

Firm-CEO Matching with Pre-CEO Executive Skill Accumulation

Seungjin Han*

September 17, 2016

Abstract

Executives have high stakes riding on executive skill accumulated in their non-CEO executive positions in order to establish themselves as talented before being on the competitive CEO market. This paper considers an assignment model augmented by costly pre-CEO executive skill accumulation on the job. As the firm size distribution is more spread to the right, causing the increase in firm size, the proportion of CEO pay attributed to the informational effect (the direct productivity effect) of executive skill accumulation uniformly decreases (increases). Subsequently, the equilibrium converges to the competitive equilibrium with complete information, improving efficiency at the cost of worsening income equality. In addition, the paper shows the far-reaching impact of pay limits to control CEO pay. If the CEO pay limit is so severe that the minimum pay cut required for the highest-paid CEO is more than the lowest-paid CEO's expected pay net of her cost of working as a CEO, then it induces an equilibrium CEO pay function discontinuous at the top. This distorts high ability executives' equilibrium skill accumulation and results in a discontinuous bunching even prior to being on the CEO market. It may redistribute earnings from CEOs to shareholders only at the low-end of firm-CEO matches, given the distorted executive skill distribution. (*JEL* G32, M12, M52)

*Address: Department of Economics, McMaster University 280 Main Street West, Hamilton, Ontario L8S 4M4, Canada. Phone: +1-905-525-9140 x23818. E-mail: hansj@mcmaster.ca. Acknowledgement to be added. All errors are mine alone.

Contents

1	Introduction	1
2	Preliminaries	4
2.1	Model	4
2.2	Stable matching equilibrium	6
3	Equilibrium Analysis	9
3.1	Equilibrium characterization	9
3.2	Equilibrium derivation	11
3.3	Equilibrium in Parametrized Models	13
3.3.1	Observable executive types	14
3.3.2	Unobservable executive types	16
3.3.3	Empirical strategy to estimate measure of efficiency	19
4	Policy Analysis: CEO Pay Limits	20
4.1	Moderate Pay Limits	21
4.2	Severe Pay Limits	24
5	Conclusion	33
A	Proof of Proposition 1	34
A.1	Proof of Conditions 1, 2, and 4	34
A.2	Proof of Condition 3	36
B	Proof of Theorem 3	39
B.1	Proof of Conditions 1 and 2	39
B.2	Proof of Conditions 3 and 4	42

1 Introduction

Before becoming chief executive officers (CEOs), executives perform various tasks in their non-CEO executive positions. Tasks usually fall into two categories. Type-1 tasks require firm-specific skill and type-2 general executive skill. Executives exert effort to complete their tasks and get paid in the current positions, but exerting effort for type-2 tasks has far reaching impacts beyond the current executive positions.

CEO jobs have increasingly emphasized general executive rather than firm-specific skill (Frydman (2014) and Murphy and Zbojnik (2004)). Therefore, in order to establish themselves as talented executives before being on the highly competitive CEO market, it is very important for executives to accumulate general executive skill through exerting effort for type-2 tasks in the current executive positions.¹ This is similar to “learning by doing” as in Gayle et al. (2015). An executive who accumulates general executive skill beyond the level that is required to complete type-2 tasks in the current position can be rewarded in the future if she is promoted to a CEO by the current firm or hired as a CEO by another firm.²

To study executives’ general executive skill accumulation prior to being on the CEO market, we extend the competitive assignment model (Sattinger (1993), Tervio (2008), Gabaix and Landier (2008)) by formulating the executive’s effort decision in their current positions for type-2 tasks, together with her effort decision for type-1 tasks, prior to being on the CEO market. Firms differ in their size and executives in their inherent unobservable ability, called type. As the executive exerts more effort for type-2 tasks, she can accumulate general executive skill potentially beyond the level that is required to complete type-2 tasks. Our paper highlights the executive’s incentive to exert extra effort to accumulate general executive skill beyond the level that is required in the current executive position. This incentive is created by the additional reward that she can receive as a CEO in the future. This additional reward is the discounted value of the expected CEO pay net of the cost of working in the CEO position.

¹The high demand for CEO talent in the market is well understood (Bebchuk and Fried (2004), Kaplan (2013)).

²This is very similar to how markets for other professionals operate (e.g. corporate lawyers, academic researchers). For examples, academic researchers can accumulate research skills beyond the level that is required by the current institutions. Those additional research skills are not necessarily compensated right away by the current institutions but bring them hard-to-resist offers from other, often better, institutions. It may lead the current institutions to make better offers or just let them go for the offers from those competing institutions. This is the reward for the additional research skill accumulated in the current research institution.

A CEO's overall talent is composed of both her inherent *unobservable type* and *general executive skill* that she accumulates in her previous non-CEO executive position. Firm size and CEO's overall talent are complementary in the firm's earnings that its CEO generates. Subsequently, larger firms can pay out huge sums of compensation to their CEOs who had already established themselves as talented executives with plenty of prior public observations on their general executive skills. Therefore, our paper is interested in a *separating stable matching equilibrium* where an executive of a higher type accumulates a higher level of executive skill in the current position, to distinguish herself from others and the market's perception of the executive's overall talent is correct.

In the CEO market that executives enter with their accumulated general executive skill, equilibrium matching is stable in the sense that there are no pairs of a firm and a CEO who, by matching and transferring part of the firm's earnings to the CEO, can make themselves strictly better off. We first present the closed-form analysis for the separating stable matching equilibrium in a parametrized model. Given the equilibrium property of the assortative matching between firm size and CEO's type, we assume that the size x of the firm that a CEO of type θ works for in equilibrium follows the functional form, $x = k\theta^q$, where k is the "shift" parameter and q is the "relative spacing" parameter. Given k , the relative spacing parameter q shows the relative heterogeneity of firm size to the executive's inherent ability. This functional form can be derived under several reasonable distributions for firm size and inherent ability such as Weibull, exponential, normal, Pareto, etc.

We show that the market pay for CEO consists of two parts. First of all, there is part of the market pay attributed to the direct productivity effect of executive skills accumulated prior to becoming a CEO. Secondly, more executive skills accumulated in previous non-CEO executive positions improve the market's perception on the executive's type (and hence on overall talent). Therefore, there is part of the market pay attributed to this informational effect of executive skills, which would be absent under complete information about the executive's type.

Given the executive type distribution, a higher q implies the spread of the firm size distribution to the right, causing an increase in the size of the firm that each CEO works for. We show that as $q \rightarrow \infty$, the proportion of the market pay attributed to the informational effect (the direct productivity effect) of executive skills is decreasing (increasing) in q and uniformly converges to zero (one). As the proportion of the market pay attributed to the information effect uniformly decreases, the separating stable matching equilibrium with unobservable types converges to the competitive equilibrium that would prevail with observable types. Therefore, as q increases, the equilibrium outcome uniformly converges to

the fully Pareto efficient outcome. The proportion of the market pay attributed to the direct productivity effect is indeed the measure of efficiency in that how close the equilibrium outcome is to the fully Pareto efficient outcome positively depends on this proportion. We present the empirical strategy to estimate this measure of efficiency.

Gabaix and Landier (2008) show the increase in CEO pay over time is largely attributed to the increase in the firm size, based on the assignment model without market imperfection and skill accumulation. Mishel and Davis (2014) show that CEOs of major U.S. companies earned 20 times more than a typical worker in 1965. This ratio has grown dramatically over time.³ From its 2009 low, the CEO-to-worker compensation ratio in 2013 had reached the value of 303.4-to-1, a rise of 107.6 since 2009. Given their results, our paper is the first to show that as the firm size distribution is spread to the right, causing the increase in firm size, efficiency in the CEO market improves at the cost of worsening income inequality and hence it creates increasing tension between efficiency and equality.

Not surprisingly, the huge increase in CEO pay has been under public scrutiny especially since the Great Recession. A heated debate on CEO pay limits followed. According to a Gallop Poll survey conducted during the later part of the Great Recession (June, 2009), the majority of Americans (59%) endorsed government action to limit CEO pay. Subsequently, we observed various legislations intended to limit CEO pay, such as “say-on-pay” provisions in the Dodd-Frank Act in the U.S. and hard cap on the CEO compensation in the public sector in Ontario, Canada through the Broader Public Sector Executive Compensation Act.⁴

In the economics literature though, a theoretical analysis of the CEO pay limits impact on welfare and efficiency is less developed. Most recently, Cebon and Hermalin (2015) studies CEO pay limits in a single firm-CEO match in isolation with agency costs related to observable but non-verifiable effort. Because a CEO works for a firm over time, the firm can offer a relational contract that makes the current pay contingent on the CEO’s observable but non-verifiable effort (Baker, et al. (1994), Fong and Li (2015), Levin (2003), Malcomson (2013), Pearce and Stacchetti (1998)). They show that CEO pay limits can lower the

³Mishel and Davis (2014) compute CEO annual compensation by using the “options realized” compensation series, which includes salary, bonus, restricted stock grants, options exercised, and long-term pay outs for CEOs at the top 350 U.S. firms ranked by sales based on Compustat’s ExecuComp database, Current Employment Statistics program, and the Bureau of Economic Analysis NIPA tables.

⁴France planned to cap executive pay in the private sector in 2012, but backed off their plans in the next year. In the 2013 referendum, Swiss voters rejected a proposal to cap the CEO pay at 12 times that of a company’s lowest wage.

future expected profit when the firm has to offer the formal contract after it reneges on its promise to honor the relational contract in the current period. If the shareholders are sufficiently patient, they will always honor the relational contract, which can lead to the first-best outcome.⁵

Our framework is sufficiently tractable to study the overall impact of the CEO pay limit on the executive's incentive to accumulate executive skill prior to being on the CEO market. We show that executive pay limits can lead to efficiency-neutral redistribution of earnings from CEOs to shareholders only when they are sufficiently small, such that the minimum pay cut required for the highest-paid CEO in the market is less than the pay minus the outside option of the lowest-paid CEO. Otherwise, they induce redistribution of earnings only at the low end of firm-CEO matches at the cost of distortion of highly able executives' skill accumulation due to discontinuous bunching at the top end.

Because the lowest paid CEO's pay net of her cost of working at the bottom match cannot go below her outside option, a new market pay function with a pay limit must start at a level as high as the outside option of the CEO in the bottom match. The problem with a severe pay limit is as follows. As the pay increases in the executive's skill, it will hit the pay limit before it reaches the highest match. A top group of talented executives foresee this problem in their current positions. They are discouraged to accumulate executive skill beyond a certain level and this leads to a discontinuous bunching of executive skill at the top.

2 Preliminaries

2.1 Model

There is a continuum of executives who are currently working in non-CEO executive positions across firms. They differ in their inherent ability, called type, denoted by $\theta \in [\underline{\theta}, \bar{\theta}]$. $G(\theta)$ denotes the measure of executives whose types are less than or equal to θ for all $\theta \geq \underline{\theta}$. The total measure of executives is one. G is differentiable with support $[\underline{\theta}, \bar{\theta}]$.

An executive of type θ needs to perform tasks to get paid in the current position. There are two types of tasks. Type-1 task requires firm-specific skill and type-2 task general executive skill (henceforth simply executive skill). Execu-

⁵Jewitt et al. (2008) show how to solve the standard one-shot moral hazard problem with non-observable effort (Mirrlees(1976, 1999), Holmstrom (1979)) when the firm faces the constraints of bounded payments. In this standard one-shot setting, the constraints of bounded payments such as pay limits cannot increase the net surplus between the firms and the manager.

tives exert effort to complete their tasks. We assume that the costs of exerting effort for the two types of tasks are separable. By completing those tasks, the executives get paid from the firms that hire them for the current positions. Executives can exert extra effort associated with type-2 tasks in order to accumulate executive skill beyond the level that is required in the current position. The cost of this extra effort is not necessarily compensated in the current position. It is compensated when the executive is either promoted to a CEO by the current firm or hired as a CEO by another firm. This is because the CEO's overall talent $t(y, \theta)$ is composed of her type $\theta \in [\underline{\theta}, \bar{\theta}]$ but also the executive skill $y \geq y_o$ that she had accumulated beyond the level y_o , required in the previous non-CEO executive position. For simplicity, we normalize $y_o = 0$.

Let us formulate the extra utility that an executive expects by accumulating executive skill. Suppose that an executive becomes a CEO for a firm in the future. She can shirk or work hard to generate positive earnings for the firm. For simplicity, we assume that the contract offered by a firm can induce its CEO to work so we do not explicitly consider the contracting problem between the firm and its CEO, contrary to Cebon and Hermalin (2015) or Jewitt et al. (2008). Let $k > 0$ denote the utility cost of working as a CEO.⁶ Let p be the expected CEO pay conditional on working. The CEO pay is the total monetary value of various forms of compensation such as salary, bonus, restricted stock grants, options, perks, etc. Let δ denote the executive's discount rate. Then, the executive's *extra utility* by accumulating executive skill on the job prior to being on the CEO market is the discounted value of the expected CEO pay net of the cost of working in the CEO position, $\delta(p - k)$, minus the utility cost of accumulating the general executive skill $c(y, \theta)$:

$$U(y, p, \theta) = \delta(p - k) - c(y, \theta). \quad (1)$$

After accumulating the executive skill, executives are implicitly on the CEO market where firms hire them as CEOs. Firms are different in terms of size. Firm x has its size $x \in [\underline{x}, \bar{x}]$ where \underline{x} and \bar{x} are positive numbers. $H(x)$ denotes the measure of firms whose sizes are less than or equal to x for any $x \geq \underline{x}$. The total measure of firms is one. H is differentiable with support $[\underline{x}, \bar{x}]$. If firm x hires an executive with talent t as its CEO, she can generate earnings $f(x, t)$ by working for the firm. Therefore, when an executive with talent t works for

⁶Because the market pay function is characterized by a differential equation, what matters is the cost of working for the least talented CEO, which is the lower bound for the initial condition of the differential equation. For simplicity, we assume that the cost of working as a CEO is k to all executives but we believe that a more general setup for the cost of working as a CEO is possible.

firm x as its CEO, the firm's profit becomes

$$\Pi(x, t, p) = f(x, t) - p.$$

The reservation profit for firms is normalized to zero. We assume a *limited liability* property for both firms and executives.

We make the following assumptions on the earnings function f , the talent function t , and the cost function c . We assume that the talent function $t(y, \theta)$ is increasing in y given any $\theta \neq \underline{\theta}$ and t is increasing in θ given any $y \neq 0$. The lowest possible talent of an executive on the CEO market is then denoted by $\underline{t} = t(\underline{\theta}, 0)$. We also assume that t is differentiable and that it is supermodular: If $y > y'$ and $\theta > \theta'$,

$$t(y, \theta) + t(y', \theta') > t(y, \theta') + t(y', \theta).$$

We also assume the cost function $c(y, \theta)$ is increasing in y given any $\theta > \underline{\theta}$, but decreasing in θ given any $y > 0$ and that $c(y, \underline{\theta}) = \infty$ for all $y > 0$. We assume that c is differentiable and convex in y given any θ . Finally, c is assumed to be submodular: If $y > y'$ and $\theta > \theta'$,

$$c(y, \theta) + c(y', \theta') < c(y, \theta') + c(y', \theta).$$

Finally, we assume that the earnings function $f(x, t)$ is increasing in x given any $t \neq \underline{t}$ and in t given any $x \neq 0$. Note that for any given x and θ , $f(x, t(\cdot, \theta))$ is a function of y . We assume that $f(x, t(\cdot, \theta))$ is concave in y . We also assume that f is differentiable and it is supermodular. We also assume that $f(\underline{x}, \underline{t}) > k$ so that an executive on the CEO market can generate positive earnings net of the utility cost of working as a CEO for any firm.

