0

to CEO 1, and we have the following inequalities:

$$u(s_1, a_1, v_1) \geq u(s_1, a_0, v_0) \quad (4)$$

$$u(s_0, a_0, v_0) \geq u(s_0, a_1, v_1)$$

It's easy to understand that above inequalities are also sufficient conditions for PAM. By the functional form of $u(s_i, a_j, v_j)$, which is defined in Equation (3), the above two inequalities can hold if and only if $\frac{ps_0(a_1-a_0)}{1-\frac{(1-p)k}{ps_0a_1}} \leq v_1 \leq \frac{ps_1(a_1-a_0)}{1-\frac{(1-p)k}{ps_1a_1}}$, which implies

$$k \leq \frac{p}{1-p} \frac{1}{\frac{1}{s_0a_1} + \frac{1}{s_1a_1}} \quad (5)$$

which is the sufficient and necessary condition for PAM to hold in this two by two case.

The two by two case shows that in the CEO market with information fiction, complementarity in firm production is not sufficient to guarantee positive assortative matching. There is an equilibrium that the large firm matches with the low talent CEO when monitoring CEO becomes more costly. The intuition relies on the understanding that size brings both advantage and disadvantage to the firm. The advantage of the large firm is, it has higher marginal product from CEO talent because of the complementary production technology. The large firm tends to outbid the small firm due to higher surplus from matching with the high talent CEO. The disadvantage of the large firm is: higher probability of monitoring is needed in the large firm because more stakes on the table for the CEO to appropriate, then the large firm tends to underbid the smaller firm because higher monitoring cost reduces matching surplus. As the marginal cost of monitoring k increases, advantage will not be affected (since it is a result of production technology) and disadvantage increases.

And when k is large enough, disadvantage outweighs advantage, large firm underbids small firm, NAM will emerge as the equilibrium sorting pattern.

To be more precise, we can consider a situation where both firms bid for the high talent CEO. The highest bid is the wage paid to the CEO when the firm is indifferent from matching with both CEOs. For firm $i \in \{0, 1\}$, denote the maximum bid as b_i , then b_i can be solved from $u(s_i, a_1, b_i) \geq u(s_i, a_0, 0)$. We have $b_i = \frac{ps_i(a_1 - a_0)}{1 - \frac{(1-p)^k}{ps_i a_1}}$, $i \in \{0, 1\}$. Then there is PAM if and only if $b_1 \geq b_0$. Note that the numerator of b_i measures how much firm output increases when firm i matches with high talent CEO instead of the low talent one and it's increasing in firm size. However, the denominator is also increasing in firm size. Thus size brings both advantage and disadvantage to firm i . As k increases, denominator decreases, both firms will bid more aggressively for the high talent CEO. This is because it's more costly to monitor the CEO and boards use higher CEO compensation to reduce monitoring cost. However, the large firm will fail the competition eventually if k continues to grow.

2.2.2 The General Case

The above result from two by two case is easy to be generalized to the market with a continuum of firms and CEOs. The difference is that in a market with more than three firms and CEOs, we have more than two sorting patterns, thus failure of PAM will not imply NAM.³ Instead of a cutoff value for PAM and NAM as in the two by two case, there exists a cutoff value for PAM and the failure of PAM.

Proposition 2. There is PAM between firm size and CEO talent if and only if $k \leq$

$$\frac{ps_0 a_0}{2(1-p)}$$

³It's easy to see that in a market with n firms and n CEOs, there are a total number of $n!$ sorting patterns. When we have a continuum of firms and CEOs, $n \rightarrow \infty$ and $n! \rightarrow \infty$, thus we have infinite number of sorting patterns.

The above proposition shows that in the market with a continuum of firms and CEOs, when marginal cost of monitoring k is higher than the cutoff value $\frac{ps_0a_0}{2(1-p)}$, PAM fails. The intuition carries from the two by two case. When $k = \frac{ps_0a_0}{2(1-p)}$, there can be other sorting patterns which are not PAM but are pay-off equivalent to PAM (Legros and Newman, 2007), which means these sorting patterns generate the same payoffs for boards and CEOs as PAM. The cutoff value $\frac{ps_0a_0}{2(1-p)}$ is distribution free and more specifically, it is not related to size of firms and talent of CEOs other than the bottom firm and CEO.

Note that PAM requires a perfect positive monotonic relationship between size and talent for all firms and CEOs, thus it is a fairly strong requirement. However, we might be able to obtain a finer structure when we consider subsets of firms and CEOs: it's possible that when PAM fails, it is still the equilibrium sorting pattern for a subset of firms and CEOs when it fails for another subset of firms and CEOs. To proceed, we first define sorting patterns on a subset of firms and CEOs. If PAM is the equilibrium sorting pattern for firms and CEOs with ranks $[m, n] \subseteq [0, 1]$, then we say there is PAM on $[m, n]$; if PAM fails for firms and CEOs with ranks $[m, n] \subseteq [0, 1]$, then we say PAM fails on $[m, n]$. One thing to notice here is that even if PAM fails on $[m, n]$, there might still exist a subset $[\alpha, \beta] \subset [m, n]$ such that there is PAM on the proper subset $[\alpha, \beta]$ and PAM fails on a non-empty complement subset $[m, n] \setminus [\alpha, \beta]$. In the proposition below, we show that when PAM fails, for some k , we can find a cutoff rank such that there is PAM for firms and CEOs above that cutoff rank and there is a failure of PAM for firms and CEOs below that cutoff rank.

Proposition 3. (PAM on top) For $k \in (\frac{ps_0a_0}{2(1-p)}, \min\{\frac{ps_0a_0}{1-p}, \frac{ps_0a_1}{2(1-p)}\})$ and $\bar{i} = \{i : k = \frac{ps_0a_i}{2(1-p)}\}$, there is PAM on $[\bar{i}, 1]$ and PAM fails on $[0, \bar{i}]$

The above proposition shows that for given distributions of firms and CEOs, when k is larger than the cutoff value $\frac{ps_0a_0}{2(1-p)}$, PAM will first fail for smaller firms and lower

talent CEOs, but PAM is still the equilibrium sorting pattern for larger firms and more talented CEOs above the cutoff value. However, as k increases, the region for PAM will shrink as the cutoff value \bar{i} is an increasing function of k . This result provides a robustness condition for empirical applications relying on PAM, which will be discussed in latter sections.

We emphasize here that the above proposition is a sufficient condition for “PAM on top”. It’s still possible that we can find a subset of firms and CEOs on $[0, \bar{i}]$ such that there is PAM on that subset. In fact, if the matching function is continuous, because PAM is the equilibrium sorting pattern on $[\bar{i}, 1]$, we can find a $\epsilon > 0$ and construct such subset as $[\bar{i} - \epsilon, \bar{i}]$ and there is PAM on $[\bar{i} - \epsilon, \bar{i}]$ due to continuity. What we show here is that when PAM fails, it must not fail on $[\bar{i}, 1]$, and must fail on (a subset of) $[0, \bar{i}]$. Note that \bar{i} is unique because CEO talent is strictly increasing in rank. $k < \min\{\frac{ps_0a_0}{1-p}, \frac{ps_0a_1}{2(1-p)}\}$ ensures our assumption of $k < \frac{ps_0a_0}{1-p}$ can be satisfied and $k < \frac{ps_0a_1}{2(1-p)}$, thus the solution to $\bar{i} = \{i : k = \frac{ps_0a_i}{2(1-p)}\}$ is not empty.