2.2 Stable matching equilibrium

Before being on the CEO market, executives first accumulate executive skill in their various positions, together with the firm-specific skill. Then, firms hire executives as their CEOs in the market. We formulate this process by one-to-one matching between them through transferring part of the earnings from firms to their CEOs. Because the inherent type is not observable, there may be multiple equilibria, depending on how the market perceives an executive's type conditional on her executive skill. In practice, CEOs have established themselves as talented executives in the market with plenty of observable executive skills. Therefore, it is natural to focus on a *separating stable matching equilibrium* where an executive of a higher type accumulates a higher level of executive skill

in her previous position and the market's perception of the executive's talent conditional on her executive skill is correct

Let $y(\theta)$ be the executive skill function that characterizes the level of executive skill that each CEO accumulates in her previous position as a function of her type θ . The image set of function $y(\cdot)$, denoted by Y , includes all executive skill levels. Let $\mu(y)$ be every firm's belief function on the executive's type, that is, $\mu(y)$ denotes the executive's type that the firm believes conditional on her executive skill y . Given the belief $\mu(\cdot)$ on each executive's type conditional on her executive skill, each firm x can infer the perceived earnings $f(x, t(y, \mu(y)))$ that can be generated by hiring a particular executive as its CEO.

The *stable job matching* takes place in the CEO market if there are no pairs of a firm and a CEO who, by matching and transferring part of the perceived earnings to the CEO, can make themselves strictly better off. Let $x(y)$ be the matching function that specifies the size of the firm for whom the CEO works as an injective function of her executive skill level y , for all, $y \in Y$. Let $p(y)$ be a pay function that specifies an (expected) pay for the CEO conditional on working as a function of her executive skill level y , for all $y \geq 0$. If $\{x(\cdot), p(\cdot)\}$ leads to a stable job matching given $\{y(\cdot), \mu(\cdot)\}$, they are called the market matching function and the market pay function respectively.

Now we formally describe the notion of an equilibrium. First, consider how stable job matching occurs in the market given $\{y(\cdot), \mu(\cdot)\}$. Given $\{y(\cdot), \mu(\cdot)\}$, suppose that firm x considers hiring an executive with y as its CEO. The firm believes that she can work for firm $x(y)$ at the market pay $p(y)$ according to $\{x(\cdot), p(\cdot)\}$. This means that if firm x wants to hire the executive with y as its CEO, it must offer her an expected pay at least as high as the market pay $p(y)$. Therefore firm x will hire the executive with y^* at pay p^* if (y^*, p^*) solves the following problem:

$$\max_{(y,p)} [f(x, t(y, \mu(y))) - p] \text{ subject to } p \geq p(y). \quad (2)$$

Given $\{y(\cdot), \mu(\cdot)\}$ consider the decision of an executive with y . Let $\tilde{y}(x)$ be the executive skill level of the CEO who works for firm x in equilibrium given $x(\cdot)$, that is,

$$x(\tilde{y}(x)) = x.$$

Since firm x believes that $f(\tilde{y}(x), \mu(\tilde{y}(x)), x) - p(\tilde{y}(x))$ is his profit, the CEO with y can work for firm x if she is willing to accept a pay p that satisfies

$$f(x, t(y, \mu(y))) - p \geq f(x, t(\tilde{y}(x), \mu(\tilde{y}(x)))) - p(\tilde{y}(x)).$$

Therefore the CEO with y will work for firm \hat{x} at pay \hat{p} if (\hat{x}, \hat{p}) solves the

following problem:

$$\max_{(x,p)} p \text{ subject to } f(x, t(y, \mu(y)) - p \geq f(x, t(\tilde{y}(x), \mu(\tilde{y}(x)))) - p(\tilde{y}(x)). \quad (3)$$

Given $\{y(\cdot), \mu(\cdot)\}$, $x(\cdot)$ and $p(\cdot)$ are the market matching function and the market pay function if $\{\tilde{y}(x), p(\tilde{y}(x))\}$ is a solution to problem (2) for firm x and $\{x(y), p(y)\}$ is a solution to problem (3) for the CEO with characteristic y . This is because $\{x(\cdot), p(\cdot)\}$ derived by solving problems (2) and (3) ensures that there are no pairs of a firm and a CEO who, by matching and transferring part of the perceived output to the CEO, can make themselves strictly better off. Therefore the CEO with y works for firm $x(y)$ at pay $p(y)$.

Consider an executive's effort exerting decision to accumulate executive skill on the job in her current position prior to being on the CEO market. Given the pay that she is supposed to receive by completing both types of tasks in the current position, we assume that it is optimal for her to complete them by exerting effort. This leaves the executive's extra effort decision to accumulate the executive skill beyond the level that is required in the current position. Given $p(\cdot)$, the executive of type θ solves the following effort exerting problem in the current position to accumulate executive skill:

$$\max_y [\delta(p(y) - k) - c(y, \theta)]. \quad (4)$$

Equation (4) captures the *investment component* of accumulating executive skill y in the current executive position. Now, we define a separating stable matching equilibrium based on the notion of perfect Bayesian equilibrium.

Definition 1 $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ constitutes a separating stable matching equilibrium if

1. if $\theta \neq \theta'$, $y(\theta) \neq y(\theta')$
2. for all θ , $\mu(y(\theta)) = \theta$
3. for all θ , $y(\theta)$ solves the executive's problem (4),
4. Given $\{y(\cdot), \mu(\cdot)\}$, $\{x(\cdot), p(\cdot)\}$ is the pair of the market matching function and the market pay function if $\{\tilde{y}(x), p(\tilde{y}(x))\}$ is a solution to problem (2) for firm x and $\{x(y), p(y)\}$ is a solution to problem (3) for the executive with y .

Part 1 of Definition 1 implies that every executive's skill fully reveals her type. Part 2 shows that the firm's belief on the type of an executive is based on Bayes' rule. Part 3 defines the executive's equilibrium effort exerting decision to acquire executive skill and part 4 defines the market matching function and the market pay function that lead to a stable job matching in the CEO market.

3 Equilibrium Analysis

3.1 Equilibrium characterization

We characterize a separating stable matching equilibrium. Proposition 1 fully characterizes the executive's effort exertion to accumulate executive skill on her non-CEO job prior to being on the CEO market and the subsequent stable matching in the CEO market. Given the set of all executive skill levels Y chosen by executives, let

$$\bar{y} := \max Y.$$

An executive skill level higher than \bar{y} is never observed. Therefore, it is arbitrary to set up the firm's belief and market pay functions beyond \bar{y} . We are interested in a monotone equilibrium in the sense that $\mu(\cdot)$ and $p(\cdot)$ are non-decreasing even beyond \bar{y} .

Proposition 1 *Any (monotone) separating stable matching equilibrium $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ is characterized as follows:*

1. $y(\cdot)$ is a strictly increasing function; $y(\underline{\theta}) = 0$ and, for all $\theta > \underline{\theta}$, the necessary and sufficient condition for an executive of type θ to acquire $y(\theta)$, given the market pay function $p(\cdot)$ specified in Condition 3.(b) and Condition 4, is

$$\delta p'(y) - c_y(y, \theta) = 0 \text{ at } y = y(\theta) \quad (5)$$

2. for all θ , $\mu(y(\theta)) = \theta$ and, for all $y > \bar{y}$, $\mu(y) = \bar{\theta}$.
3. Given $\{y(\cdot), \mu(\cdot)\}$, $\{x(\cdot), p(\cdot)\}$ is the pair of the market matching function and the market pay function if and only if it satisfies (a) and (b) below

$$(a) \text{ for all } \theta, 1 - H(x(y(\theta))) = 1 - G(\theta),$$

$$(b) \text{ for all } y \in Y$$

$$p'(y) = f_t(x(y), t(y, \mu(y)))t_y(y, \mu(y)) + f_t(x(y), t(y, \mu(y)))t_\theta(y, \mu(y))\mu'(y) \quad (6)$$

and $p(0)$ satisfies

$$k \leq p(0) \leq f(\underline{x}, \underline{t}). \quad (7)$$

4. Finally, $p(y)$ for all $y > \bar{y}$ satisfies

$$f(\bar{x}, t(y, \bar{\theta})) - p(y) = f(\bar{x}, t(\bar{y}, \bar{\theta})) - p(\bar{y}) \quad (8)$$

Proof. See on-line appendix. ■

Condition 1 characterizes the executive skill function. An executive makes her effort exerting decision to accumulate executive skill given the correct belief about the market pays conditional on the executive skill, which are characterized by the market pay function $p(\cdot)$. The executive with the lowest type does not exert effort to accumulate extra executive skill to, distinguish herself from the rest. Given the market pay function $p(\cdot)$, the submodularity of c ensures that (5) is not only the necessary condition, but also the sufficient condition for the executive of type $\theta > \underline{\theta}$ to acquire in equilibrium the executive skill $y(\theta)$ through costly effort exertion. The executive skill accumulated is increasing in the executive's type. Condition 2 is about the firm's belief function $\mu(y)$. In a separating equilibrium, the executive skill fully reveals the executive's type and hence we have that, for all θ , $\mu(y(\theta)) = \theta$. Given Condition 1, the firm's belief function is increasing up until \bar{y} and $\mu(\bar{y}) = \bar{\theta}$. Because $\mu(\bar{y}) = \bar{\theta}$, $\mu(y)$ for all $y > \bar{y}$ cannot be lower than $\bar{\theta}$ given the monotonicity. Since $\bar{\theta}$ is the highest type level, it implies that for all $y > \bar{y}$, $\mu(y) = \bar{\theta}$.

Conditions 1 and 2 show how the executive's skill $y(\cdot)$ and the firm's belief $\mu(\cdot)$ are formed given the market pay function that executives expect. Condition 3 shows how matching $x(\cdot)$ and pay formation $p(\cdot)$ take place, given $\{y(\cdot), \mu(\cdot)\}$. For a distribution of executives' skill levels and the firm's belief on the executive's type, market matching and market pays are determined at the point where there are no firm-CEO pairs who, by matching and transferring part of the perceived earnings to the CEO, can make themselves strictly better off. A firm can hire an executive as its CEO if it offers a pay that is at least as high as the existing market pay for her (see the firm's hiring problem (2)). An executive can work for a firm as its CEO if she is willing to accept a pay that ensures the firm's profit as high as the one that the firm earns with the executive that it is supposed to hire in equilibrium (see the executive's working problem (3)). Those conditions for a successful match between a firm and a CEO give us a way to check the property of the market matching function $x(\cdot)$. Given the increasing executive skill and belief functions $\{y(\cdot), \mu(\cdot)\}$, the supermodular assumptions on f and t make the market matching function $x(\cdot)$ increasing in executive skill. Therefore, the equilibrium matching must be positively assortative between the firm size and the CEO's executive skill. Given the increasing executive skill function $y(\cdot)$, it also implies positively assortative matching between the firm size and the CEO's type. Those matching patterns are characterized in Condition 3.(a).

Now consider the market pay determination. The market pay function specifies the return on an executive skill accumulation on the job prior to being on

the CEO market. A marginal increase in an executive's skill y has two effects. The first effect is the marginal increase in a firm's earnings due to the increase in her talent directly through the increase in her executive skill. This is the *direct productivity effect* of executive skill accumulation and the first term of the right-hand-side of (6) captures it:

$$f_t(x(y), t(y, \mu(y)))t_y(y, \mu(y)).$$

The second effect is due to the change in a firm's perception $\mu(y)$ on her type conditional on y . As a firm hires an executive with a slightly higher y , it *believes* that the executive's type is slightly higher. Because an executive's talent is also increasing in her type, there is a marginal increase in a firm's earnings due to an increase in the executive's talent through a marginal increase in her type. This is the *informational effect* of executive skill accumulation and the second term on the right-hand-side of (6) captures it:

$$\dots + f_t(x(y), t(y, \mu(y)))t_\theta(y, \mu(y))\mu'(y).$$

This is the information rent that a CEO can extract even though her type and hence her overall talent is correctly revealed in the CEO market. An increase in the market pay that an executive can receive from a marginal increase in her executive skill should correctly reflect the sum of these two, as in (6).

The CEO at the bottom match does not exert effort to accumulate any extra executive skill prior to being on the CEO market so her talent is $\underline{t} := t(\underline{\theta}, 0)$. Because this CEO can create earnings of $f(\underline{x}, \underline{t})$, her pay $p(0)$ cannot be greater than $f(\underline{x}, \underline{t})$ and less than k . This is reflected in (7). The notion of stability does not pin down the pay at the bottom match. As long as $p(0)$ is somewhere between $f(\underline{x}, \underline{t})$ and k , it can be supported in an equilibrium. The proof in the Appendix shows that items (a) and (b) in Condition 3 are the sufficient and necessary equilibrium conditions for a firm $x(y)$ and the CEO with y to match at pay $p(y)$.

Finally, condition 4 specifies the market pay function for the executive skill beyond the highest level observed in the equilibrium. The right hand side of (8) is the equilibrium profit for the largest firm. Because, for all $y \geq \bar{y}$, $\mu(y) = \bar{\theta}$, (8) shows that $p(y)$ for all $y > \bar{y}$ tracks down the largest firm's iso profit curve associated with hiring the executive of the highest type in the space of y and p . The increasing property of $p(\cdot)$ is preserved even beyond \bar{y} .

3.2 Equilibrium derivation

Let us show how to derive $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ given the characterization of a separating stable matching equilibrium in Proposition 1.

Executive skill function: We start with deriving the executive skill function. Define

$$\tilde{x}(\theta) := x(y(\theta)) \quad (9)$$

as the size of the firm who hires the executive of type θ as its CEO in equilibrium. This is derived by solving $1 - H(\tilde{x}(\theta)) = 1 - G(\theta)$ for $\tilde{x}(\theta)$ at all θ by Condition 3.(a) in Proposition 1. Since every executive fully reveals her type by accumulating the executive skill in equilibrium (i.e., $\mu(y(\theta)) = \theta$), we have $\mu'(y(\theta)) = 1/y'(\theta)$. Given $\tilde{x}(\theta)$ and $\mu'(y(\theta)) = 1/y'(\theta)$, combining (5) and (6) yields

$$\delta f_t(\tilde{x}(\theta), t(y, \theta)) \left(t_y(y, \theta) + \frac{t_\theta(y, \theta)}{y'} \right) - c_y(y, \theta) = 0. \quad (10)$$

This yields a first-order differential equation $y' = \phi(y, \theta)$, where

$$\phi(y, \theta) := \frac{-\delta f_t(\tilde{x}(\theta), t(y, \theta)) t_\theta(y, \theta)}{\delta f_t(\tilde{x}(\theta), t(y, \theta)) t_y(y, \theta) - c_y(y, \theta)} \quad (11)$$

The initial condition is $y(0) = 0$ according to Condition 1. If $\phi(\cdot, \cdot)$ is continuous in θ and Lipschitz continuous in y , then the first-order differential equation $y' = \phi(y, \theta)$ has the unique solution for y given an initial condition, according to the Picard-Lindelof Theorem (See Teschl (2012)).

Firm's belief function: The executive skill function is a strictly increasing function in the range of Y according to Condition 1. We can derive the firm's belief function $\mu(y)$ by deriving the inverse function of the executive skill function in the range of Y , i.e., $\mu(y(\theta)) = \theta$. If $y > \bar{y}$, we set up $\mu(y) = \bar{\theta}$ as stated in Condition 2.

Market matching function: Given the firm's belief function $\mu(y)$, we can derive the market matching function $x(y)$ according to Condition 3.(a), $1 - F(x(y)) = 1 - G(\mu(y))$.

Market pay function: Finally, consider the market pay function. For $y \in Y$, it can be derived by integrating the right-hand-side of (6) with the initial condition $k \leq p(0) \leq f(\underline{x}, \underline{t})$ at the bottom match and, for $y \geq \bar{y}$, it keeps track of the iso profit curve of the largest firm according to (8). Therefore, the market pay function can be specified as follows:

$$p(y) = \begin{cases} \int_0^y f_t(x(s), t(s, \mu(s))) (t_y(s, \mu(s)) + t_\theta(s, \mu(s)) \mu'(s)) ds + p(0) & \text{if } y \leq \bar{y}, \\ f(\bar{x}, t(y, \mu(y))) - f(\bar{x}, t(\bar{y}, \mu(\bar{y}))) + p(\bar{y}) & \text{if } y > \bar{y}. \end{cases} \quad (12)$$

The notion of stability does not pin down the pay $p(0)$ at the bottom match, even when we have the unique executive skill function with the usual Lipschitz condition. Therefore, we have a continuum of separating stable matching equilibria where one market pay function can be derived by shifting another market pay function by the same amount at each possible y . This implies that competing for talented CEOs alone does not pin down the market pays for CEOs. All CEOs will prefer the market pay function with $p(0) = f(\underline{x}, \underline{t})$, but all firms will prefer the one with $p(0) = k$. The market pay function realized in the equilibrium will depend on the bargaining power of each side of the market, which is affected by many different factors.

3.3 Equilibrium in Parametrized Models

This section provides the equilibrium analysis in parametrized models with a broad class of functions that are often used in the literature. An executive with talent t creates extra earnings $\alpha x^\gamma t$ for firm x when she works as a CEO for the firm, compared to what the executive with the lowest talent can ($\alpha > 0$). An executive talent takes the functional form, $t(y, \theta) = \theta y$.⁷

A firm's profit function is given by

$$f(x, t(y, \theta)) = \alpha x^\gamma \theta y + \tau,$$

where $\tau \geq k$ is the earnings that a CEO with the lowest talent ($\underline{t} = 0$ in this case) can create. If $\gamma = 1$, CEO impact on extra earnings exhibits constant returns to scale with respect to firm size. Since Gabaix and Landier (2006) show that $\gamma = 1$ is empirically consistent with data, we assume $\gamma = 1$ but all the qualitative results in this section go through with a general γ .

For simplicity we assume that an executive's discount rate $\delta = 1$. It is reasonable to assume that because we consider that executive skill accumulation by those executives who are very close to being CEOs. To incorporate the convexity and supermodularity of the executive's utility cost function $c(y, \theta)$ for executive skill accumulation, we adopt the commonly-used functional form for $c(y, \theta)$ as follows.

$$c(y, \theta) = \beta \left(\frac{y}{\theta} \right)^m,$$

where $\beta > 0$ and $m > 1$.

⁷We can alternatively think of the talent function as $t(y, \theta) = \theta^\delta y$, where δ measures the returns to scale of the executive skill with respect to his type in talent formation. We parsimoniously assume $\delta = 1$. Alternatively, we can reparametrize θ^δ as the executive's effective type in talent formation. All the qualitative results in this section go through with a general δ .

Finally, we need to consider the distributions of firm size and executive type. In equilibrium, firms and CEOs are matched assortatively in terms of firm size and executive skill. Because the executive skill function is increasing in executive type, firms and CEOs are also matched assortatively in terms of size x and type θ . Then, we can consider the functional form $\tilde{x}(\theta)$ such that

$$1 - H(\tilde{x}(\theta)) = 1 - G(\theta).$$

We assume that $\tilde{x}(\theta)$ follows the following specific form:

$$\tilde{x}(\theta) = k\theta^q, \tag{13}$$

where $k > 0$ and $q > 0$. k is the “shift” parameter and q is the “relative spacing” parameter. Given k , the relative spacing parameter q shows the relative heterogeneity of firm size to the executive’s inherent ability. This functional form can be derived under several reasonable distributions for firm size and inherent ability. For example, assume that the distributions of firm size and inherent ability follow a class of Weibull distributions. Then, we have $1 - G(\theta) = \exp[-(\theta/\lambda_1)^{z_1}]$ and $1 - H(x) = \exp[-(x/\lambda_2)^{z_2}]$. If $k_i = 1$, it is the exponential distribution. If $k_i = 3, 4$, it is close to the normal distribution. In this case, $q = z_1/z_2$ and $k = \lambda_2/\lambda_1^{z_1/z_2}$.

Suppose that the distribution of firm size follows a class of Pareto distributions, so does the distribution of the executive’s inherent ability. Then, we have

$$1 - G(\theta) = (\theta/\theta_m)^{-z_1}, \tag{14}$$

$$1 - H(x) = (x/x_m)^{-z_2}, \tag{15}$$

where θ_m is the mode of the executive’s inherent ability, x_m is the mode of the firm size. Then, $q = z_1/z_2$ and $k = \lambda_2/\lambda_1^{z_1/z_2}$. Following the equilibrium derivation explained in the previous section, we can characterize the separating stable matching equilibrium in the parametrized model as follows.

3.3.1 Observable executive types

We first characterize the equilibrium allocation with complete information about executive types. This is the case where the executive’s inherent ability is publicly observable and the equilibrium allocation is Pareto optimal. This outcome serves our benchmark outcome.

Since there is no private information on executive types, the equilibrium is characterized by $\{y_e(\cdot), p_e(\cdot), x_e(\cdot)\}$. $y_e(\theta)$ is determined according to Condition

1 in Proposition 1. The marginal increase in the market pay $p'_e(y)$ from accumulating executive skills is only attributed to the direct productivity effect of executive skill accumulation due to the absence of its informational effect given that the executive type is publicly observable. Therefore, the marginal increase in the market pay includes only the firm term on the right hand side of (6) but not the second term:

$$p'_e(y) = f_t(x(y), t(y, \theta))t_y(y, \theta). \quad (16)$$

$p_e(\cdot)$ and $x_e(\cdot)$ are determined according to Conditions (3) and (4) after replacing (6) with (16). combining (34) in Condition 1 in Proposition 1 and (16) yields

$$f_t(x(y), t(y, \theta))t_y(y, \theta) - c_y(y, \theta) = 0.$$

Given the fact that $\tilde{x}(\theta) = x(y(\theta)) = k\theta^q$, solving the equation above yields the executive skill function as follows:

$$y_e(\theta) = \left[\frac{\alpha k}{m\beta} \right]^{\frac{1}{m-1}} \theta^{\frac{q+m+1}{m-1}}. \quad (17)$$

The inverse of $y_e(\theta)$ is

$$\mu_e(y) = \left[\frac{m\beta}{\alpha k} \right]^{\frac{1}{q+m+1}} y^{\frac{m-1}{q+m+1}}. \quad (18)$$

From (16), the market pay function is

$$\begin{aligned} p_e(y) - p_e(0) &= \int_0^y f_t(x(s), t(s, \mu_c(s)))t_y(s, \mu_c(s))ds \\ &= \frac{q+m+1}{qm+2m} \alpha k \left[\frac{m\beta}{\alpha k} \right]^{\frac{q+1}{q+m+1}} y^{\frac{qm+2m}{q+m+1}}, \end{aligned} \quad (19)$$

with $k \leq p_e(0) \leq \tau$. As one can see, the amount of market pay for the CEO with executive skill y net of the pay for the CEO with the lowest skill level is fully attributed to the direct productivity effect from executive skill accumulation.

Using $\mu_e(y)$, we can derive the market matching function $x_e(y)$ as

$$x_e(y) = \tilde{x}(\mu_e(y)) = k \left[\frac{m\beta}{\alpha k} \right]^{\frac{q}{q+m+1}} y^{\frac{(m-1)q}{q+m+1}} \quad (20)$$

(17), (19) and (20) completely characterize the stable matching equilibrium with complete information. Since there is no private information, the equilibrium allocation is fully Pareto efficient.

3.3.2 Unobservable executive types

This is our focus in this paper. Following the steps in Section 3.2, the separating stable matching equilibrium with incomplete information on executive types is derived as follows.

Proposition 2 *Given the parametrized model, the separating stable matching equilibrium $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ is given by*

$$\begin{aligned} y(\theta) &= \left[\frac{\alpha k(q+2m)}{m\beta(q+m+1)} \right]^{\frac{1}{m-1}} \theta^{\frac{q+m+1}{m-1}}. \\ \mu(y) &= \left[\frac{m\beta(q+m+1)}{\alpha k(q+2m)} \right]^{\frac{1}{q+m+1}} y^{\frac{m-1}{q+m+1}} \\ x(y) &= \tilde{x}(\mu(y)) = k \left[\frac{m\beta(q+m+1)}{\alpha k(q+2m)} \right]^{\frac{q}{q+m+1}} y^{\frac{(m-1)q}{q+m+1}} \end{aligned}$$

and

$$\begin{aligned} &p(y) - p(0) \\ &= \int_0^y f_t(x(s), t(s, \mu(s))) t_y(s, \mu(s)) ds + \int_0^y f_t(x(s), t(s, \mu(s))) t_\theta(s, \mu(s)) \mu'(s) ds \\ &= \frac{q+m+1}{qm+2m} \alpha k \left[\frac{m\beta(q+m+1)}{\alpha k(q+2m)} \right]^{\frac{q+1}{q+m+1}} y^{\frac{qm+2m}{q+m+1}} + \\ &\quad \frac{m-1}{qm+2m} k \left[\frac{m\beta(q+m+1)}{\alpha k(q+2m)} \right]^{\frac{q+1}{q+m+1}} y^{\frac{qm+2m}{q+m+1}}, \quad (21) \end{aligned}$$

with $k \leq p(0) \leq \tau$.

The first three functions represent the executive skill function, the firm's belief function on the executive's type, and the market matching function. The last function characterizes the market pay function. In characterizing the equilibrium in terms of closed-form solutions, it is most important to derive the executive skill function $y(\cdot)$ from the differential equation (10) (or equivalently (11)). We use the method of try and error, guessing the solution for $y(\theta)$ as $A\theta^z$, and determining A and z in terms of model parameters that satisfy (10). The solutions are

$$A = \left[\frac{\alpha k(q+2m)}{m\beta(q+m+1)} \right]^{\frac{1}{m-1}} \quad \text{and} \quad z = \frac{q+m+1}{m-1}.$$

The market pay for the CEO with executive skill y net of the pay to the CEO with the lowest skill, i.e., $p(y) - p(0)$, consists of the two parts. The first part is the part of the market pay attributed to the *direct productivity effect* of executive skills accumulated prior to being CEOs. The second part is the part of the market pay attributed to the *informational effect* of executive skills that improve the market perception on the executive's type.

Proposition 3 *The proportion of the market pay attributed to the informational effect (direct productivity effect) of executive skills uniformly decreases (increases) in q .*

Proof. From the market pay function in (21), we can derive the proportion of the market pay attributed to the information effect of executive skills is the ratio of the second term on the right-hand side of (21) to the sum of the two terms on the right-hand side:

$$\frac{\text{information rent}}{p(y) - p(0)} = \frac{m - 1}{q + 2m} \quad (22)$$

This is strictly decreasing in q and uniformly converges to 0 as $q \rightarrow \infty$.

On the other hand, the proportion of the market pay attributed to the direct productivity effect is the ratio of the first term on the right-hand side of (24) to the sum of the two terms on the right-hand side:

$$\frac{\text{pay due to direct productivity effect}}{p(y) - p(0)} = \frac{q + m + 1}{q + 2m} \quad (23)$$

Because $m > 1$, this is strictly increasing in q and uniformly converges to 1 as $q \rightarrow \infty$. ■

From our specification for the relation between the firm size and executive type in (13), we can see that a higher q implies that the firm size distribution is more spread to the right, given the executive type distribution. This implies that the size of the firm that any given CEO works for is larger than before. The direct productivity effect of executive skills is related to how much extra the CEO can produce from the current firm who considers hiring her, whereas the informational effect is related to how much big firm she can work for by improving the market's perception on her type. Our comparative statics show that as q increases, the current firm size levels up, the direct productivity effect becomes the major incentive to increase executive skills, but the informational effect becomes secondary and negligible. This has the major impact of the equilibrium outcome as q increases.

As q increases, there are efficiency gains because the separating stable matching equilibrium under incomplete information converges to the stable matching equilibrium that would prevail with complete information. As shown in the proof of Theorem 1, how close it is to the stable matching equilibrium under complete information is positively related to the proportion of the market pay attributed to the direct productivity effect of executive skill accumulation. Therefore, this proportion

$$\mathcal{E} := \frac{q + m + 1}{q + 2m}$$

represents the measure of efficiency and it is strictly increasing in q .⁸

Theorem 1 *As $q \rightarrow \infty$, the separating stable matching equilibrium under incomplete information converges to the stable matching equilibrium under complete information.*

Proof. Consider the two equilibrium, one under incomplete information and the other under complete information. Then, we can see that

$$\begin{aligned} \lim_{q \rightarrow \infty} \frac{y(\theta)}{y_e(\theta)} &= \lim_{q \rightarrow \infty} \left[\frac{q + 2m}{q + m + 1} \right]^{\frac{1}{m-1}} = 1 \\ \lim_{q \rightarrow \infty} \frac{x(y)}{x_e(y)} &= \lim_{q \rightarrow \infty} \left[\frac{q + m + 1}{q + 2m} \right]^{\frac{q}{q+m+1}} = 1 \\ \lim_{q \rightarrow \infty} \frac{p(y) - p(0)}{p_e(y) - p_e(0)} &= \lim_{q \rightarrow \infty} \left[\frac{q + m + 1}{q + 2m} \right]^{\frac{q+1}{q+m+1}} + \lim_{q \rightarrow \infty} \frac{m-1}{\alpha(q+m+1)} \left[\frac{q + m + 1}{q + 2m} \right]^{\frac{q+1}{q+m+1}} = 1 + 0 \\ \lim_{q \rightarrow \infty} \frac{\mu(y)}{\mu_e(y)} &= \lim_{q \rightarrow \infty} \left[\frac{q + m + 1}{q + 2m} \right]^{\frac{1}{q+m+1}} = 1 \end{aligned}$$

The limit results above show the convergence of the separating stable matching equilibrium under incomplete information to the stable matching equilibrium under complete information as $q \rightarrow \infty$. ■

⁸Note that profit and talent functions have parameters γ and δ . We set them equal to one given that $\gamma = 1$ empirically consistent with data and $\delta = 1$ seems reasonable because we can always reparametrize θ^δ as effective type. Such parametric specification does not lose the generality of our qualitative results. In particular, with general γ and δ , the measure of efficiency becomes

$$\mathcal{E} = \frac{\gamma q + m + \delta}{\gamma q + 2m + \delta - 1}$$

This measure also converges to 1 as $q \rightarrow \infty$. The closed-form characterization of the equilibrium with general γ and δ is available upon request.

Our equilibrium analysis in the parametrized model provides the efficiency implication of the spread of the firm size distribution to the right over time. As it is more spread to the right, causing an increase in firm size, efficiency improves because it increases (decreases) the proportion of the market pay attributed to the direct productivity effect (the informational effect) of executive skill accumulation.

3.3.3 Empirical strategy to estimate measure of efficiency

Can we quantitatively identify the measure of efficiency from data? This is very interesting but also challenging. This depends on how to estimate q and m from the empirical work given that the measure of efficiency is a function of them (See 22) and (23)). While executive skills are well observed in the market, they may not be well measured by econometricians. However, the firm size is well defined and measured (e.g., a firm's market value, market capitalization, total assets, etc). Therefore, it may be more interesting to convert the market pay as a function of the size of the firm that the CEO is working for. Let $p^*(x)$ be the market pay from the firm with size x . Because the market matching function $x(y)$ is strictly increasing, it has its inverse function $y(x)$. By using the inverse of the market matching function, we can derive $p^*(x) = p(y(x))$ as follows:

$$p^*(x) - p^*(\underline{x}) = \frac{q + m + 1}{qm + 2m} \alpha k^{\frac{-(qm+m+1)}{m-1}} \left[\frac{\alpha k(q + 2m)}{m\beta(q + m + 1)} \right]^{\frac{1}{m-1}} x^{\frac{qm+2m}{q(m-1)}} + \frac{m - 1}{qm + 2m} \alpha k^{\frac{-(qm+m+1)}{m-1}} \left[\frac{\alpha k(q + 2m)}{m\beta(q + m + 1)} \right]^{\frac{1}{m-1}} x^{\frac{qm+2m}{q(m-1)}}, \quad (24)$$

where the first term on the right hand side is the part of the market pay attributed to the direct productivity effect of executive skills and the second term is the part of the market pay attributed to the informational effect of executive skills.