Note that both cutoff values $\frac{ps_0a_0}{2(1-p)}$ and \bar{i} are decreasing in p , thus Propositions 2 and 3 show how sorting pattern is affected by p , which is the probability of “high output”, a measure of profitability of firm production. Mathematically, the effect from lowering p on sorting pattern is similar to the effect from increasing k . Intuitively, when p is small, complementarity is weaker, thus PAM is more likely to fail. To be more specific, if we fix k and decreases p , PAM will fail for bottom firms and CEOs first. As p continues to decrease, PAM will fail for more and more bottom firms and CEOs, and the set of firms and CEOs on which PAM is the equilibrium sorting pattern will become smaller. We will discuss how this explains why CEO turnover is connected to industry shock.

3 COMPENSATING AND MONITORING CEOS

In previous section, we have shown that considering information frictions in CEO labor market will generate surprising results on the allocation of CEOs: PAM may not be the equilibrium sorting pattern even with complementary production technology and unidimensional matching (firms are only different in size and CEOs are only different in talent). And in this section, we will show such consideration also generates novel insights on how CEOs are compensated and monitored. Unlike previous literature focusing nearly exclusively on PAM, we will also examine how failure of PAM affects CEO compensation and monitoring of CEOs.

We start with two lemmas. The first lemma shows that once PAM fails, there is negative assortative matching (NAM) on a subset of firms and CEOs. Note here we assume the matching function M is continuously differentiable. The second lemma presents closed form solutions of CEO compensation and the probability of monitoring.

Lemma 1. When PAM fails on $[0, 1]$, there exists a subset of firms $(a, b) \subseteq [0, 1]$ with $M'(i) < 0, i \in (a, b)$.

The proof to the above lemma is very straightforward. Intuitively, if there doesn't exist such subset, PAM must be the equilibrium sorting pattern on $[0, 1]$, which contradicts with the failure of PAM.

In the next lemma, we solve CEO compensation and the probability of monitoring. Our solution is different from previous literature (Gabaix and Landier, 2008; Edmans et al., 2009; Tervio, 2008; Dicks, 2012) from two aspects: first, our closed-form solution can be applied to any sorting pattern, not only PAM; second, our solution does not rely on specific distributions of CEO talent and firm size, this challenges us with

some tractability issue, however, it also provides us the opportunity to examine how changes in distributions of talent and size can affect how CEOs are monitored.

Lemma 2. Given sorting pattern M , CEO j 's wage and the probability of monitoring at firm i are

$$v_j = \int_0^j \varphi(l, j, k, M) p s_{M(l)} a'_l dl$$

$$g_i = 1 - \frac{v_{M(i)}}{p s_i a_{M(i)}}$$

where $\varphi(l, j, k, M) = \exp[-\int_l^j \frac{\frac{(1-p)k}{p s_{M(x)} a_x'}{1 - \frac{(1-p)k}{p s_{M(x)} a_x}} dx] / (1 - \frac{(1-p)k}{p s_{M(l)} a_l})$

Note that in Section 2.1, the probability of monitoring is independent of marginal cost of monitoring k . However, CEO compensation is endogenously determined in the matching market and it is a function of k . From the optimal contract in single firm's problem, we know that how intensively to monitor the CEO is affected by how much to compensate the CEO, thus CEO compensation is affected by the marginal cost k through CEO compensation, which is clearly shown in $g_i = 1 - \frac{v_{M(i)}}{p s_i a_{M(i)}}$. The above Lemma also highlights the dependence of CEO compensation and the probability of monitoring on sorting pattern M . The next two sections will show such dependence in a clearer way.

3.1 Compensating CEOs

A central debate in CEO compensation is whether CEOs are paid higher merely because they work for larger firms. Previous literature shows that under PAM, higher talent CEOs match with larger firms, and firm size augments CEO pay, which pushes CEO compensation higher when firm size increases. However, it's not clear how

crucially this result depends on sorting pattern. The following proposition shows that CEO compensation is increasing in CEO talent in any sorting pattern, and will not increase in firm size if PAM fails.

Proposition 4. (Reward for talent) Regardless of sorting pattern

- i. CEO compensation is strictly increasing in CEO talent
- ii. CEOs of the same talent matching with firms of different size have the same wage

The result above implies that CEO compensation will not increase in firm size if PAM fails and a CEO will not be paid with a higher wage solely because she works for a larger firm. The result here seemingly contradicts with the empirical finding of positive correlation between CEO compensation and firm size. Without actually testing the sorting pattern of CEO market from available data, this can be reconciled from two aspects. The first is, PAM is indeed the equilibrium sorting pattern in the CEO market, thus CEO compensation is an increasing function of firm size because PAM between firm size and CEO talent implies PAM between firm size and CEO compensation. The second is, PAM actually fails, however, there is still positive correlation between CEO compensation and firm size if the failure of PAM is not too extreme (such as NAM). This is because correlation measures linear relationship, but the above proposition discusses monotonic relationship, and either will imply the other. Thus positive correlation between CEO compensation and firm size will not contradict with the fact that CEO compensation does not increase in firm size.

Overall, Proposition 4 shows that talent will always be rewarded even in a frictional CEO labor market where PAM may fail. In fact, this result is true as long as boards prefer a high talent CEO to a low talent CEO. To understand this, we assume there are two CEOs, CEO 0 and CEO 1. If both CEOs have the same talent, then they must be paid with the same wage, if not, both boards will prefer the one with a

lower wage, thus the matching cannot be stable. If CEO 1 has strictly higher talent than CEO 0 but is paid with a weakly lower wage than CEO 0, then the board who hires CEO 0 will be better off by hiring CEO 1 and paying her CEO 0's wage. CEO 1 will accept her new job offer because now she is paid with a higher wage. Thus matching is unstable and a rematch occurs until CEOs with strictly higher talent are paid with a strictly higher wage.

Previous literature focuses on frictionless CEO labor market without any information friction. Denote \bar{v}_j as CEO j 's compensation under frictionless CEO market. In such market, complementary production technology will be sufficient for PAM, and thus it's easy to solve CEO j 's wage $\bar{v}_j = \int_0^j ps_l a'_l dl$.⁴ Here \bar{v}_j crucially depends on firm size, however, we will show that CEO compensation will also be affected by information friction besides firm size in the frictional CEO market with information frictions and monitoring cost. Instead of discussing CEO compensation under different sorting patterns in frictional CEO market, we restrict our discussion to PAM. This provides us an opportunity to examine how information friction affects CEO compensation by holding sorting pattern fixed (PAM). By Lemma 2, we can solve CEO j 's wage under PAM with information friction v_j^P by substituting $j = M(j)$.

Proposition 5. Denote v_j^P and \bar{v}_j as CEO j 's compensation under frictional and frictionless CEO markets respectively, under PAM, we have

$$v_j^P = \bar{v}_j + \delta_j, \frac{v_j^P}{\partial j} - \frac{\bar{v}_j}{\partial j} = \frac{\delta_j}{\partial j}$$

with $\delta_j = \int_0^j [\varphi(l, j, PAM) - 1] ps_l a'_l dl > 0$ and $\frac{\partial \delta_j}{\partial j} > 0$ if $k \neq 0$

⁴In a frictionless CEO market, board i 's utility function when she hires CEO j is $u(s_i, a_j, v_j) = ps_i a_j - v_j$. Since PAM is the equilibrium sorting pattern, CEO j matches with firm $i = j$. Because hiring CEO j is an optimal choice for board $i = j$, first order condition yields $ps_j a'_j - v'_j = 0$. Integrating along CEO's rank gives out $v_j = \int_0^j ps_l a'_l dl$. Note here we assume $v_0 = 0$, i.e., the CEO with rank 0 gets 0 equilibrium wage.