(13) and (24) are important in estimating q and m . We do not observe executive types θ since they are their inherent ability. Nonetheless, it seems reasonable to assume that the distribution of executive types does not change over time. Then, we pick an executive type distribution, for example, a Pareto distribution with some specific values for relevant parameters and draw sample observations from the distribution that matches the number of CEOs available from the data. The data for firm size is available such as market capitalization. Then, given these data, we can estimate the log-linearized equation (13):

$$\ln \tilde{x}(\theta) = a_0 + a_1 \ln \theta$$

The coefficient a_1 of $\ln \theta$ provides the estimate of q :

$$a_1 = q \tag{25}$$

(24) follows the form of $p^*(x) - p^*(\underline{x}) = K \times x^{\frac{qm+2m}{q(m-1)}}$, where K is constant. Since the data for CEO pay and firm size are available, we can estimate the log-linearized version of (24):

$$\ln(p^*(x) - p^*(\underline{x})) = b_0 + b_1 \ln x$$

The coefficient b_1 of $\ln x$ represents $\frac{qm+2m}{q(m-1)}$:

$$b_1 = \frac{qm + 2m}{q(m - 1)} \tag{26}$$

Given $q = a_1$, we can derive m from (26) as $m = \frac{a_1 b_1}{a_1 b_1 - a_1 - 2}$. Let q_t and m_t be the estimates of q and m at time t . It will determine the proportion of market pay contributed to the information effect of executive skills (and direct productivity effect). We can then check whether the proportion has increased or decreased over time.

4 Policy Analysis: CEO Pay Limits

In contrast to the assignment model adopted in the literature on CEO (Tervio (2008), Gabaix and Landier (2008)), we explicitly formulate *executive skill accumulation* in the non-CEO executive positions under *private information* on the executive's inherent ability. Our parametrized model shows that the spread of the firm size distribution to the right, causing the increase in the firm size, has significant efficiency implications. In particular, as the firm size distribution is more spread to the right, the proportion of the CEO pay attributed to the informational effect (direct productivity effect) of executive skill accumulation uniformly decreases (increases). Since the direct productivity effect dominates the informational effect more and more, the efficiency improves.

As Gabaix and Landier (2008) show, the increase in firm size also results in the dramatic increase in CEO pay over time, which is often criticized due to income inequality it may have caused. Therefore, the spread of the firm size distribution to the right creates increasing tension between efficiency and inequality. This section provides the policy analysis of the impact of pay limits to control CEO pays in the market.

4.1 Moderate Pay Limits

Fix a separating stable matching equilibrium $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ without pay limits. Suppose now that a government imposes a CEO pay limit. Let $\bar{p}_c > 0$ be the maximum pay that a firm can pay to its CEO in any match. Recall that Y denotes the image set of the executive skill function $y(\cdot)$. Assume that this is a compact set. Given the original equilibrium $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$, we can derive the maximum pay \bar{p} that the market pay function $p(\cdot)$ induces:

$$\bar{p} := \max_{y \in Y} p(y) = p(\bar{y}).$$

If $\bar{p}_c \geq \bar{p}$, then the pay limit is not binding and it is the same as no pay limit, so we focus only on the case where $\bar{p}_c < \bar{p}$. In this case, a group of top talented executives cannot expect the market pays that they would have been paid when they are hired as CEOs in the original equilibrium. Let $\Delta_c := \bar{p} - \bar{p}_c$. This is the *minimum pay cut required for the highest paid CEO at the top match*. There are two possible situations: either $\Delta_c \leq p(0) - k$ or $\Delta_c > p(0) - k$. If the minimum pay cut required by a pay limit for the highest paid CEO in the market is no greater than the pay for the lowest paid CEO, net of her cost of working as a CEO (i.e., $\Delta_c \leq p(0) - k$), then we call it a moderate pay limit.⁹ The other case we call a severe pay limit.

We start with a moderate pay limit ($\Delta_c \leq p(0) - k$). Given the market pay function $p(\cdot)$ in the original equilibrium $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$, let us define a pay function $p_c(\cdot)$ as follows: for all y ,

$$p_c(y) := \min[p(y) - \Delta, \bar{p}_c]. \quad (27)$$

where Δ is some positive constant in $[\Delta_c, p(0) - k]$. Then, we can characterize the new separating stable matching equilibrium as follows.

Theorem 2 *Let $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ be the original separating stable matching equilibrium with no pay limit. If $\Delta \in [\Delta_c, p(0) - k]$, then $\{y(\cdot), \mu(\cdot), x(\cdot), p_c(\cdot)\}$ is a new separating stable matching equilibrium with pay limit*

Proof. Because $\Delta_c \leq p(0) - k$, we have

$$k \leq p(0) - \Delta \leq f(\underline{x}, \underline{t}), \quad (28)$$

⁹Note that the executive's reservation utility is normalized to zero. This means that her outside option is zero. Generally, we should interpret $p(0) - k$ here as the expected pay net of the cost of working as a CEO at the bottom match minus outside option of the lowest-paid CEO in the market.

for all $\Delta \in [\Delta_c, p(0) - k]$. This also means that under the new market pay function defined in (27), the firm's profit at the bottom pair is non-negative and the executive's pay without exerting effort to accumulate executive skill is non-negative.

Because $\Delta \geq \Delta_c = \bar{p} - \bar{p}_c$, we have that $\bar{p} - \Delta \leq \bar{p}_c$. Because $\bar{p} = p(\bar{y})$, it implies that $p_c(y) := p(y) - \Delta$ for all $y \in Y$ satisfies

$$k \leq p_c(y) \leq \bar{p}_c \text{ for all } y \in Y.$$

The shape of the new market pay function $p_c(\cdot)$ is the same as that of $p(\cdot)$ in the range of Y , but it is shifted by Δ at all $y \in Y$. Therefore, if every executive acquires her executive skill that she would have in the original equilibrium, then

$$y(\underline{\theta}) = 0, \tag{29}$$

and, for all $\theta > \underline{\theta}$,

$$\delta p'_c(y) - c_y(y, \theta) = \delta p'(y) - c_y(y, \theta) = 0 \text{ at } y = y(\theta), \tag{30}$$

where the first equality is due to $p'_c(y) = p'(y)$ for all $y \in Y$. Equations (29) and (30) are identical to those in Condition 1 of Proposition 1. Furthermore, for all $y > Y$,

$$p_c(y) := \min[p(y) - \Delta, \bar{p}_c] \leq p(y),$$

which implies that if an executive chooses $y > \bar{y}$, the corresponding market pay is no higher than what she could have earned in the original equilibrium. It means that if an executive has no incentives to increase her executive skill level above \bar{y} given $p(\cdot)$ in the original equilibrium, she also has no incentives given $p_c(\cdot)$. Therefore, given $p_c(y)$, $y(\cdot)$ is also the executive's skill level in the new separating equilibrium with a pay limit.

Given the original equilibrium $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$, let us replace only $p(\cdot)$ with $p_c(\cdot)$ to get a collection of mappings $\{y(\cdot), \mu(\cdot), x(\cdot), p_c(\cdot)\}$. Applying the proof of Proposition 1, we can show that $\mu(\cdot)$ is the firm's belief function given the monotonicity imposed on the firm's belief and that $\{x(\cdot), p_c(\cdot)\}$ characterizes stable job matching given $\{y(\cdot), \mu(\cdot)\}$. Therefore, $\{y(\cdot), \mu(\cdot), x(\cdot), p_c(\cdot)\}$ is a separating stable matching equilibrium with a pay limit. ■

If the minimum pay cut Δ_c required for the highest paid CEO is smaller than $p(0) - k$, the pay for the lowest paid CEO net of her cost of working as a CEO at the bottom match, then the pay limit can shift the whole market pay function down by $\Delta \in [\Delta_c, p(0) - k]$ up to \bar{y} because $p(\bar{y}) - \Delta \leq \bar{p}_c$: the new market pay function $p_c(\cdot)$ simply takes the minimum between $p(y) - \Delta$ and \bar{p}_c after \bar{y} .

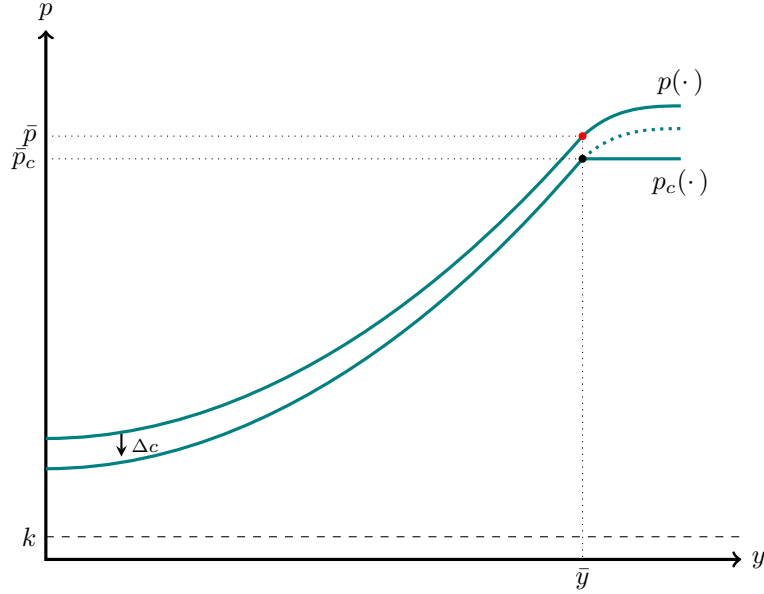


Figure 1: Market pay function with a moderate pay limit

Figure 1 shows the new market pay function with $\Delta = \Delta_c$. The new market pay function preserves the derivative of the original market pay function up until y hits \bar{y} . As y increases beyond \bar{y} , the new market pay cannot be higher than the original pay for each $y > \bar{y}$. Therefore, every CEO's executive skill will be the same as before.

Shifting down the pay function in this way also does not change the pattern of the assortative matching $x(\cdot)$ because the assortativeness comes from the super-modular property of the earnings and talent functions given $\{y(\cdot), \mu(\cdot)\}$. Finally, the marginal condition of the new market pay function for stable matching in (6) with the new initial condition (28) is confirmed because the new market pay function has the same shape as the original market pay function.

It shows that imposing a moderate pay limit *redistributes the match surplus* from the CEO to shareholders at least by Δ_c in a new separating stable matching equilibrium. Because the matching pattern and the executive's effort exerting decision to accumulate executive skill prior to being on the CEO market, do not change, the sum of the total surplus in the economy does not change with a moderate pay limit. A moderate pay limit is *efficiency neutral*; it changes nothing except for the match surplus redistributed by the same amount in all matches.

4.2 Severe Pay Limits

Let $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ be the original separating stable matching equilibrium with no pay limit. Suppose that the pay limit \bar{p}_c is severe so that the minimum pay cut required for the highest paid CEO in the market is greater than the pay for the lowest paid CEO net of her cost of working as a CEO (i.e., $\Delta_c = \bar{p} - \bar{p}_c > p(0) - k$). In this case, there is no separating stable matching equilibrium.

Lemma 1 *If the pay limit \bar{p}_c satisfies $\Delta_c = \bar{p} - \bar{p}_c > p(0) - k$, then there is no separating stable matching equilibrium*

Proof. Any separating stable matching equilibrium shares the same $y(\cdot), \mu(\cdot)$, and $x(\cdot)$ according to Proposition 1. The only difference in any two separating stable matching equilibria is that their market matching functions have the different initial conditions because (7) in Condition 3.(b) of Proposition 1 does not pin down the pay at the bottom match. Given the original separating stable matching equilibrium $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ with no pay limit, suppose that \bar{p}_c satisfies $\Delta_c = \bar{p} - \bar{p}_c > p(0) - k$. Consider a candidate for the market pay function $p_c(\cdot)$ with the pay limit. Let it satisfy

$$\max_{y \in Y} p_c(y) = \bar{p}_c. \quad (31)$$

Since one market pay function can be derived by shifting another one by a constant over the image set Y of the executive skill function $y(\cdot)$, (31) implies that $p_c(\cdot)$ is derived by shifting the original market pay function by $\Delta_c = \bar{p} - \bar{p}_c$ over the image set Y , that is, for all $y \in Y$,

$$p_c(y) = p(y) - \Delta_c. \quad (32)$$

Then, $p_c(0) = p(0) - \Delta_c < k$. Because pay cannot be lower than the utility cost of working as a CEO, $p_c(\cdot)$ satisfying (32) cannot be a market pay function.

Consider $p_c(\cdot)$ that satisfies $p_c(0) = k$. Because one market function can be derived by shifting another one by a constant over the image set Y , it implies that $p_c(\cdot)$ is derived by shifting the original market pay function by $p(0) - k$ over the image set. That is, if for all $y \in Y$,

$$p_c(y) = p(y) - (p(0) - k), \quad (33)$$

then, $\max_{y \in Y} p_c(y) = \bar{p} - (p(0) - k) > \bar{p}_c$. Since the pay for an executive with the highest executive skill level in the market exceeds the pay limit \bar{p}_c , $p_c(\cdot)$ satisfying (33) cannot be a market pay function. Therefore, there does not exist a market pay function $p_c(\cdot)$ that satisfies, for all $y \in Y$,

$$k \leq p_c(y) \leq \bar{p}_c.$$

Because Y is the image set of $y(\cdot)$, it means that there is no separating stable matching equilibrium if \bar{p}_c satisfies $\Delta_c = \bar{p} - \bar{p}_c > p(0) - k$. ■

Lemma 1 implies that if the minimum pay cut required for the highest paid CEO in the market is greater than the pay for the lowest paid CEO net of her cost of working as a CEO (i.e., $\Delta_c = \bar{p} - \bar{p}_c > p(0) - k$), it is necessary that there is bunching, in the sense that a positive measure of executives have the same executive skill level. Given the original separating stable matching equilibrium $\{y(\cdot), \mu(\cdot), x(\cdot), p(\cdot)\}$ with no pay limit, suppose that the market pay function is shifted down by $\Delta \in [0, p(0) - k]$ over the image set Y . If $\Delta = 0$, then the market pay function is not shifted at all and the CEO with the lowest executive skill level will have the same pay that she would have received in the original equilibrium. If $\Delta = p(0) - k$, the pay for the CEO with the lowest executive skill level is simply her cost of working as a CEO. For any given $\Delta \in [0, p(0) - k]$, the top talented CEOs' pays are too high, so that their pays are cut to \bar{p}_c . This induces bunching at the top in the sense that executives of type θ_Δ or higher all have the same executive skill level y_Δ .

As shown later, the executive skill function is non-decreasing with a discrete jump at y_Δ . This means that there is a range of the executive skill levels, below y_Δ , that are never observed in equilibrium. Because Bayesian updating cannot be applied to this range, the firm's belief and market pay functions can be arbitrary in this range. We are interested in a *monotone stable matching equilibrium* in which the firm's belief and market pay functions are non-decreasing, on and off the equilibrium path. Let $\{y_c(\cdot), \mu_c(\cdot), x_c(\cdot), p_c(\cdot)\}$ be a monotone stable matching equilibrium with a severe pay limit. Then, θ_Δ and y_Δ satisfy the following condition for the executive skill function $y_c(\cdot)$:

$$y_c(\theta) = y_\Delta \text{ for all } \theta \geq \theta_\Delta.$$

Recall that $y(\cdot)$ is the executive skill function in the original equilibrium. Therefore, $y(\theta_\Delta)$ denotes the executive skill that an executive of type θ_Δ accumulates in the original equilibrium. Before the executive's skill and her type reach $y(\theta_\Delta)$ and θ_Δ respectively, the executive skill function, the firm's belief function, and the market matching function are shown to be identical to those in the original equilibrium, but the market pay function is shifted downward by $\Delta \in [0, p(0) - k]$: $p_c(y) = p(y) - \Delta$ if $y < y(\theta_\Delta)$.