$\delta_j = v_j^P - \bar{v}_j$ is the difference between v_j^P and \bar{v}_j to CEO j due to her superior information, thus we call δ_j as the information rent to CEO j . $\delta_j > 0$ and $\frac{\partial \delta_j}{\partial j} > 0$ show that information rent is always positive and higher for a more talented CEO.

$v_j^P = \bar{v}_j + \delta_j$ shows that CEO compensation can be decomposed into two parts: frictionless CEO wage \bar{v}_j and information rent δ_j . A positive information rent implies CEO compensation is higher in frictional CEO labor market than that in frictionless CEO labor market. This is because in our model, under optimal contract, paying a CEO higher wage implies a lower probability of monitoring, and since monitoring is costly, the board thus tends to overpay the CEO to reduce total monitoring cost. The result here is interesting at least for two reasons. First, it contributes to CEO compensation literature by showing that information rent incorporated into matching market can better explain why CEO compensation is so high. Second, traditional CEO compensation literature either attributes CEO wage to efficient contracting OR rent extraction. Our result here combines two explanations to show CEO compensation is determined both from efficient contracting (the frictionless CEO wage) AND rent extraction (information rent).

$\frac{v_j^P}{\partial j} - \frac{\bar{v}_j}{\partial j} = \frac{\delta_j}{\partial j} > 0$ shows that under frictional CEO labor market, due to information friction, not only CEO wage is higher, but CEO wage increases faster as CEO rank increases, which is a result of a more talented CEO extracting higher information rent. There are two reasons for this. The first is, any information rent extracted by lower talent CEOs will be at least partially captured by higher talent CEOs, by noting that a lower talent CEO's compensation is a higher talent CEO's outside option. The second is, under PAM, a higher talent CEO matches with a larger firm, and in a larger firm, monitoring cost is even higher because of higher probability of monitor (again from optimal contract), thus board at a larger firm has even higher incentive

to overpay the CEO, which pushes information rent even higher for a more talented CEO.

3.2 Monitoring CEOs

How board of directors monitor the CEO is an important component of corporate governance, which can be defined as a system to prevent management from making decisions that benefit themselves but are detrimental to shareholders and other stakeholders. In this section, we interpret probability of monitoring as corporate governance. We acknowledge that corporate governance is a multi-dimensional measure (Zingales, 1998; Shleifer and Vishny, 1997; Gillan and Starks, 1998), and our model only attempts to capture one very specific aspect of corporate governance, which is how intensely boards monitor CEOs. Besides, our interpretation is also similar to Acharya and Volpin (2010) and Dicks (2012).

Boards of directors are always criticized by regulators, shareholders and the public that they fail to monitor the CEO and serve their fiduciary role as a watchdog for shareholders. A widely accepted view is that boards of directors are reluctant to monitor CEOs due to CEOs' managerial power and influence. As Kieff and Paredes (2013) put it, "reluctance of monitoring boards to cross their CEOs is legendary". In this section, we provide an alternative explanation to why corporate governance might be weak from the view of optimal contract. To state the proposition, we denote the density functions for firm size and CEO talent as $f(s)$ and $t(a)$.

Proposition 6. Given sorting pattern M , for firm i ,

- i. When $M'(i) > 0$, $\frac{\partial g_i}{\partial i} < (>)0$ if and only if $t(a_{M(i)})$ is small (large) enough compared with $f(s_i)$
- ii. When $M'(i) < 0$, $\frac{\partial g_i}{\partial i} > 0$ regardless of firm size and CEO talent distribution.

The above proposition shows that both sorting pattern and distributions of firm size and CEO talent can affect corporate governance. By noting that density function measures scarcity and smaller density indicates a higher level of scarcity, (i.) shows a seemingly striking result: larger firms can have weaker corporate governance than smaller firms, despite of the fact that corporate governance might be more important at a large firm because of more stakes on table. The central message here is that board of directors will not impose a strong corporate governance when CEO talent is sufficiently scarce. We emphasize here that what matters is relative scarcity: the scarcity of CEO talent vs the scarcity of large firms.

The intuition for (i.) is as follows. Larger firms tend to have higher corporate governance because higher probability of monitoring is needed at larger firms to prevent CEOs from misreporting. On the other hand, larger firms match with higher talent CEOs under PAM, and because higher talent CEOs are paid with a higher compensation (Proposition 4) in equilibrium, it implies that larger firms tend to have lower corporate governance. When talent is sufficiently scarce, CEO compensation will be sufficiently high due to equilibrium force and the effect from compensation to lower governance outweighs the effect from size to increase governance, and corporate governance decreases with firm size. Here we provide an example to illustrate the relation between corporate governance and scarcity.

Example 1. Assume $k \leq \frac{ps_0a_0}{2(1-p)}$, thus PAM is the equilibrium sorting pattern and $M'(i) > 0, \forall i \in [0, 1]$. According to Lemma 2, under PAM, corporate governance at firm i is $g_i = 1 - \frac{\int_0^j \varphi(l, j, k, PAM) s_l a_l' dl}{s_i a_i}$ with $\varphi(l, j, k, PAM) = \exp[-\int_l^j \frac{\frac{(1-p)k}{2} a_x'}{1 - \frac{(1-p)k}{ps_x a_x}} dx] / (1 - \frac{(1-p)k}{ps_l a_l})$. We consider two extreme scenarios.

First, CEO talent is scarce: all firms have the same size s and CEOs have different talent. It's easy to show that $\varphi(l, j, k, PAM) = \frac{1}{1 - \frac{(1-p)k}{ps a_j}}$. And thus corporate

governance at firm i is $g_i = 1 - \frac{1 - \frac{a_0}{a_i}}{1 - \frac{(1-p)k}{psa_i}}$. Simple algebra shows that g_i is decreasing in a_i .⁵ Thus corporate governance will be weaker for a firm matching with a more talented CEO.

Second, CEO talent is in sufficient supply: all CEOs have the same talent a , and all firms have different size. Proposition 4 implies CEOs with the same talent are paid with the same wage, which is denoted as v_a , then $g_i = 1 - \frac{v_a}{s_i a}$. Thus corporate governance is increasing in firm size.

(ii.) shows that corporate governance increases in firm size when $M'(i) < 0$ regardless of distributions of firm size and CEO talent. This is because when $M'(i) < 0$, larger firms match with lower talented CEOs. Because a lower talented CEO is bounded to be paid with a lower wage according to Proposition 4, the board at a larger firm will impose a stronger corporate governance to prevent the CEO from misreporting, as now incentive compatibility is harder to satisfy with lower CEO compensation.

Proposition 6 shows that firm-level corporate governance is a function of aggregate market-level characteristics such as talent scarcity and sorting pattern. This explains why empirical research on the relation between firm size and corporate governance gives mixed results without controlling for the relative scarcity of CEO talent and sorting pattern between CEOs and firms. For example, Klapper and Love (2004) shows that firm size positively affect corporate governance and Black et al. (2006) find a negative correlation between firm size and corporate governance index using Korea data.

⁵Note that $\frac{\partial g_i}{\partial a_i} = \frac{ps}{(a_i ps + k(p-1))^2} (k(1-p) - a_0 ps)$ and by $k \leq \frac{ps_0 a_0}{2(1-p)}$, we have $k(1-p) - a_0 ps \leq \frac{ps_0 a_0}{2(1-p)} (1-p) - a_0 ps < 0$, thus $\frac{\partial g_i}{\partial a_i} < 0, \forall i \in [0, 1]$.

3.3 Comparative Statics

We have examined how sorting pattern is related to marginal cost of monitoring k and the probability of “high state” p , and it will also be interesting to see how CEO compensation and corporate governance is affected by k and p . However, it’s difficult to fully pin down the comparative statics result because changing k and p might simultaneously change sorting pattern, CEO compensation and corporate governance. One special case is when k and p satisfy $k \leq \frac{ps_0a_0}{2(1-p)}$, from Proposition 2, we know PAM is the equilibrium sorting pattern. Thus when k changes in the region of $[0, \frac{ps_0a_0}{2(1-p)}]$, sorting pattern will be the same and it gives us an opportunity to see how CEO compensation and corporate governance will be affected while fixing equilibrium sorting pattern as PAM. We thus have the following proposition:

Proposition 7. When $\{k, p\} \in \{k, p | 0 \leq k \leq \frac{ps_0a_0}{2(1-p)}\}$

- i. Information rent for each CEO will be higher as k increases,
- ii. Corporate governance at each firm will be weaker as p decreases or k increases

Thus a higher k will generate a higher information rent, which is easy to understand because a larger k means monitoring is more costly, thus the board will tend to overpay the CEO and reduce corporate governance to save monitoring cost. Thus information rent is higher (and also CEO compensation) and corporate governance is lower when k is larger. How p affects CEO compensation and corporate governance is different from k . Note that a larger p will affect CEO compensation from two aspects, first, the CEO will get paid better because marginal product is higher, second, it will be paid worse because a larger p is similar to a smaller k which means the board will prefer a lower compensation and higher corporate governance. However, the effect from higher marginal product will be cancelled out in setting corporate governance

and thus it's easy to see corporate governance will be stronger as p increases, or in other words, corporate governance will be weaker as p decreases.

4 EXTENSIONS

A salient feature of our model is that all boards have the same marginal cost of monitoring. As we discussed in previous section, the reason for such assumption is because of technical difficulties in dealing with multi-dimensional types if firms are heterogenous in both size and marginal monitoring cost and our central focus on how level instead of heterogeneity of such cost will affect sorting pattern. Nevertheless, for completeness, we discuss two special cases when firms have different marginal monitoring cost.

4.1 Two by Two Case

The setup here is similar to Subsection 2.2.1 discussing the equilibrium sorting condition with two firms and two CEOs except now we assume the two firms have different marginal costs of monitoring. Firm 0 and firm 1's marginal costs of monitoring are $k - \delta$ and $k + \delta$ respectively, with $\delta \leq k$. Note that we don't rule out $\delta < 0$. When $\delta > 0$, firm 1 has higher marginal cost of monitoring than firm 0 and vice versa when $\delta < 0$. Similarly as in Subsection 2.2.1, we can solve the equilibrium sorting condition: there is PAM between firm size and CEO talent if and only if

$$k \leq \frac{p}{(1-p) \left(\frac{1}{a_1 s_1} + \frac{1}{a_1 s_0} \right)} + \delta \frac{s_1^2 + s_0^2}{s_1^2 - s_0^2}$$

Note here k is the average marginal cost of monitoring which measures the "level" and δ measures "heterogeneity" of marginal cost. It's clear from the above inequality that both level and heterogeneity of marginal cost are important in determining

sorting pattern. The similarity here with Subsection 2.2.1 (when both firms have the same marginal cost) is, a higher k will push the market from PAM to NAM while holding δ the same, as long as δ is not too large, even a non-zero δ implies heterogeneity between marginal cost. While holding k fixed, a larger δ will always push sorting towards PAM. Note that as δ increases, large firm has larger marginal cost and smaller firm has smaller marginal cost. Next we discuss a special case when marginal cost is increasing in firm size and firms have different firm size.

4.2 Size-dependent Marginal Cost

We assume here monitoring cost is $ks^\theta g$, $\theta \in [0, 1]$, and thus marginal cost of monitoring is ks^θ , which is obviously size-dependent. To understand such size dependence, we can consider that conditional on auditing, CEO's report is more complicated at a larger firm, thus the board needs to exert more effort in verifying its authenticity.

By applying similar steps as in solving Subsection 2.1's single firm problem, we can solve the optimal contract, which is exactly the same as in Subsection 2.1. This is not surprising because we have shown that optimal contract can be fully pinned down by CEO's incentive constraints, and the only change here is the cost function. The board's utility function is:

$$u(s, a, v) = psa - (1 - p)ks^\theta - \left[1 - \frac{(1 - p)ks^{\theta-1}}{pa}\right]v \quad (6)$$

which can be used to solve equilibrium sorting pattern. The proof is essentially the same as the proof for Proposition 2, which is presented in Appendix A. We can show that there is positive assortative matching between firm size and CEO talent if and only if $k \leq \frac{ps_0^{1-\theta} a_0}{(1-p)(2-\theta)}$.

Note that boards' participation constraint requires $k < \frac{ps_0^{1-\theta}a_0}{(1-p)}$. If $\theta \neq 1$ and $k \in (\frac{ps_0^{1-\theta}a_0}{(1-p)(2-\theta)}, \frac{ps_0^{1-\theta}a_0}{(1-p)})$, PAM fails. When $\theta = 1$, $(\frac{ps_0^{1-\theta}a_0}{(1-p)(2-\theta)}, \frac{ps_0^{1-\theta}a_0}{(1-p)})$ is empty, and in fact, PAM is always the equilibrium sorting pattern irrespective of the value of k . This shows that our result on the failure of PAM applies to a large scope of questions.

5 IMPLICATIONS AND EMPIRICAL PREDICTIONS

5.1 Implications

CEO turnover. Empirical research shows that CEO turnover is strongly tied to economic shock. However, according to relative performance evaluation literature, market-level shock should be filtered out in the decision of retaining and firing CEOs. By discussing how sorting pattern is related to p , our result can explain the seemingly puzzling empirical findings. Notice that p measures the probability of "high output", then a decrease in p can be considered as a negative industry shock and an increase in p can be considered as a positive industry shock. We have discussed that a decrease in p causes PAM to fail and as p continues to decrease, PAM will fail for more and more bottom firms and CEOs. We can consider a change in sorting pattern as CEO turnover, thus our result shows that negative industry shock causes CEO turnover in the following direction: worse CEOs get hired by larger firms and better CEOs work for smaller firms.

Robustness of PAM. A set of papers using PAM as the basis for empirical applications has generated fruitful results(e.g., Tervio, 2008; Gabaix and Landier, 2008; Pan, 2010; Eisfeldt and Kuhnen, 2013). PAM is the equilibrium sorting pattern in a frictionless CEO market with complementary production technology. However, when we consider a more realistic setting in which CEOs have superior information and monitoring CEOs is costly, PAM is the equilibrium sorting pattern only when marginal

cost of monitoring is below certain threshold according to Proposition 2. Thus PAM is not a “take-as-given” result, it can not be assumed or implied even with complementary production technology. Proposition 3 thus provides a robustness check, by showing that when PAM fails, it can still exist as the equilibrium sorting pattern for larger firms and more talented CEOs. This implies that empirical applications relying on PAM are more robust with a sample of large firms (e.g., Gabaix and Landier (2008) use S&P 500 data).