Because CEOs cannot get paid higher than the pay limit \bar{p}_c , the executive skill accumulated by those executives of type θ_Δ or higher are bunched at the same level y_Δ . That group of CEOs will receive the same pay \bar{p}_c . We first show the determination of θ_Δ and y_Δ in a monotone stable matching equilibrium with $\Delta \in [0, p(0) - k]$.

Lemma 2 *In a monotone stable matching equilibrium with bunching, θ_Δ and y_Δ must satisfy the following equations:*

$$\delta(\bar{p}_c - k) - c(y_\Delta, \theta_\Delta) = \delta(p(y(\theta_\Delta)) - \Delta - k) - c(y(\theta_\Delta), \theta_\Delta) \quad (34)$$

$$\mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c = f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) - (p(y(\theta_\Delta)) - \Delta), \quad (35)$$

where $\mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta]$ is the expected profit for the firm of size $\tilde{x}(\theta_\Delta)$, conditional on hiring the executive with y_Δ as its CEO. Furthermore, we have that

$$p(y(\theta_\Delta)) - \Delta < \bar{p}_c < \bar{p} \text{ and } y(\theta_\Delta) < y_\Delta < \bar{y}.$$

Proof. Note that $p(\cdot)$ and $y(\cdot)$ are the market pay function and the executive skill function in the original equilibrium and that $\tilde{x}(\cdot)$ is defined as $x(y(\cdot))$ in the original equilibrium according to (9). Then $\tilde{x}(\theta_\Delta) = x(y(\theta_\Delta))$ denotes the size of the firm that hires the executive of type θ_Δ as its CEO in the original equilibrium with no pay limit, given the assortativeness of stable matching in terms of firm size and CEO's type.

Suppose that (34) is not satisfied. If the right-hand-side is greater than the left-hand-side, the executive of type θ_Δ or $\theta_\Delta + \epsilon$, for small $\epsilon > 0$, can increase her payoff by decreasing her executive skill to $y(\theta_\Delta)$. If the left-hand-side is greater, the executive of type $\theta_\Delta - \epsilon$, for small $\epsilon > 0$, can increase her payoff by increasing her executive skill to y_Δ . Therefore, (34) must be satisfied in the equilibrium with bunching.

Suppose that (35) is not satisfied. If the right-hand-side is greater than the left-hand-side, then a firm with size $\tilde{x}(\theta_\Delta)$ or $\tilde{x}(\theta_\Delta) + \epsilon$, for small $\epsilon > 0$, can increase its profit by hiring an executive with $y(\theta_\Delta)$ as its CEO by offering the market pay $p_c(y(\theta_\Delta)) = p(y(\theta_\Delta)) - \Delta$. If the left-hand-side is greater, then a firm with size $\tilde{x}(\theta_\Delta) - \epsilon$, for small $\epsilon > 0$, can increase its profit by hiring an executive with y_Δ by offering the maximum pay allowed \bar{p}_c . Therefore, (35) must be satisfied.

Because $\bar{p}_c \geq p(y(\theta_\Delta)) - \Delta$, (34) implies $y_\Delta \geq y(\theta_\Delta)$. Because $y_\Delta \geq y(\theta_\Delta)$, we have

$$t(y_\Delta, \theta) \geq t(y(\theta_\Delta), \theta_\Delta) \text{ for all } \theta \geq \theta_\Delta, \text{ with strict inequality for all } \theta > \theta_\Delta. \quad (36)$$

(36) implies that

$$\mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] > f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) \quad (37)$$

(35) and (37) imply that $\bar{p}_c > p(y(\theta_\Delta)) - \Delta$. Of course $\bar{p}_c < \bar{p}$.

Because $\bar{p}_c > p(y(\theta_\Delta)) - \Delta$, we have that $y_\Delta > y(\theta_\Delta)$ from (34). To show that $y_\Delta < \bar{y}$, note that in the original equilibrium without pay limit, we have

$$\begin{aligned} & \delta(p(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta_\Delta) > \delta(p(\bar{y}) - k) - c(\bar{y}, \theta_\Delta) \\ \iff & \delta(p(y(\theta_\Delta)) - \Delta - k) - c(y(\theta_\Delta), \theta_\Delta) > \delta(p(\bar{y}) - \Delta - k) - c(\bar{y}, \theta_\Delta). \end{aligned} \quad (38)$$

Because $\bar{p}_c > p(y(\theta_\Delta)) - \Delta$ and $p(\bar{y}) > p(y(\theta_\Delta))$, (38) implies that

$$\bar{p}_c - c(y(\theta_\Delta), \theta_\Delta) > p(y(\theta_\Delta)) - \Delta - c(\bar{y}, \theta_\Delta) \quad (39)$$

(34) and (39) imply that $y_\Delta < \bar{y}$. ■

θ_Δ and y_Δ can be derived by solving the two equations, (34) and (35), because those are the only unknowns in the two equations. Because $y_\Delta > y(\theta_\Delta) > y(\theta) = y_c(\theta)$ for all $\theta < \theta_\Delta$ and $\bar{p}_c > p(y_\Delta) - \Delta > p(y) - \Delta = p_c(y)$ for all $y < y(\theta_\Delta)$, the severe pay limit generates an exclusive group of top talented CEOs with the same executive skill y_Δ , which is strictly higher than the executive skill level of any other CEO. Because they all have the same skill level, they receive the same pay level \bar{p}_c , which is the maximum pay allowed under the pay limit.

The set of skills that executives acquire in equilibrium is $[0, y(\theta_\Delta))$ and y_Δ so that any executive skill level y in $[y(\theta_\Delta), y_\Delta)$ or greater than y_Δ is never observed in equilibrium. Therefore, we have to properly define the firm's belief function and the market pay function off the equilibrium path. We completely characterize a monotone stable matching equilibrium with bunching.

Theorem 3 *A monotone stable equilibrium $\{y_c(\cdot), \mu_c(\cdot), x_c(\cdot), p_c(\cdot)\}$ with bunching is characterized as follows:*

1. *The executive skill function $y_c(\cdot)$: $y_c(\theta) = 0$ at $\theta = \underline{\theta}$, and $y_c(\theta)$ for all $\theta \in (\underline{\theta}, \theta_\Delta)$ satisfies*

$$\delta p'_c(y_c(\theta)) - c_y(y_c(\theta), \theta) = 0 \quad (40)$$

and $y_c(\theta) = y_\Delta$ for all $\theta \geq \theta_\Delta$

2. *The firm's belief function, $\mu_c(\cdot)$:*

$$\mu_c(y) = \begin{cases} \theta & \text{if } y = y_c(\theta) < y(\theta_\Delta) \\ \theta_\Delta & \text{if } y(\theta_\Delta) \leq y < y_\Delta \\ G(\theta | \theta \geq \theta_\Delta) & y = y_\Delta \\ \bar{\theta} & y > y_\Delta \end{cases} \quad (41)$$

3. The market matching correspondence $x_c(\cdot)$: $x_c(y)$ is a singleton for $y < y(\theta_\Delta)$ such that

$$H(x_c(y)) = G(\theta) \text{ with } y = y_c(\theta) \text{ for all } y < y(\theta_\Delta) \quad (42)$$

and $x_c(y_\Delta) = \{x \geq \tilde{x}(\theta_\Delta)\}$.

4. The market pay function $p_c(\cdot)$: $p_c(0) = p(0) - \Delta$ with $\Delta \in [0, p(0) - k]$ and, for all $y < y(\theta_\Delta)$, $p_c(y)$ satisfies

$$p'_c(y) = f_t(x_c(y), t(y, \mu_c(y))) (t_y(y, \mu_c(y)) + t_\theta(y, \mu_c(y))\mu'_c(y)), \quad (43)$$

and for all $y \in [y(\theta_\Delta), y_\Delta)$, $p_c(y)$ satisfies

$$f(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta)) - p_c(y) = \mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c \quad (44)$$

and for $y \geq y_\Delta$, $p_c(y) = \bar{p}_c$.

Proof. See Appendix ■

In a monotone equilibrium, the firm's belief and market pay functions are non-decreasing on and off the equilibrium path. As shown in the above proof, an executive of a type less than θ_Δ fully reveals her type by accumulating the same executive skill that she would have accumulated in the original equilibrium with no pay limit. Therefore, the firm's belief on an executive's type is θ if $y = y_c(\theta) < y(\theta_\Delta)$. If y belongs to $[y(\theta_\Delta), y_\Delta)$, then it is a level of the executive skill that is not supposed to be observed in equilibrium. In this case, the firm believes that an executive's type is θ_Δ . Therefore, $\mu_c(\cdot)$ is non-decreasing as y increases. If $y = y_\Delta$, then firm believes, according to Bayes' rule, that an executive's type θ follows the probability distribution $G(\theta | \theta \geq \theta_\Delta)$, conditional on $\theta \geq \theta_\Delta$ because every executive of type θ_Δ or higher accumulates the same executive skill level y_Δ . Because the support of $G(\cdot | \theta \geq \theta_\Delta)$ is not less than θ_Δ , $\mu_c(\cdot)$ again satisfies the non-decreasing property. Finally, if $y > y_\Delta$, then this is again a level of the executive skill off the equilibrium. In this case, the firm believes that an executive's type is the highest possible type $\bar{\theta}$. Therefore, the firm's belief function in Condition 2 is non-decreasing.

Condition 3 shows the equilibrium matching pattern. (42) shows that one-to-one positive assortative matching takes place in terms of firm size and CEO's executive skill level; from the bottom match up to the matches that involve CEOs with executive skill levels lower than $y(\theta_\Delta)$, which is the skill that the executive of type θ_Δ would have accumulated in the original equilibrium with no pay limit. There is no executive with skill in $[y(\theta_\Delta), y_\Delta)$ but there is a positive measure of executives with y_Δ . According to Condition 3, those executives

are hired, as CEOs, by firms with size no less than $\tilde{x}(\theta_\Delta) = x(y(\theta_\Delta))$, which, according to the definition of $\tilde{x}(\cdot)$ in (9), is the size of the firm that would have hired the executive with $y(\theta_\Delta)$ in the original equilibrium.

It is critical to understand the formation of the market pay function in a monotone stable matching equilibrium. If executive skill y belongs to $[y(\theta_\Delta), y_\Delta)$, then pay $p_c(y)$ satisfies (44). Note that if a firm sees an executive with $y \in [y(\theta_\Delta), y_\Delta)$, then according to (41) it believes the executive's type is $\mu_c(y) = \theta_\Delta$. Then, for any $y \in [y(\theta_\Delta), y_\Delta)$, a firm infers that the executive's talent is $t(y, \theta_\Delta)$, given its belief on the executive's type. Suppose that amongst all firms hiring an executive with y_Δ , the firm with the smallest size $\tilde{x}(\theta_\Delta)$ hires an executive with $y \in [y(\theta_\Delta), y_\Delta)$ as its CEO. If it hires such an executive, it believes that its earnings will be $f(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta))$. For all $y \in [y(\theta_\Delta), y_\Delta)$, pay $p_c(y)$ in (44) shows that it is just enough to make the firm with size $\tilde{x}(\theta_\Delta)$ believe that its profit associated with hiring an executive with y , as its CEO, is the same as the profit it would get by hiring an executive with y_Δ at pay \bar{p}_c . That is, in the range of $[y(\theta_\Delta), y_\Delta)$, the market pay function reflects the iso profit curve for the firm with size $\tilde{x}(\theta_\Delta)$ associated with hiring the executive of type θ_Δ who has hypothetically an alternative skill level in the range. We can then show the following corollary on the market pay function.

Corollary 1 *Given $\Delta \in [0, p(0) - k]$, the market pay function specified in Condition 4 of Theorem 3 is nondecreasing with a kink at $y(\theta_\Delta)$ and continuous everywhere except y_Δ :*

$$p_c(y) = \begin{cases} p(y) - \Delta & \text{if } y < y(\theta_\Delta), \\ f(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta)) - \mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] + \bar{p}_c & \text{if } y \in [y(\theta_\Delta), y_\Delta), \\ \bar{p}_c & \text{if } y \geq y_\Delta. \end{cases}$$

Proof. Combining (40) and (43) with the initial condition $y_c(\underline{\theta}) = 0$ yields the same differential equation as (10) for y , before it hits $y(\theta_\Delta)$. Therefore, $y_c(\theta) = y(\theta)$ for all $\theta \in [\underline{\theta}, \theta_\Delta)$. Because of (41), (42), and $y_c(\theta) = y(\theta)$ for all $\theta \in [\underline{\theta}, \theta_\Delta)$, the right-hand-side of (43) is the same as the right-hand-side of (6) before y hits $y(\theta_\Delta)$. Applying the initial condition $p_c(0) = p(0) - \Delta$, we have that for all $y \in [0, y(\theta_\Delta))$,

$$\begin{aligned} p_c(y) &= \int_0^y f_t(x_c(s), t(s, \mu_c(s))) (t_y(s, \mu_c(s)) + t_\theta(s, \mu_c(s))\mu'_c(s)) ds + p_c(0) \\ &= \int_0^y f_t(x(s), t(s, \mu(s))) (t_y(s, \mu(s)) + t_\theta(s, \mu(s))\mu'(s)) ds + p_c(0) \quad (45) \\ &= p(y) - \Delta, \end{aligned}$$

where the first equality comes from (43) with the initial condition $p_c(0)$, the second equality comes from the fact that the right hand side of (43) is the same as the right-hand-side of (6) before y hits $y(\theta_\Delta)$, and the last equality is due to the first part of the right-hand-side of (12).

How about the pay $p_c(y(\theta_\Delta))$, the pay in the new equilibrium for the executive who would have accumulated $y(\theta_\Delta)$ in the original equilibrium? Applying $y = y(\theta_\Delta)$ to (44) yields

$$\begin{aligned} p_c(y(\theta_\Delta)) &= \bar{p}_c + f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) - \mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] \\ &= p(y(\theta_\Delta)) - \Delta, \end{aligned} \quad (46)$$

where the second equality is due to (35). (45) and (46) imply that $p_c(y)$ has its left-hand limit and right-hand limit and they are the same as $p_c(y(\theta_\Delta))$:

$$\lim_{y \nearrow y(\theta_\Delta)} p_c(y) = p_c(y(\theta_\Delta)) = \lim_{y \searrow y(\theta_\Delta)} p_c(y). \quad (47)$$

Therefore, $p_c(\cdot)$ is continuous at $y = y(\theta_\Delta)$. It implies that $p_c(\cdot)$ is continuous at all $y < y_\Delta$ because $p(\cdot)$ is continuous at all $y < y(\theta_\Delta)$ and $f(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta))$ is continuous at all $y \in [y(\theta_\Delta), y_\Delta)$.

However, it is discontinuous at $y = y_\Delta$. To see this point, note that

$$\lim_{y \nearrow y_\Delta} p_c(y) = f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta_\Delta)) - \mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] + \bar{p}_c < \bar{p}_c,$$

because $f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta_\Delta)) < \mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta]$. Therefore, there is a jump at $y = y_\Delta$.