5.2 Empirical Predictions

From Proposition 1, CEO’s expected wage is $pw_1 = (1 - g)sa$ and thus $1 - g$ is the dollar change in the CEO’s wage associated with one dollar change in firm revenue, which is CEO’s pay-performance sensitivity, as defined in Jensen and Murphy (1990). We can also note that as g decreases, the CEO’s expected wage increases. This implies corporate governance and incentive contract are substitutes and CEO ownership can reverse the negative impact of weak governance (Lilienfeld-Toal, 2014). Thus we have the following prediction about the relation between CEO’s pay-performance sensitivity and CEO compensation with corporate governance, which has been supported by the empirical evidence from Fahlenbrach (2009).

1. Pay-performance sensitivity and CEO compensation decrease with corporate governance.

Another interesting and important empirical question is, is there PAM in CEO labor market? If a number of empirical applications reply on PAM, then, it’s at least necessary to understand if we can truly observe PAM in data. A more interesting way to address this is, how the existence of PAM is connected to various explanatory variables? Proposition 2 connects sorting pattern to marginal cost of monitoring, which

can be proxied by board characteristics. As we discussed in Section 2.1, board of directors suffers from disutility because of their lack of independence, and in fact, it's not hard to show that this cost is also decreasing in board ownership and reputational concerns.⁶ Thus we can connect sorting pattern to board independence, ownership and reputational concerns. Cremers and Grinstein (2013) show that CEO labor market is likely to be fragmented across different industries, thus the fragmentation of CEO labor market provides us an opportunity for cross-industry tests. Proposition 2 also predicts PAM is more likely to fail under negative industry shock, which has been discussed in Section 2.2.2 and in the discussion of CEO turnover. Proposition 3 shows that PAM fails for smaller firms and lower talent CEOs first. We thus have the following empirical predictions on sorting pattern:

2. PAM between firm size and CEO talent is more likely to fail

- (i.) for an industry with lower board independence, board ownership or reputational concerns,
- (ii.) in economic downturn,
- (iii.) for smaller firms and lower talent CEOs.

The first thing to notice here is that testing PAM is equivalent to testing a monotonic relationship between firm size and CEO talent, which does not imply or is implied by a positive correlation. Therefore, instead of using correlation between firm size and CEO talent, we should use rank coefficient (e.g., Spearman's ρ or Kendall's τ) as the empirical measure of PAM. In order to construct the rank coefficient, we need data on firm size and CEO talent. Empirical literature uses market capitalization,

⁶Assume board has utility function $u = \alpha(q - w) - cg$, with α increases with board's stock ownership and reputational concerns and c increases with board's lack of independence. Here board independence can be considered as percentage of independent directors among all directors. Independent directors are board members who do not have any obvious relationship with the firm or its senior executives that potentially would give rise to a conflict of interest. A simple monotone transformation yields $u = q - w - k$ with $k = \frac{c}{\alpha}$. Thus k is higher when board has lower independence, ownership or reputational concerns.

assets, or sales to measure firm size. The difficulty in this test is the unobservable nature of CEO talent. To measure CEO talent, we can use CEO's educational attainment (Cole and Mehran, 2010), media coverage, age of first job as CEO, and undergraduate school ranking (Falato et al., 2012, 2013). However, testing result would be questionable if we are not sure how much it is affected by the measurement errors in constructing proxy measures for talent.

Proposition 4 shows that CEO compensation is always an increasing function of CEO talent regardless of whether PAM is true, this implies that there is higher strength of monotonicity between CEO compensation and CEO talent than between CEO compensation and firm size. As we discussed before, the strength of monotonicity can be tested by rank coefficient. Define $\frac{v_j^P}{\partial j}$ as wage spread between different CEOs, then Proposition 5 states that wage spread will be higher for higher talent CEOs. Under PAM, higher talent CEOs match with larger firms, thus wage spread increases with firm size under PAM. Comparative statics result from Proposition 7 shows that CEO compensation increases with marginal cost of monitoring because information rent is higher, and we have shown that marginal cost of monitoring is decreasing in board independence, board ownership and board reputational concerns. We thus have the following empirical predictions on CEO compensation:

3. i. Monotonicity between CEO compensation and CEO talent is stronger than that between CEO compensation and firm size
 - ii. Under PAM, wage spread increases with firm size
 - iii. Under PAM, CEO compensation is higher for an industry with lower boards' independence, boards' ownership or reputational concerns

A board with higher ownership can be a board with more blockholders as directors. Core et al. (1999) find that when a firm has an independent director who owns at least 5% of the shares, CEO compensation will be lower. Agrawal and Nasser (2012) show

that firms with independent blockholder directors have lower CEO compensation. Becker et al. (2011), Cyert et al. (2002), and Bebchuk et al. (2010) have similar findings. Chhaochharia and Grinstein (2009) document that board independence is associated with a reduction in CEO compensation.

According to Proposition 6, under PAM, when CEO talent becomes sufficiently scarce, corporate governance can decrease in firm size. One thing to notice is that we are dealing with relative scarcity here: when there are very few high talent CEOs compared with large firms, we call there is a scarcity of CEO talent. Of course, the opposite can happen: when there is a scarcity of large firms, larger firms will have stronger corporate governance than smaller firms. Proposition 7 shows that when p is smaller, corporate governance is weaker, thus corporate governance is weaker in economic downturn. We state our predictions on corporate governance as follows:

4. Under PAM

- i. When there is a scarcity of CEO talent (large firms), larger firms have weaker (stronger) corporate governance than smaller firms
- ii. Corporate governance is weaker at economic downturns

6 CONCLUSION

This paper models a CEO labor market where boards of directors make decisions on which CEO to hire, how much to pay the CEO and how often to monitor the CEO. CEOs have private information about firm production and can appropriate firm resources by misreporting firm output. Boards of directors monitor CEOs with a cost, which is affected by board independence, ownership and reputational concerns.

We have three novel findings on allocation of CEO talent, CEO compensation and corporate governance. First, positive assortative matching between firm size and

CEO talent can fail when marginal cost of monitoring is large enough or profitability of firm production is low enough, which explains why CEO turnover is connected with industry shock. A more important result is once PAM fails, it will fail first for smaller firms and lower talent CEOs, which implies empirical applications relying on PAM are more robust when using a subsample of large firms. Second, CEO compensation can be decomposed into frictionless competitive market pay and information rent, thus CEO compensation in our model is higher than that solved from frictionless CEO labor market, which helps explain why CEO compensation is so high. Third, firm-level corporate governance is affected by aggregate market characteristics such as the scarcity of CEO talent and sorting pattern. Larger firms can have weaker corporate governance than smaller firms when CEO talent is sufficiently scarce. This is in sharp contrast with the traditional view which attributes weak corporate governance solely to board of directors' failure to monitor the CEO.

In this paper, we assume marginal cost of monitoring CEO is the same to all boards in the CEO labor market, which is a fairly restrictive assumption. We show that assuming this provides us tractability in analyzing complicated matching problem and answering important questions on how the level of marginal monitoring cost (instead of the difference in marginal monitoring cost among firms) would affect allocating, compensating and monitoring CEOs. However, it's a promising direction to consider a market where boards of directors in different firms have different characteristics and examine the effects of both the level and heterogeneity of board characteristics among firms on the allocation of CEO talent. However, a general and comprehensive approach requires solving multi-dimensional matching problem, which isn't an easy task, we thus only provide a rudimentary analysis on this matter in the extension by considering two special cases.

References

- Acharya, Viral V, Marc Gaborro, and Paolo F Volpin (2012), “Competition for managers, corporate governance and incentive compensation.” *Corporate Governance and Incentive Compensation* (April 2012). CEPR Discussion Paper No. DP8936.
- Acharya, V.V. and P.F. Volpin (2010), “Corporate governance externalities*.” *Review of Finance*, 14, 1–33.
- Agrawal, Anup and Tareque Nasser (2012), “Blockholders on boards and ceo compensation, turnover and firm valuation.” In CELS 2009 4th Annual Conference on Empirical Legal Studies Paper.
- Bebchuk, Lucian A, Yaniv Grinstein, and Urs Peyer (2010), “Lucky ceos and lucky directors.” *The Journal of Finance*, 65, 2363–2401.
- Becker, Bo, Henrik Cronqvist, and Rüdiger Fahlenbrach (2011), “Estimating the effects of large shareholders using a geographic instrument.” *Journal of Financial and Quantitative Analysis*, 46, 907–942.
- Becker, Gary S (1973), “A theory of marriage: Part i.” *The Journal of Political Economy*, 813–846.
- Black, Bernard S, Hasung Jang, and Woochan Kim (2006), “Does corporate governance predict firms’ market values? evidence from korea.” *Journal of Law, Economics, and Organization*, 22, 366–413.
- Border, Kim C and Joel Sobel (1987), “Samurai accountant: A theory of auditing and plunder.” *The Review of economic studies*, 54, 525–540.
- Chade, Hector, Jan Eeckhout, and Lones Smith (2014), “Search and matching models in microeconomics.” working paper.
- Chhaochharia, Vidhi and Yaniv Grinstein (2009), “Ceo compensation and board structure.” *The Journal of Finance*, 64, 231–261.
- Cole, Rebel and Hamid Mehran (2010), “What can we learn from privately held firms about executive compensation?”
- Core, John E, Robert W Holthausen, and David F Larcker (1999), “Corporate governance, chief executive officer compensation, and firm performance.” *Journal of financial economics*, 51, 371–406.
- Cremers, KJ Martijn and Yaniv Grinstein (2013), “Does the market for ceo talent explain controversial ceo pay practices?” *Review of Finance*, rft024.

- Cyert, Richard M, Sok-Hyon Kang, and Praveen Kumar (2002), “Corporate governance, takeovers, and top-management compensation: Theory and evidence.” *Management Science*, 48, 453–469.
- Dicks, David L (2012), “Executive compensation and the role for corporate governance regulation.” *Review of Financial Studies*, 25, 1971–2004.
- Edmans, A. and X. Gabaix (2011), “The effect of risk on the ceo market.” *Review of Financial Studies*, 24, 2822–2863.
- Edmans, Alex, Xavier Gabaix, and Augustin Landier (2009), “A multiplicative model of optimal ceo incentives in market equilibrium.” *Review of Financial Studies*, 22, 4881–4917.
- Eisfeldt, Andrea L and Camelia M Kuhnen (2013), “Ceo turnover in a competitive assignment framework.” *Journal of Financial Economics*.
- Fahlenbrach, Rüdiger (2009), “Shareholder rights, boards, and ceo compensation*.” *Review of Finance*, 13, 81–113.
- Falato, Antonio, Dan Li, and Todd Milbourn (2012), “Ceo pay and the market for ceos.”
- Falato, Antonio, Dan Li, and Todd T Milbourn (2013), “Which skills matter in the market for ceos? evidence from pay for ceo credentials.” *Evidence from Pay for CEO Credentials* (April 1, 2013).
- Gabaix, X. and A. Landier (2008), “Why has ceo pay increased so much?” *The Quarterly Journal of Economics*, 123, 49–100.
- Gillan, Stuart and Laura T Starks (1998), “A survey of shareholder activism: Motivation and empirical evidence.” *Contemporary Finance Digest*, 2, 10–34.
- Hermalin, B.E. and M.S. Weisbach (1998), “Endogenously chosen boards of directors and their monitoring of the ceo.” *American Economic Review*, 96–118.
- Jensen, Michael C and Kevin J Murphy (1990), “Performance pay and top-management incentives.” *Journal of political economy*, 225–264.
- Klapper, Leora F and Inessa Love (2004), “Corporate governance, investor protection, and performance in emerging markets.” *Journal of corporate Finance*, 10, 703–728.
- Legros, Patrick and Andrew F Newman (2007), “Beauty is a beast, frog is a prince: assortative matching with nontransferabilities.” *Econometrica*, 75, 1073–1102.
- Lilienfeld-Toal, Stefan, Ulf Vonand Ruenzi (2014), “Ceo ownership, stock market performance, and managerial discretion.” *The Journal of Finance*, 69, 1013–1050.

- Pan, Yihui (2010), "The determinants and impact of executive-firm matches." Unpublished working paper. University of Minnesota.
- Shleifer, Andrei and Robert W Vishny (1997), "A survey of corporate governance." *The journal of finance*, 52, 737–783.
- Tervio, M. (2008), "The difference that ceos make: An assignment model approach." *The American Economic Review*, 98, 642–668.
- Zingales, Luigi (1998), "Corporate governance, in the new palgrave dictionary of economics and the law, p. newman ed."

APPENDIX A

PROOFS MAY-2015

A PROOFS

Proof to Proposition 2

Proof. First, we prove that $v_j \geq v_0, \forall j \in [0, 1]$, where v_j is CEO j 's compensation and v_0 is CEO 0's compensation. Assume there exists a CEO $j' > 0$ such that $v_{j'} < v_0$. Then the matching cannot be stable because the firm initially matching with CEO 0 will choose to match with CEO j' by paying CEO j' a wage of v_0 . The firm will be better off because $\frac{\partial u(s,a,v)}{\partial a} > 0$ implies the board will get a strictly higher utility by shifting to CEO j' and CEO j' will be better off because of a higher wage.

We use the sufficient and necessary condition for PAM from Chade et al. (2014), which states that PAM is the equilibrium sorting pattern if and only if $u_{sv}u_a - u_{sa}u_v \geq 0, \forall s, a, v$. u_{sv} is the cross-partial of the board's utility function (2) on s and v . u_a, u_{sa}, u_v can be similarly defined. Simple algebra yields the sufficient and necessary condition as

$$-2\frac{(1-p)k}{p}pa^{-1}s^{-1} + p + \frac{(1-p)k}{p}s^{-2}a^{-2}v \geq 0, \forall s, a, v$$

Because we have $v \geq v_0, \forall v$, the above inequality is equivalent to

$$-2\frac{(1-p)k}{p}pa^{-1}s^{-1} + p + \frac{(1-p)k}{p}s^{-2}a^{-2}v_0 \geq 0, \forall s, a$$

which is equivalent to $\frac{sa}{2-\frac{v_0}{psa}}\frac{p}{1-p} \geq k, \forall s, a$. Because $s \geq s_0, a \geq a_0, \forall s, a$, we have

$\frac{s_0a_0}{2-\frac{v_0}{ps_0a_0}}\frac{p}{1-p} \geq k$. When $v_0 = 0$, the sufficient and necessary condition for PAM is

$$k \leq \frac{s_0a_0}{2} \frac{p}{1-p}$$

□

Proof to Proposition 3

We first prove a lemma to show the necessary condition for NAM among any two pairs of firms and CEOs.