To see why the market pay function has a kink at $y = y(\theta_\Delta)$, note that the left-hand derivative of $p_c(y)$ at $y = y(\theta_\Delta)$ is

$$p'_{c-}(y) = f_t(x(y), t(y, \mu(y))) (t_y(y, \mu(y)) + t_\theta(y, \mu(y))\mu'(y)) \text{ with } x(y) = \tilde{x}(\theta_\Delta).$$

On the other hand, the right-hand limit of $p_c(y)$ at $y = y(\theta_\Delta)$ is

$$p'_{c+}(y) = f_t(x(y), t(y, \mu(y)))t_y(y, \mu(y)) \text{ with } x(y) = \tilde{x}(\theta_\Delta).$$

Because $p'_{c-}(y) > p'_{c+}(y)$ at $y = y(\theta_\Delta)$, $p_c(\cdot)$ has a kink at $y = y(\theta_\Delta)$ and it is getting flatter as it passes $y = y(\theta_\Delta)$. ■

Figure 2 shows the market pay function for the case with $\Delta = p(0) - k$. The market pay function up to $y(\theta_\Delta)$ has the identical shape as the original market pay function, but it is simply shifted downward by Δ . In the range of $[y(\theta_\Delta), y_\Delta)$, the market pay function keeps track of the iso profit curve for the firm with size $\tilde{x}(\theta_\Delta)$ associated with hiring the executive of type θ_Δ who has hypothetically an

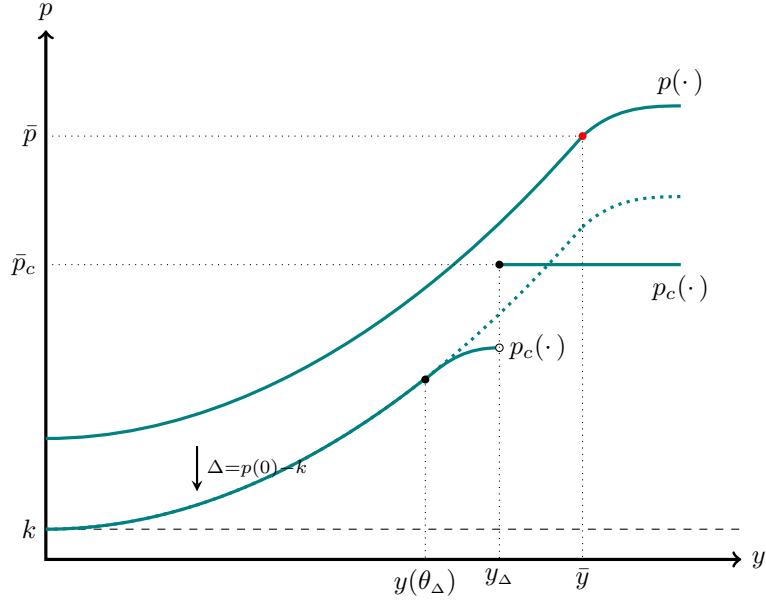


Figure 2: Market pay function with a severe pay limit

alternative skill level y in the range. Because we fix the executive's type to θ_Δ , a change in the pay due to a marginal change in y in this range is given by

$$p'_c(y) = f_t(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta)) t_y(y, \theta_\Delta).$$

It only reflects the change in earnings due to a change in the talent, given by a marginal change in y , but *no information rent*, i.e., the change in earnings due to a change in the talent given by a change in the perception on the executive's type. The reason is that the firm believes that the executive's type is θ_Δ as long as her executive skill belongs to $[y(\theta_\Delta), y_\Delta)$.

While the Appendix includes the full proof of Theorem 3 to show that no one has an incentive to part from the equilibrium matching partner, it is helpful to use Figure 2 to explain what happens at the high end of firm-CEO matches. The executive of type θ_Δ is indifferent between $(p_c(y(\theta_\Delta)), y(\theta_\Delta))$ and (\bar{p}_c, y_Δ) according to (34) in Lemma 2, but acquires y_Δ for \bar{p}_c . Therefore, her utility is indeed equal to

$$\delta p_c(y(\theta_\Delta)) - c(y(\theta_\Delta), \theta_\Delta). \quad (48)$$

Even if the market pay function $p_c(\cdot)$ had the same shape as the original market pay function in the range of $[y(\theta_\Delta), y_\Delta)$, the executive of type θ_Δ could not have a higher utility level than (48) by choosing an alternative skill level in this range. However, the market pay function is even flatter due to no information rent in

this range. It implies that the executive of type θ_Δ has clearly no incentive to choose y in $[y(\theta_\Delta), y_\Delta)$.

Because the executive of type θ_Δ is indifferent between $(p_c(y(\theta_\Delta)), y(\theta_\Delta))$ and (\bar{p}_c, y_Δ) , any executive of a type higher than θ_Δ prefers (\bar{p}_c, y_Δ) to $(p_c(y(\theta_\Delta)), y(\theta_\Delta))$, given the submodular property of the utility cost c of exerting effort to accumulate executive skill. Because the executive of type θ_Δ has no incentive to choose y in $[y(\theta_\Delta), y_\Delta)$, we can also apply the submodular property of c to show that any executive of a type higher than θ_Δ also does not want to choose any executive skill in the range of $[y(\theta_\Delta), y_\Delta)$. Therefore, they all choose to acquire y_Δ for \bar{p}_c . Of course, no executive will acquire executive skill above y_Δ because it incurs a higher utility cost, but cannot yield a pay higher than \bar{p}_c . Therefore, all executives of type θ_Δ acquire the same executive skill y_Δ . In Figure 2, the portion of the market pay function for y greater than $y(\theta_\Delta)$, except y_Δ , is off the path so that it is never realized in equilibrium.

The firm with size $\tilde{x}(\theta_\Delta)$ is indifferent between hiring an executive with y_Δ at pay \bar{p}_c and hiring an executive with $y(\theta_\Delta)$ at pay $p_c(y(\theta_\Delta))$ according to (35) in Lemma 2. Because there is no executive with a skill level that is higher than y_Δ or in the range of $(y(\theta_\Delta), y_\Delta)$, the only possible deviation for the firm is to hire an executive with a skill level that is lower than $y(\theta_\Delta)$. Because the market pay function is shifted downward by the same amount in the range $[0, y(\theta_\Delta)]$, we can show that the firm of size $\tilde{x}(\theta_\Delta)$ has no incentive to hire an executive with skill lower than $y(\theta_\Delta)$ as its CEO. Then, we can use the supermodular property of the firm's earnings function f to show that any firm with a bigger size than $\tilde{x}(\theta_\Delta)$ has also no incentive to hire an executive with skill lower than $y(\theta_\Delta)$.

In those matches that involve CEOs of type lower than θ_Δ , there is redistribution of earnings from CEOs to shareholders by the same amount Δ and there is no distortion of the executive's effort decision on executive skill accumulation prior to being on the CEO market because their original effort decisions are preserved. However, in those matches that involve CEOs of type θ_Δ or higher, there is distortion of executive's effort decisions on executive skill accumulation. In this group of executives, y_Δ is lower than the levels of the skills that executives of higher types would have acquired with no pay limit, but higher than the levels of skills that executives of lower types would have acquired with no pay limit. Therefore, a severe pay limit (i.e., $\Delta_c = \bar{p} - \bar{p}_c > p(0) - k$) leads to redistribution of earnings in the low end of firm-CEO matches at the cost of distortion of highly able executive's skill accumulation in their non-CEO executive positions prior to being on the CEO market.

5 Conclusion

This paper studies executive skill accumulation on the non-CEO executive job prior to being on the CEO market, and the subsequent pay formation in the CEO market where firms differ in their size and executives in their inherent unobservable type. Given the high demand for CEO talent, executives have high stakes riding on establishing themselves as talented prior to being on the CEO market. This creates incentives to accumulate executive skill beyond the level that is required in their current positions because better executive skill brings them hard-to-resist market offers to work as a CEO. The equilibrium analysis shows that, as the firm size distribution is spread to the right, causing an increase in firm size, the proportion of the market pay attributed to the informational effect (direct productivity effect) of executive skill accumulation decreases (increases). This brings efficiency gains to the economy as the separating stable matching equilibrium under incomplete information converges to the stable matching equilibrium as the firm size distribution is more spread to the right. This paper also presents the empirical strategy on how to estimate the measure of efficiency (i.e., the proportion of the market pay attributed to the direct productivity effect) from data.

As Gabaix and Landier (2008) show that increases in firm size also brings dramatic increase in CEO pay over time and it contributes to income inequality. Therefore, the spread of the firm size distribution creates increasing tension between efficiency and inequality. CEO pay limits draw a great deal of attention given the rapidly increasing income inequality. Although there is a large literature on CEO pays there is still very little theoretical analysis on the impact of CEO pay limits on the economy. The analysis based on a contract theory with agency cost highlights how CEO pay limits can mitigate the agency cost that arises from the misalignment between the CEO and the shareholders' objectives when the CEO's effort is observable but non-verifiable. However, it does not show the overall impact on the economy because it focuses on the current firm-CEO match in isolation. We believe that CEO pay limits should be analyzed in the context of market competition because we clearly observe the high demand for CEO talent.

However, if the CEO pay limit is so severe that the minimum pay cut required for the highest-paid CEO is more than the lowest-paid CEO's expected pay net of her cost of working as a CEO, then it induces an equilibrium CEO pay function discontinuous at the top. This distorts high ability executives' equilibrium skill accumulation and results in a discontinuous bunching even prior to being on the CEO market. Therefore, redistribution of earnings may come at the low end of firm-CEO matches given this distorted executive skill

distribution.

A Proof of Proposition 1

A.1 Proof of Conditions 1, 2, and 4

We first prove the property of the executive skill function $y(\theta)$ stated in Condition 1 of Proposition 1. characteristics differ across executives in a separating stable equilibrium. It is clear that $y(\underline{\theta}) = 0$ because the utility cost of exerting effort to accumulate executive skill beyond the level that is required in the current position is infinitely large for the executive with the lowest type $\underline{\theta} = 0$.

Consider two executives, one of type θ and the other of type θ' with $\theta > \theta'$. Since each executive can earn the other executive's pay by accumulating the same level of executive skill, we must have

$$\begin{aligned} \delta(p(y(\theta)) - k) - c(y(\theta), \theta) &\geq \delta(p(y(\theta')) - k) - c(y(\theta'), \theta) \\ \delta(p(y(\theta')) - k) - c(y(\theta'), \theta') &\geq \delta(p(y(\theta)) - k) - c(y(\theta), \theta'), \end{aligned}$$

in equilibrium. Summing up the inequalities yields

$$c(y(\theta), \theta) + c(y(\theta'), \theta') \leq c(y(\theta'), \theta) + c(y(\theta), \theta').$$

This condition is satisfied if and only if $y(\theta) > y(\theta')$ in a separating stable equilibrium by the submodular property of c . Therefore, we have $y(\theta) > y(\theta')$ if $\theta > \theta'$.

Consider the skill accumulation decision by the executive of type θ . Since the market pay function is given by $p(\cdot)$, it must satisfy the first-order condition in (5) in the range of Y :

$$\delta p'(y(\theta)) - c_y(y(\theta), \theta) = 0.$$

Since the skill accumulation decision by the executive of type θ' also satisfies the first-order condition, we have

$$\delta p'(y(\theta')) - c_y(y(\theta'), \theta') = 0. \tag{49}$$

Suppose that $\theta' < \theta$. By the submodular property of c , (49) implies

$$\delta p'(y(\theta')) - c_y(y(\theta'), \theta) > 0. \tag{50}$$

Suppose that $\theta' > \theta$. By the submodular property of c , (49) implies

$$\delta p'(y(\theta')) - c_y(y(\theta'), \theta) < 0. \tag{51}$$

(50) and (51) show that no executive has an incentive to change her skill level to those that other executives accumulate if and only if (5) is satisfied for all θ .

What if an executive deviates to acquire $y > \bar{y} = \max Y$? The market pay function defined in Condition 4 for $y > \bar{y}$ should provide no incentive for such a deviation. Note that the market pay function defined in Condition 4 incorporates the firm's belief $\mu(y) = \bar{\theta}$ on the executive's type, conditional on $y > \bar{y}$ and it is based on the iso profit curve for the firm with the biggest size associated with hiring the executive of the highest type as its CEO, when her executive skill level is $y > \bar{y}$. First, consider a deviation by the executive of type $\bar{\theta}$. Applying (8) yields the right-hand limit as follows:

$$\delta p'_+(\bar{y}) - c_y(\bar{y}, \bar{\theta}) = \delta f_t(\bar{x}, t(\bar{y}, \bar{\theta}))t_y(\bar{y}, \bar{\theta}) - c_y(\bar{y}, \bar{\theta}).$$

This measures the rate of change in the executive's utility when she marginally increases her skill level above \bar{y} . Combining (5) and (6), we can induce

$$\delta (f_t(\bar{x}, t(\bar{y}, \bar{\theta}))t_y(\bar{y}, \bar{\theta}) + f_t(\bar{x}, t(\bar{y}, \bar{\theta}))t_\theta(\bar{y}, \bar{\theta})\mu'(\bar{y})) - c_y(\bar{y}, \bar{\theta}) = 0. \quad (52)$$

(52) implies that

$$\delta p'_+(\bar{y}) - c_y(\bar{y}, \bar{\theta}) = \delta f_t(\bar{x}, t(\bar{y}, \bar{\theta}))t_y(\bar{y}, \bar{\theta}) - c_y(\bar{y}, \bar{\theta}) < 0. \quad (53)$$

Therefore, the utility for the executive of type $\bar{\theta}$ indeed decreases as she marginally increase her skill level above \bar{y} . Because f is concave in t , t is concave in y , and c is convex in y , (53) implies that for all $y > \bar{y}$, we have

$$\delta p'(y) - c_y(y, \bar{\theta}) = \delta f_t(\bar{x}, t(y, \bar{\theta}))t_y(y, \bar{\theta}) - c_y(y, \bar{\theta}) < 0.$$

Therefore, the utility for the executive of type $\bar{\theta}$ monotonically decreases as she increases her skill level beyond \bar{y} . Therefore, we can conclude that, for all $y > \bar{y}$

$$\delta(p(y) - k) - c(y, \bar{\theta}) < \delta(p(\bar{y}) - k) - c(\bar{y}, \bar{\theta}). \quad (54)$$

Due to the submodular property of c , (54) induces that for all $\theta < \bar{\theta}$ and all $y > \bar{y}$

$$\delta(p(y) - k) - c(y, \theta) < \delta(p(\bar{y}) - k) - c(\bar{y}, \theta). \quad (55)$$

Because no executive wants to change her skill level to any other executive's skill level, we have that for all $\theta < \bar{\theta}$,

$$\delta(p(\bar{y}) - k) - c(\bar{y}, \theta) < \delta(p(y(\theta)) - k) - c(y(\theta), \theta). \quad (56)$$

(55) and (56) yield that for all $\theta < \bar{\theta}$ and all $y > \bar{y}$

$$\delta(p(y) - k) - c(y, \theta) < \delta(p(y(\theta)) - k) - c(y(\theta), \theta).$$

Thus, utility levels for all other executives by acquiring $y > \bar{y}$ are also lower than their equilibrium payoffs. Therefore, no executive has incentives to acquire $y > \bar{y}$.

Condition 2 is about the firm's belief function $\mu(y)$. In a separating equilibrium, the executive's skill fully reveals her type and hence we have that, for all θ , $\mu(y(\theta)) = \theta$. Given Condition 1, the firm's belief function is increasing up until \bar{y} and $\mu(\bar{y}) = \bar{\theta}$. Because $\mu(\bar{y}) = \bar{\theta}$, $\mu(y)$ for all $y > \bar{y}$ cannot be lower than $\bar{\theta}$ given the monotonicity. Since $\bar{\theta}$ is the highest type level, it implies that for all $y > \bar{y}$, $\mu(y) = \bar{\theta}$.