Lemma 3. There is NAM for firms $i > j$ and CEOs $x > y$ only if $k \geq \frac{p}{(1-p)} \frac{1}{\frac{1}{s_j a_x} + \frac{1}{s_i a_x} - \frac{v_y}{ps_i a_x s_j a_y}}$, where v_y is CEO y 's compensation.

Proof. Stability under NAM requires

$$\begin{aligned} ps_j a_y - \left[1 - \frac{(1-p)k}{ps_j a_y}\right] v_y &\leq ps_j a_x - \left[1 - \frac{(1-p)k}{ps_j a_x}\right] v_x \\ ps_i a_x - \left[1 - \frac{(1-p)k}{ps_i a_x}\right] v_x &\leq ps_i a_y - \left[1 - \frac{(1-p)k}{ps_i a_y}\right] v_y \end{aligned}$$

The above inequalities imply

$$\frac{ps_i a_x - ps_i a_y + \left[1 - \frac{(1-p)k}{ps_i a_y}\right] v_y}{1 - \frac{(1-p)k}{ps_i a_x}} \leq v_x \leq \frac{ps_j a_x - ps_j a_y + \left[1 - \frac{(1-p)k}{ps_j a_y}\right] v_y}{1 - \frac{(1-p)k}{ps_j a_x}}$$

which solves the necessary condition for NAM

$$k \geq \frac{p}{(1-p)} \frac{1}{\frac{1}{s_j a_x} + \frac{1}{s_i a_x} - \frac{v_y}{ps_i a_x s_j a_y}}$$

□

Next we use the above lemma to prove our proposition.

Proof. We prove the proposition in two steps, first we prove any firm or CEO in $[\bar{i}, 1]$ will not match with any firm or CEO in $[0, \bar{i}]$. We prove it by contradiction. We assume there exists a firm $i \in [\bar{i}, 1]$ and CEO $y \in [0, \bar{i}]$ such that firm i matches with CEO y . Because we have the same measure of firms and CEOs on $[\bar{i}, 1]$ and all firms

and CEOs will be matched, then there must exist a firm $j \in [0, \bar{i}]$ and a CEO $x \in [\bar{i}, 1]$ such that x and j match with each other. Because $i > j$ and $x > y$, we thus have NAM here. Because $i > \bar{i} > j$ and $x > \bar{i} > y$, we have $a_x > a_{\bar{i}}$ and $s_i > s_j > s_0$, thus $\frac{2}{s_0 a_{\bar{i}}} > \frac{1}{s_0 a_x} + \frac{1}{s_i a_x} > \frac{1}{s_j a_x} + \frac{1}{s_i a_x}$. And by $\bar{i} = \{i : k = \frac{ps_0 a_i}{2(1-p)}\}$,

$$\begin{aligned} k &= \frac{p}{\frac{2}{s_0 a_{\bar{i}}}(1-p)} \\ &< \frac{p}{(1-p)} \frac{1}{\frac{1}{s_j a_x} + \frac{1}{s_i a_x}} \\ &< \frac{p}{(1-p)} \frac{1}{\frac{1}{s_j a_x} + \frac{1}{s_i a_x} - \frac{v_y}{ps_i a_x s_j a_y}} \end{aligned}$$

which violates condition for NAM according to Lemma 3.

Second step, we prove there is PAM on $[\bar{i}, 1]$. We use the sufficient condition for PAM from Chade et al. (2014), which states that if $u_{sv}u_a - u_{sa}u_v \geq 0, \forall s, a, v$, then PAM is the equilibrium sorting pattern. u_{sv} is the cross-partial of the director's utility function (2) on s and v . u_a, u_{sa}, u_v can be similarly defined. Simple algebra yields the sufficient condition: $\forall s, a, v$ on $[\bar{i}, 1]$

$$-2 \frac{(1-p)k}{p} p a^{-1} s^{-1} + p + \frac{(1-p)k}{p} s^{-2} a^{-2} v \geq 0$$

Because $s \geq \bar{s}$ and $a \geq \bar{a}$ and $v \geq 0$ for firms and CEOs on $[\bar{i}, 1]$, the above inequality can be satisfied if

$$-2 \frac{(1-p)k}{p} p a_{\bar{i}}^{-1} s_{\bar{i}}^{-1} + p \geq 0$$

Substitute $k = \frac{ps_0 a_{\bar{i}}}{2(1-p)}$ to the above inequality, we have $-2 \frac{(1-p)k}{p} p a_{\bar{i}}^{-1} s_{\bar{i}}^{-1} + p = p(1 - \frac{s_0}{s_{\bar{i}}}) \geq 0$. \square

Proof to Lemma 1

Proof. We can first prove that once PAM fails, there exists a firm x such that $M'(x) < 0$. This can be proved by contradiction. If there doesn't exist such an x , then $M'(x) > 0$ on $[0, 1]$, which is exactly PAM. Because M is continuously differentiable, there must exist a neighborhood of x , i.e., there exists a $\delta > 0$ such that $M'(x) < 0, i \in (x - \delta, x + \delta)$. Denote $a = x - \delta$ and $b = x + \delta$, thus when PAM fails on $[0, 1]$, we can find a subset of firms (a, b) such that NAM will be the equilibrium sorting pattern. \square

Proof to Lemma 2

Proof. We denote CEO j 's equilibrium wage as v_j under matching M . Then board at firm i 's utility if she hires CEO j is

$$u(s_i, a_j, v_j) = ps_i a_j - (1 - p)k - [1 - \frac{(1 - p)k}{ps_i a_j}]v_j$$

For sorting pattern M , CEO j matches with firm $M(j)$. Because CEO j is the optimal choice for the board at firm $M(j)$, the following first order condition must be satisfied:

$$\frac{\partial u(s_i, a_j, v_j)}{\partial j} \Big|_{i=M(j)} = ps_{M(j)} a'_j - [1 - \frac{(1 - p)k}{ps_{M(j)} a_j}]v'_j - \frac{(1 - p)k}{ps_{M(j)} a_j^2} a'_j v_j = 0$$

The solution to the above differential equation is $v_j = \int_0^j \varphi(l, j, k, M) ps_{M(l)} a'_l dl$, where $\varphi(l, j, k, M) = exp[-\int_l^j \frac{\frac{(1-p)k}{ps_{M(x)} a_x^2} a'_x}{1 - \frac{(1-p)k}{ps_{M(x)} a_x}} dx] / (1 - \frac{(1-p)k}{ps_{M(l)} a_l})$ \square

Proof to Proposition 4

Proof. The first order condition as shown in the proof of Lemma 2 is

$$\frac{\partial u(s_i, a_j, v_j)}{\partial j} \Big|_{i=M(j)} = ps_{M(j)} a'_j - \left[1 - \frac{(1-p)k}{ps_{M(j)} a_j}\right] v'_j - \frac{(1-p)k}{ps_{M(j)} a_j^2} a'_j v_j = 0$$

which solves $v'_j = [ps_{M(j)} a_j - \frac{(1-p)k}{ps_{M(j)} a_j} v_j] \frac{a'_j a_j^{-1}}{1 - \frac{(1-p)k}{ps_{M(j)} a_j}}$. And we have $ps_{M(j)} a_j \geq v_j$ and $\frac{(1-p)k}{ps_{M(j)} a_j} < 0$, thus when $a'_j > 0$, we have $\frac{v_j}{\partial j} > 0$. And when $a'_j = 0$, we have $\frac{v_j}{\partial j} = 0$. \square

Proof to Proposition 5

Proof. We only need to prove $\delta_j \geq 0$ and $\frac{\partial \delta_j}{\partial j} > 0$. First, to prove $\delta_j \geq 0$, we prove $\varphi(l, j, k, PAM) \geq 1$. Note that we have

$$\begin{aligned} \frac{\varphi(l, j, k, PAM)}{\exp\left[\int_l^j \frac{\frac{(1-p)k}{ps_x^2} s'_x}{1 - \frac{(1-p)k}{ps_x a_x}} dx\right]} &= \frac{1}{1 - \frac{(1-p)k}{ps_l a_l}} \exp\left[-\int_l^j \frac{\frac{(1-p)k}{ps_x a_x^2} a'_x + \frac{(1-p)k}{ps_x^2 a_x} s'_x}{1 - \frac{(1-p)k}{ps_x a_x}} dx\right] \\ &= \frac{1}{1 - \frac{(1-p)k}{ps_l a_l}} \exp\left[-\int_l^j \frac{d\left(1 - \frac{(1-p)k}{ps_x a_x}\right)}{1 - \frac{(1-p)k}{ps_x a_x}}\right] = \frac{1}{1 - \frac{(1-p)k}{ps_j a_j}} \end{aligned}$$

Obviously $\exp\left[\int_l^j \frac{\frac{(1-p)k}{ps_x^2} s'_x}{1 - \frac{(1-p)k}{ps_x a_x}} dx\right] \geq 1$ and $\frac{1}{1 - \frac{(1-p)k}{ps_j a_j}} \geq 1$, thus

$$\varphi(l, j, k, PAM) = \exp\left[\int_l^j \frac{\frac{(1-p)k}{ps_x^2} s'_x}{1 - \frac{(1-p)k}{ps_x a_x}} dx\right] \frac{1}{1 - \frac{(1-p)k}{ps_j a_j}} \geq 1$$

It's easy to see $\varphi(l, j, k, PAM) = 1$ if and only if $\exp\left[\int_l^j \frac{\frac{(1-p)k}{ps_x^2} s'_x}{1 - \frac{(1-p)k}{ps_x a_x}} dx\right] = 1$ and $\frac{1}{1 - \frac{(1-p)k}{ps_j a_j}} = 1$, which is true if and only if $k = 0$. Thus $\delta_j \equiv \int_0^j [\varphi(l, j, PAM) - 1] ps_l a'_l dl \geq 0$ and $\delta_j = 0$ if and only if $k = 0$.

To prove $\frac{\partial \delta_j}{\partial j} > 0$, notice that we have $\varphi(l, j, k, PAM) = \frac{\exp[-\int_l^j \frac{(1-p)k}{ps_x a_x^2} a_x']}{1 - \frac{(1-p)k}{ps_l a_l}}$, thus

$$\frac{\partial \varphi(l, j, k, PAM)}{\partial j} = -\frac{\frac{(1-p)k}{ps_j a_j^2} a_j'}{1 - \frac{(1-p)k}{ps_j a_j}} \varphi(l, j, k, PAM).$$

Then the partial derivative of δ_j with respect to j is

$$\begin{aligned} \frac{\partial \delta_j}{\partial j} &= \int_0^j \frac{\partial \varphi(l, j, k, PAM)}{\partial j} ps_l a_l' dl + \left(\frac{1}{1 - \frac{(1-p)k}{ps_j a_j}} - 1 \right) ps_j a_j' \\ &= \left(-\frac{\frac{(1-p)k}{ps_j a_j^2} a_j'}{1 - \frac{(1-p)k}{ps_j a_j}} \right) \int_0^j \varphi(l, j, k, PAM) ps_l a_l' dl + \left(\frac{(1-p)k}{1 - \frac{(1-p)k}{ps_j a_j}} \right) ps_j a_j' \\ &= \frac{\frac{(1-p)k}{ps_j a_j^2} a_j'}{1 - \frac{(1-p)k}{ps_j a_j}} (ps_j a_j - v_j) = \frac{(1-p)k}{1 - \frac{(1-p)k}{ps_j a_j}} \pi_j > 0 \end{aligned}$$

And thus $\frac{\partial \delta_j}{\partial j} > 0$. □

Proof to Proposition 6

Proof. Under matching M , CEO $M(i)$ matches with firm i , and by the first order condition derived in Section 3, we have $v'_{M(i)} = [ps_i a_{M(i)} - \frac{(1-p)k}{ps_i a_{M(i)}} v_{M(i)}] \frac{a'_{M(i)} a_{M(i)}^{-1}}{1 - \frac{(1-p)k}{ps_i a_{M(i)}}}$. Note that $g_i = 1 - \frac{v_{M(i)}}{ps_i a_{M(i)}}$, we have $\frac{\partial g_i}{\partial i} = -\frac{v'_{M(i)} M'(i) ps_i a_{M(i)} - p(s'_i a_{M(i)} + s_j a'_{M(i)} M'(i)) v_{M(i)}}{(ps_i a_{M(i)})^2}$, and substitute $v'_{M(i)}$ into $\frac{\partial g_i}{\partial i}$, we have $\frac{\partial g_i}{\partial i} = \frac{A_i s'_i - B_i a'_{M(i)} M'(i)}{(ps_i a_{M(i)})^2}$ with $A_i = pa_{M(i)} v_{M(i)}$ and $B_i = [\frac{ps_i a_{M(i)} - v_{M(i)}}{1 - \frac{(1-p)k}{ps_i a_{M(i)}}}] ps_i$. It's easy to see that $A_i \in [0, ps_1 a_1]$ and $B_i \in [0, \frac{ps_1 a_1}{1 - \frac{(1-p)k}{ps_0 a_0}} ps_1]$, where $[a_0, a_1]$ and $[s_0, s_1]$ are supports for talent and size distributions respectively. Thus both A_i and B_i are bounded functions of i on $[0, 1]$ for sorting pattern M .

Denote $F(\cdot), T(\cdot)$ as the distribution functions and $f(\cdot)$ and $t(\cdot)$ as the density functions for firm size and CEO talent. Then the size of firm i is $s_i = F^{-1}(i)$, thus $s'_i = \frac{1}{F'(s_i)} = \frac{1}{f(s_i)}$. Similarly, $a'_{M(i)} = \frac{1}{t(a_{M(i)})}$. Therefore, we have

$$\frac{\partial g_i}{\partial i} = \frac{\frac{A_i}{f(s_i)} - \frac{B_i M'(i)}{t(a_{M(i)})}}{(ps_i a_{M(i)})^2}$$

When $M'(i) < 0$, it's easy to see that $\frac{\partial g_i}{\partial i} > 0$. When $M'(i) > 0$, if $f(s_i)$ is sufficiently close to zero and $t(a_{M(i)})$ is sufficiently large, then $\frac{\partial g_i}{\partial i} > 0$; if $f(s_i)$ is sufficiently large and $t(a_{M(i)})$ is close to zero or $M'(i)$ is sufficiently large, then $\frac{\partial g_i}{\partial i} < 0$. \square