A.2 Proof of Condition 3

Consider the property of the market matching function $x(y)$ in Condition 3. For condition 3.(a), consider two executives, one with executive skill level y and the other with y' . Suppose that the executive with y works for firm x and the executive with y' works for firm x' in equilibrium such that $x > x'$. If firm x' were to hire the executive with y , the maximum pay that it is willing to offer is $f(x', t(y, \mu(y))) - [f(x', t(y', \mu(y')))] - p(y')$ because it will not hire the executive with y if the profit associated with hiring her is less than the profit associated with hiring the executive with y' at the market pay $p(y')$. Therefore, if the executive with y wants to work for firm x at the market pay $p(y)$ instead of working for firm x' , then the following condition is satisfied in equilibrium:

$$p(y) \geq f(x', t(y, \mu(y))) - [f(x', t(y', \mu(y')))] - p(y'). \quad (57)$$

Similarly, if the executive with y' wants to work for firm x' at the market pay $p(y')$ instead of working for firm x , the following condition is satisfied in equilibrium as well:

$$p(y') \geq f(x, t(y', \mu(y')))] - f(x, t(y, \mu(y))) + p(y). \quad (58)$$

Summing up (57) and (58) yields

$$f(x', t(y', \mu(y')))] + f(x, t(y, \mu(y))) \geq f(x', t(y, \mu(y))) + f(x, t(y', \mu(y')))]. \quad (59)$$

Since the firm's belief function $\mu(y)$ is an increasing function, $t(y, \mu(y)) \geq t(y', \mu(y'))$ if $y \geq y'$. Because $x > x'$ and f satisfies the supermodular assumption, (59) implies that $y > y'$. Therefore, an executive with a higher skill level matches with a firm with a bigger size.

Consider two firms, one hiring the executive with y and the other hiring the executive with y' such that $y > y'$, that is, firms $x(y)$ and $x(y')$. Since each firm

can hire an executive by paying a pay that is at least as high as the market pay, we must have

$$f(x(y), t(y, \mu(y))) - p(y) \geq f(x(y), t(y', \mu(y'))) - p(y'), \quad (60)$$

$$f(x(y'), t(y', \mu(y'))) - p(y') \geq f(x(y'), t(y, \mu(y))) - p(y). \quad (61)$$

in equilibrium. Summing up (60) and (61) yields

$$f(x(y), t(y, \mu(y))) + f(x(y'), t(y', \mu(y'))) \geq f(x(y), t(y', \mu(y'))) + f(x(y'), t(y, \mu(y))) \quad (62)$$

Since $y > y'$, we have $\mu(y) > \mu(y')$ and hence $t(y, \mu(y)) > t(y', \mu(y'))$. Because f satisfies the supermodular assumption, (62) implies that $x(y) > x(y')$. Therefore, a firm with a bigger size matches with an executive with a higher skill level. From both sides on the market, we can then conclude that the market matching function must be positively assortative, which is stated in Condition 3.(a)

For Condition 3.(b). Consider firm $x(y)$ who hires the executive with $y = y(\theta)$ for each $y \in Y$. It is necessary that the first-order condition (7) holds for the firm's hiring problem (2):

$$f_t(x(y), t(y, \mu(y))) (t_y(y, \mu(y)) + t_\theta(y, \mu(y))\mu'(y)) - p'(y) = 0.$$

Now consider the executive with some skill level y in Y . The first-order condition for the solution to the executive's working problem (3) is

$$\begin{aligned} & f_x(x, t(y, \mu(y))) \\ & - f_t(x, t(\tilde{y}(x), \mu(\tilde{y}(x)))) [t_y(\tilde{y}(x), \mu(\tilde{y}(x))) + t_\theta(\tilde{y}(x), \mu(\tilde{y}(x)))\mu'(\tilde{y}(x))] \tilde{y}'(x) \\ & - f_x(x, t(\tilde{y}(x), \mu(\tilde{y}(x)))) + p'(\tilde{y}(x))\tilde{y}'(x) = 0. \end{aligned} \quad (63)$$

at $\tilde{y}(x) = y$. Because

$$f_x(x, t(y, \mu(y))) = f_x(x, t(\tilde{y}(x), \mu(\tilde{y}(x))))$$

at $\tilde{y}(x) = y$ and $\tilde{y}(\cdot)$ is an increasing function, (63) is equivalent to (7) at $\tilde{y}(x) = y$. This shows that it is also necessary that the first-order condition (7) holds for the executive's working problem (3). Finally, we must have $k \leq p(y(\underline{\theta})) \leq f(\underline{x}, 1)$ in an equilibrium because $f(\underline{x}, 1)$ is the earnings created at the bottom match between the smallest firm and the least talented executive with $t(0, 0) = 1$ and k is her cost of working as a CEO. Therefore, Condition 3 (i.e., 3.(a) and 3.(b)) is the necessary condition for $\{x(\cdot), p(\cdot)\}$ to be the pair of the market matching function and the market pay function.

Let us prove that Condition 3 is the sufficient condition as well. Suppose that the executive with $y = \tilde{y}(x)$ wants to work for firm x' . Then the first-order condition for the executive with y at x' becomes

$$\begin{aligned} & f_x(x', t(y, \mu(y))) \\ & - f_t(x', t(\tilde{y}(x'), \mu(\tilde{y}(x')))) [t_y(\tilde{y}(x'), \mu(\tilde{y}(x'))) + t_\theta(\tilde{y}(x'), \mu(\tilde{y}(x')))\mu'(\tilde{y}(x'))] \tilde{y}'(x') \\ & \quad - f_x(x', t(\tilde{y}(x'), \mu(\tilde{y}(x')))) + p(\tilde{y}(x'))\tilde{y}'(x'). \end{aligned} \quad (64)$$

Since (7) must be satisfied for the executive with $y' = \tilde{y}(x')$, (64) is equivalent to

$$f_x(x', t(y, \mu(y))) - f_x(x', t(\tilde{y}(x'), \mu(\tilde{y}(x')))). \quad (65)$$

If $x' > x$, $\tilde{y}(x') > y = \tilde{y}(x)$ and $\mu(\tilde{y}(x')) > \mu(y)$ because $\tilde{y}(\cdot)$ is an increasing function given Condition 3.(a) and $\mu(\cdot)$ is a increasing function. Therefore, if $x' > x$, then $t(\tilde{y}(x'), \mu(\tilde{y}(x'))) > t(y, \mu(y))$. This implies that (65) is negative by the supermodular assumption on f . If $x' < x$, $\tilde{y}(x') < y = \tilde{y}(x)$ and $\mu(\tilde{y}(x')) < \mu(y)$. Therefore, if $x' < x$, then $t(\tilde{y}(x'), \mu(\tilde{y}(x'))) < t(y, \mu(y))$ (65) is positive by the supermodular assumption on f . Therefore, (7) in Condition 3.(b), together with Condition 3.(a) (i.e., positively assortative matching), is the necessary and sufficient condition for the executive with $y \in Y$ to work for firm $x(y)$ at pay $p(y)$.

Finally, consider the firm's decision. Since (7) must hold for any firm $x(y')$ who hires the executive with $y' = y(\theta')$

$$f_t(x(y'), t(y', \mu(y')))[t_y(y', \mu(y')) + t_\theta(y', \mu(y'))\mu'(y')] - p'(y') = 0. \quad (66)$$

Suppose that firm $x(y)$ considers hiring the executive with $y' = y(\theta')$. Then, the first-order condition for firm $x(y)$ at y' is

$$f_t(x(y), t(y', \mu(y')))[t_y(y', \mu(y')) + t_\theta(y', \mu(y'))\mu'(y')] - p'(y'). \quad (67)$$

If $y' \geq y$, $x(y') \geq x(y)$ because of Condition 3.(a). Then, the supermodular assumption on f implies that if $y' \geq y$, then

$$f_t(x(y'), t(y', \mu(y'))) \geq f_t(x(y), t(y', \mu(y'))). \quad (68)$$

Because $\mu'(y') > 0$, and t_y and t_θ are also positive, (68) implies that if $y' > y$, then (66) is negative. On the other hand, if $y' < y$, then (66) is positive. It means that firm $x(y)$ does not want to hire an executive other than the executive with y . Therefore, (7) in Condition 3.(b), together with Condition 3.(a), is also the necessary and sufficient condition for firm $x(y)$ to hire the executive with y at pay $p(y)$.

B Proof of Theorem 3

B.1 Proof of Conditions 1 and 2

First we tackle Condition 1. We need to show that every executive has no incentives to acquire any executive skill level other than the one she is supposed to acquire in equilibrium. We start with examining whether the executive of type θ_Δ has incentives to change her executive skill level.

1. Type- θ_Δ executive's deviation to change executive skill level

The executive's equilibrium skill level is y_Δ . There are *three types of deviations*. The *first type* is to change her skill level to some $y > y_\Delta$. Clearly she has no incentive to acquire a higher skill level than y_Δ because she cannot get a pay higher than \bar{p}_c and a higher skill level incurs extra utility costs.

The *second type* is to change her executive skill level to $y \in [y(\theta_\Delta), y_\Delta]$. Note that the payoff for the executive by acquiring y_Δ is $\delta(\bar{p}_c - k) - c(y_\Delta, \theta_\Delta)$ and (34) in Lemma 2 shows that it is

$$\begin{aligned} \delta(\bar{p}_c - k) - c(y_\Delta, \theta_\Delta) &= \delta(p(y(\theta_\Delta)) - \Delta - k) - c(y(\theta_\Delta), \theta_\Delta) \\ &= \delta(p_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta_\Delta), \end{aligned} \quad (69)$$

where the second equality is due to (46). Therefore, we can examine whether the executive has an incentive to acquire $y \in [y(\theta_\Delta), y_\Delta]$ by comparing her payoff from acquiring $y \in [y(\theta_\Delta), y_\Delta]$ with $\delta(p_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta_\Delta)$. For the comparison, note that because $p_c(y)$ is adjusted according to (44) for $y \in [y(\theta_\Delta), y_\Delta]$, taking the derivative of (44) with respect to y yields

$$p'_c(y) = f_t(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta))t_y(y, \theta_\Delta). \quad (70)$$

Applying (70), we can see that, for all $y \in [y(\theta_\Delta), y_\Delta]$

$$\delta p'_c(y) - c_y(y, \theta_\Delta) = \delta f_t(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta))t_y(y, \theta_\Delta) - c_y(y, \theta_\Delta). \quad (71)$$

In the original equilibrium, we have that for all $y > y(\theta_\Delta)$

$$\begin{aligned} \delta p'_c(y) - c_y(y, \theta_\Delta) \\ = \delta f_t(x(y), t(y, \mu(y)))(t_y(y, \mu(y)) + t_\theta(y, \mu(y))) - c_y(y, \theta_\Delta) < 0. \end{aligned} \quad (72)$$

Note that $f_t(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta)) < f_t(x(y), t(y, \mu(y)))$ because $\tilde{x}(\theta_\Delta) < x(y)$ and $t(y, \theta_\Delta) < t(y, \mu(y))$ given $\theta_\Delta < \mu(y)$. Furthermore, $t_y(y, \theta_\Delta) < t_y(y, \mu(y))$ because of the supermodular assumption given $\theta_\Delta < \mu(y)$. Because $t_\theta(y, \mu(y)) > 0$, we then have

$$f_t(\tilde{x}(\theta_\Delta), t(y, \theta_\Delta))t_y(y, \theta_\Delta) < f_t(x(y), t(y, \mu(y)))(t_y(y, \mu(y)) + t_\theta(y, \mu(y))) \quad (73)$$

Combing (72) and (73) imply that the sign of (71) is negative:

$$\delta p'_c(y) - c_y(y, \theta_\Delta) < 0.$$

This implies that the executive's payoff by acquiring any $y \in [y(\theta_\Delta), y_\Delta]$ is strictly lower than her payoff with y_Δ :

$$\delta(\bar{p}_c - k) - c(y_\Delta, \theta_\Delta) = \delta(p_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta_\Delta) > \delta(p_c(y) - k) - c(y, \theta_\Delta). \quad (74)$$

Therefore, the executive of type θ_Δ has no incentive to acquire any $y \in [y(\theta_\Delta), y_\Delta]$.

The last type is to change her executive skill level to $y \in [0, y(\theta_\Delta)]$. First of all, the executive does not have an incentive to deviate to acquire $y(\theta_\Delta)$ because her payoff in that case is simply the same as her payoff with y_Δ according to (69):

$$\delta(\bar{p}_c - k) - c(y_\Delta, \theta_\Delta) = \delta(p_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta_\Delta). \quad (75)$$

From (45) and (46), we know that $p_c(y) = p(y) - \Delta$ for all $y \in [0, y(\theta_\Delta)]$. To show that the executive of type θ_Δ has no incentive to acquire $y \in [0, y(\theta_\Delta)]$, we only need to establish

$$\begin{aligned} \delta(p(y(\theta_\Delta)) - \Delta - k) - c(y(\theta_\Delta), \theta_\Delta) &\geq \delta(p(y) - \Delta - k) - c(y, \theta_\Delta) \\ \Leftrightarrow \delta(p(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta_\Delta) &\geq \delta(p(y) - k) - c(y, \theta_\Delta). \end{aligned}$$

The last inequality is indeed satisfied with strict inequality from the proof of Condition 1 in Proposition 1 because $p(\cdot)$ is the market pay function and $y(\cdot)$ is the executive skill function in the original equilibrium. Therefore, for all $y \in [0, y(\theta_\Delta)]$

$$\delta(\bar{p}_c - k) - c(y_\Delta, \theta_\Delta) = \delta(p_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta_\Delta) > \delta(p_c(y) - k) - c(y, \theta_\Delta), \quad (76)$$

which shows that the executive of type θ_Δ has no incentive to acquire any $y \in [0, y(\theta_\Delta)]$. Therefore, the executive of type θ_Δ has no incentive to change her skill y_Δ

2. Type- θ executive's deviation to change executive skill level ($\theta > \theta_\Delta$).

The executive whose type is greater than θ_Δ acquires y_Δ . She clearly has no incentive to acquire y that is strictly above y_Δ because a higher skill level induces higher utility costs, but the same pay \bar{p}_c that y_Δ induces. In order to show that she has no incentive to acquire a skill level that is strictly less than y_Δ , note that (74) and (76) together show that the type- θ_Δ executive's payoff with $y \in [0, y_\Delta]$ is strictly lower than her equilibrium payoff: For all $y \in [0, y_\Delta]$

$$\delta(\bar{p}_c - k) - c(y_\Delta, \theta_\Delta) > \delta(p_c(y) - k) - c(y, \theta_\Delta). \quad (77)$$

Because $y_\Delta > y$ and $\theta > \theta_\Delta$, the submodular property of c leads (77) implies that, for all $y \in [0, y_\Delta)$,

$$\delta(\bar{p}_c - k) - c(y_\Delta, \theta) > \delta(p_c(y) - k) - c(y, \theta),$$

which shows that the executive of type $\theta > \theta_\Delta$ also has no incentive to acquire y that is strictly below y_Δ . Therefore, θ_Δ is the unique optimal executive skill level for the executive of type $\theta > \theta_\Delta$.

3. Type- θ executive's deviation to change executive skill level ($\theta < \theta_\Delta$).

Consider effort exerting decisions by the executives whose types are in $[\underline{\theta}, \theta_\Delta)$. Those executives' effort exerting decisions are identical to the ones in the original equilibrium. To see that, note $y_c(\underline{\theta}) = 0 = y(\underline{\theta})$. For all $\theta \in (\underline{\theta}, \theta_\Delta)$, $y_c(\theta) = y(\theta)$ because $p'_c(y)$ in (40) is the same as $p'(y)$ for all $y < y(\theta_\Delta)$. Then we can apply the proof of Condition 1 from Proposition 1 to show that (40) implies that no executive in $[\underline{\theta}, \theta_\Delta)$ wants to acquire the levels of the executive skills that other executives in $[\underline{\theta}, \theta_\Delta)$ acquire.

The only question is whether the executive has an incentive to acquire $y \geq y(\theta_\Delta)$. Suppose that the executive deviates to acquire $y = y(\theta_\Delta)$. Then, her payoff is $\delta(p_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta)$. The proof of Condition 1 of Proposition 1 shows

$$\delta(p(y(\theta)) - k) - c(y(\theta), \theta) > \delta(p(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta).$$

This inequality is equivalent to

$$\delta(p_c(y_c(\theta)) - k) - c(y_c(\theta), \theta) > \delta(p_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta) \quad (78)$$

because $y_c(\theta) = y(\theta)$ and $p_c(y) = p(y) - \Delta$ for all $y \leq y(\theta_\Delta)$. (78) implies that the executive also has no incentive to deviate to acquire $y(\theta_\Delta)$.

Suppose that the executive deviate to acquire $y \in (y(\theta_\Delta), y_\Delta)$. From (74), we have

$$\delta(\bar{p}_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta_\Delta) > \delta(p_c(y) - k) - c(y, \theta_\Delta). \quad (79)$$

Because $y > y(\theta_\Delta)$ and $\theta_\Delta > \theta$, the submodular property of c leads (79) to

$$\delta(p_c(y(\theta_\Delta)) - k) - c(y(\theta_\Delta), \theta) > \delta(p_c(y) - k) - c(y, \theta). \quad (80)$$

Combining (78) and (80) yields

$$\delta(p_c(y_c(\theta)) - k) - c(y_c(\theta), \theta) > \delta(p_c(y) - k) - c(y, \theta),$$

which shows that the executive of type $\theta \in [\underline{\theta}, \theta_\Delta)$ has no incentive to acquire $y \in (y(\theta_\Delta), y_\Delta)$.

Suppose that the executive deviates to acquire y_Δ . Because $y_\Delta > y(\theta_\Delta)$ and $\theta_\Delta > \theta$, applying the submodular property of the utility cost function $c(\cdot, \cdot)$ to (75) yields

$$\delta(\bar{p}_c - k) - c(y_\Delta, \theta) < \delta(p_c(y(\theta)) - k) - c(y(\theta_\Delta), \theta). \quad (81)$$

Combining (78) and (81) yields

$$\delta(p_c(y_c(\theta)) - k) - c(y_c(\theta), \theta) > \delta(\bar{p}_c - k) - c(y_\Delta, \theta),$$

which shows that the executive of type $\theta \in [\underline{\theta}, \theta_\Delta)$ has no incentive to acquire y_Δ . It is then clear that the executive of type $\theta \in [\underline{\theta}, \theta_\Delta)$ has no incentive to acquire $y > y_\Delta$ because such a higher skill level incurs higher utility costs but the same pay \bar{p}_c as y_Δ induces. Therefore, $y_c(\theta)$ is the unique optimal skill level for the executive of type $\theta \in [\underline{\theta}, \theta_\Delta)$. This concludes the proof of Condition 1

Now we prove condition 2. Firms correctly infer the type of an executive whose type is less than θ_Δ because she acquires the same executive skill that she would have acquired in the original equilibrium. Therefore, the firm's belief on an executive's type is θ if $y = y_c(\theta) < y(\theta_\Delta)$. If y belongs to $[y(\theta_\Delta), y_\Delta)$, then it is a skill level that is not supposed to be observed in equilibrium. In this case, the firm believes that an executive's type is θ_Δ . Therefore, $\mu_c(\cdot)$ is non-decreasing as y increases. If $y = y_\Delta$, then the firm believes, according to Bayes' rule, that an executive's type θ follows the probability distribution $G(\theta|\theta \geq \theta_\Delta)$ conditional on $\theta \geq \theta_\Delta$, because every executive of type θ_Δ or higher acquires the same θ_Δ . Because the support of $G(\cdot|\theta \geq \theta_\Delta)$ is not less than θ_Δ , $\mu_c(\cdot)$ again satisfies the non-decreasing property. Finally, if $y > y_\Delta$, then this is again a skill level off the equilibrium. In this case, the firm believes that an executive's type is the highest possible type. Therefore, the firm's belief function in Condition 2 is non-decreasing and it uses Bayes' rule on the equilibrium path.

B.2 Proof of Conditions 3 and 4

Now we prove that given $\{y_c(\cdot), \mu_c(\cdot)\}, \{x_c(\cdot), p_c(\cdot)\}$ characterizes stable matching between firms and CEOs. We need to show that no firm or CEO has incentives to change the CEO that it currently employs or the firm she currently works for. We first examine whether a CEO has incentives to work for an alternative firm. There are those CEOs whose executive skill level is exactly y_Δ and those CEOs whose executive skill is below y_Δ . First consider a CEO with $y < y_\Delta$

1 a CEO's deviation to change her firm

1.1 Deviation to change her firm by a CEO with $y < y_\Delta$

Because there is no CEO with $y \in [y_c(\theta_\Delta), y_\Delta)$, we have that $y < y_c(\theta_\Delta)$ in this case. The CEO currently works for the firm with size $x_c(y)$ at pay $p_c(y)$. There are *three types of deviations* to change her firm. She can work for a firm of size x less than $\tilde{x}(\theta_\Delta)$, or the firm of size exactly equal to $\tilde{x}(\theta_\Delta)$, or a firm of size greater than $\tilde{x}(\theta_\Delta)$. Recall that $\tilde{x}(\theta_\Delta) = x(y(\theta_\Delta))$ is the size of the firm that hires the executive of type θ_Δ as its CEO in the original equilibrium with no pay limit and that it is the size of the smallest firm amongst all the firms that hire the executive with y_Δ as its CEO in the new equilibrium with the pay limit.

First, suppose that the CEO parts away from her current firm to match with a firm with size $x < \tilde{x}(\theta_\Delta)$. Note that she receives the pay $p_c(y) = p(y) - \Delta$ from her current firm. For $y < y_c(\theta_\Delta)$ and $x < \tilde{x}(\theta_\Delta)$, market matching, executive skill levels, and the firm's belief are identical to those in the original equilibrium. The only difference is that the pay is shifted down by Δ in the range of $[0, y_c(\theta_\Delta))$. Because the pay is shifted by the same amount, it does not alter stable matching between firms and CEOs in these ranges. Therefore, by using the proof of Condition 3 of Proposition 1, we can infer that the CEO with y has no incentive to part from her current firm with size $x_c(y)$ in order to work for any other firm with size less than $\tilde{x}(\theta_\Delta)$.

Second, suppose that the CEO parts away from her current firm to work for the firm with size exactly equal to $\tilde{x}(\theta_\Delta)$. If she wants to work for the firm with size $\tilde{x}(\theta_\Delta)$, then the maximum pay, denoted by $\hat{p}(\tilde{x}(\theta_\Delta))$, she can get from the firm is determined by

$$\begin{aligned} f(\tilde{x}(\theta_\Delta), t(y, \mu_c(y)) - \hat{p}(\tilde{x}(\theta_\Delta))) &= \mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c \quad (82) \\ &= f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) - p_c(y(\theta_\Delta)). \end{aligned}$$

Rearranging terms in the equation above yields

$$\begin{aligned} \hat{p}(\tilde{x}(\theta_\Delta)) &= f(\tilde{x}(\theta_\Delta), t(y, \mu_c(y)) - f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) + p_c(y(\theta_\Delta)) \quad (83) \\ &= f(\tilde{x}(\theta_\Delta), t(y, \mu_c(y)) - f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) + p(y(\theta_\Delta)) - \Delta. \end{aligned}$$

Because $\{x(\cdot), p(\cdot)\}$ is a stable matching given $\{y(\cdot), \mu(\cdot)\}$ in the original equilibrium, the firm with size $\tilde{x}(\theta_\Delta)$ has no incentive to hire an executive with $y \neq y(\theta_\Delta)$ as its CEO. In particular, the proof of Condition 3 of Proposition 1 shows that

$$\begin{aligned} f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) - p(y(\theta_\Delta)) &> f(\tilde{x}(\theta_\Delta), t(y, \mu(y)) - p(y) \\ \Leftrightarrow p(y) &> f(\tilde{x}(\theta_\Delta), t(y, \mu(y)) - f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) + p(y(\theta_\Delta)) \quad (84) \end{aligned}$$

Because $\mu(y) = \mu_c(y)$ for $y < y(\theta_\Delta)$, (84) and (83) imply

$$\begin{aligned} p(y) - \Delta &> f(\tilde{x}(\theta_\Delta), t(y, \mu(y))) - f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) + p(y(\theta_\Delta)) - \Delta \\ &\Leftrightarrow p_c(y) > \hat{p}(\tilde{x}(\theta_\Delta)). \end{aligned}$$

Therefore, a CEO with $y < y(\theta_\Delta)$ has no incentive to work for the firm with size exactly equal to $\tilde{x}(\theta_\Delta)$.

Third, suppose that the CEO parts from her current firm to work for a firm with size x greater than $\tilde{x}(\theta_\Delta)$. Let $\hat{p}(x)$ denote the maximum pay she can get from the firm:

$$\begin{aligned} f(x, t(y, \mu_c(y))) - \hat{p}(x) &= \mathbb{E}[f(x, t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c \\ \Leftrightarrow \hat{p}(x) &= f(x, t(y, \mu_c(y))) - \mathbb{E}[f(x, t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] + \bar{p}_c. \end{aligned}$$

Taking the derivative of $\hat{p}(x)$ with respect to x yields

$$\hat{p}'(x) = f_x(x, t(y, \mu_c(y))) - \mathbb{E}[f_x(x, t(y_\Delta, \theta)) | \theta \geq \theta_\Delta].$$

Because $y < y_\Delta$ and $\mu_c(y) < \theta_\Delta$, we have $t(y, \mu_c(y)) < t(y_\Delta, \theta)$ for all $\theta \geq \theta_\Delta$. Then, the supermodular property of f implies that, for all $\theta \geq \theta_\Delta$

$$f_x(x, t(y, \mu_c(y))) < f_x(x, t(y_\Delta, \theta))$$

and hence

$$f_x(x, t(y, \mu_c(y))) < \mathbb{E}[f_x(x, t(y_\Delta, \theta)) | \theta \geq \theta_\Delta]. \quad (85)$$

(85) leads to $\hat{p}'(x) < 0$. This means that $\hat{p}(\tilde{x}(\theta_\Delta)) > \hat{p}(x)$, for all $x > \tilde{x}(\theta_\Delta)$: the maximum pay that the executive with y can get from a firm with size x greater than $\tilde{x}(\theta_\Delta)$ is even lower than the maximum pay $\hat{p}(\tilde{x}(\theta_\Delta))$ that she can receive from the firm with size exactly equal to $\tilde{x}(\theta_\Delta)$. Therefore, the CEO has no incentive to match with a firm with size x greater than $\tilde{x}(\theta_\Delta)$.

1.2. Deviation to change her firm by the CEO with $y = y_\Delta$

Currently, she works for some firm with size no less than $\tilde{x}(\theta_\Delta)$ and receives pay \bar{p}_c . Because the current firm pays her the maximum pay \bar{p}_c that is allowed under the pay limit, she cannot receive any higher pay by changing her firm. Therefore, she has no incentive to part from the current firm to work for an alternative firm. Therefore, we can conclude that any executives have no incentive to part from the current firms that hire them given the current pay offers. Now consider the firm's deviation to change its executive.

2 A firm's deviation to change its CEO

2.1 Deviation to change its CEO by firm $\tilde{x}(\theta_\Delta)$: Because the firm with size exactly equal to $\tilde{x}(\theta_\Delta)$ hires an executive with y_Δ as its CEO and there is no executive with a higher skill level, there is only one type of deviation: deviation to hire an executive with $y < y_\Delta$, which also implies $y < y(\theta_\Delta)$. Then, the minimum pay that the executive with y is willing to accept is $p_c(y)$. Therefore, the firm's maximum profit upon hiring the executive with y is

$$f(\tilde{x}(\theta_\Delta), t(y, \mu_c(y)) - p_c(y)).$$

Given $p_c(y) = p(y) - \Delta$ for all $y \leq y(\theta_\Delta)$, (35) shows that the firm's profit with its current executive is

$$\mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c = f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta) - p_c(y(\theta_\Delta))). \quad (86)$$

Because $\mu_c(y) = \mu(y)$ and $p_c(y) = p(y) - \Delta$ for all $y \leq y(\theta_\Delta)$, we can apply part of the proof of Condition 3 in Proposition 1,

$$f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta) - p(y(\theta_\Delta))) > f(\tilde{x}(\theta_\Delta), t(y, \mu(y)) - p(y)) \text{ for all } y \neq y(\theta_\Delta),$$

to show that

$$f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta) - p_c(y(\theta_\Delta))) > f(\tilde{x}(\theta_\Delta), t(y, \mu_c(y)) - p_c(y)). \quad (87)$$

(86) and (87) yield

$$\mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c > f(\tilde{x}(\theta_\Delta), t(y, \mu_c(y)) - p_c(y)). \quad (88)$$

Therefore, the firm with size exactly equal to $\tilde{x}(\theta_\Delta)$ has no incentive to hire any executive with $y < y_\Delta$ as its CEO.

2.2 Deviation to change its CEO by firm $x > \tilde{x}(\theta_\Delta)$. The firm's current CEO has the executive skill level y_Δ . The only possible deviation is to hire an executive with $y < y_\Delta$, which also implies $y < y(\theta_\Delta)$. Note that $t(y_\Delta, \theta) > t(y, \mu_c(y))$ given $y_\Delta > y$ and $\theta > \mu_c(y)$ for all $\theta \geq \theta_\Delta$. Because $x > \tilde{x}(\theta_\Delta)$, the supermodular property of f implies that for all $\theta \geq \theta_\Delta$

$$f(x, t(y_\Delta, \theta)) - f(x, t(y, \mu_c(y))) > f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) - f(\tilde{x}(\theta_\Delta), t(y, \mu_c(y))). \quad (89)$$

Taking the expectation of both sides in (89) over θ conditional on $\theta \geq \theta_\Delta$ yields

$$\begin{aligned} \mathbb{E}[f(x, t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - f(x, t(y, \mu_c(y))) > \\ \mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - f(\tilde{x}(\theta_\Delta), t(y, \mu_c(y))). \end{aligned} \quad (90)$$

Combining (88) and (90) yields

$$\mathbb{E}[f(x, t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c > f(x, t(y, \mu_c(y))) - p_c(y). \quad (91)$$

Note that the right-hand-side of (91) is the maximum profit that firm x can get by hiring an executive with $y < y_\Delta$, because $p_c(y)$ is the minimum pay that an executive with y is willing to accept. The left-hand-side is the profit that firm x gets with the current executive with y_Δ and hence the firm has no incentive to part from its current CEO to hire an executive with $y < y_\Delta$ as its CEO.

2.3 Deviation to change its CEO by firm $x < \tilde{x}(\theta_\Delta)$: Firm x has no incentive to part from the current CEO to match with any other executive with $y \in [0, y(\theta_\Delta))$. This can be easily shown by applying the proof of Condition 3 in Proposition 1 since market matching, worker's investment, and the firm's belief for $y \in [0, y(\theta_\Delta))$ are identical to those in the original equilibrium. The only difference is that the market pay is shifted down by the same amount Δ at every y . Then, the only remaining deviation that we need to consider is to part from the current CEO in order to hire an executive with y_Δ as its CEO.

Suppose that firm x deviates to hire an executive with y_Δ . Because the executive can receive \bar{p}_c , the minimum pay that she is willing to accept is \bar{p}_c . Then firm x 's profit is

$$\pi_d(x) := \mathbb{E}[f(x, t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c.$$

Because the firm's current profit without deviation is

$$\pi_o(x) := f(x, t(\tilde{y}_c(x), \mu_c(\tilde{y}_c(x)))) - p_c(\tilde{y}_c(x)),$$

where $\tilde{y}_c(x) \in [0, y(\theta_\Delta))$ denotes the skill level of the executive that firm x hires without deviation. Note that according to (46)

$$f(\tilde{x}(\theta_\Delta), t(y(\theta_\Delta), \theta_\Delta)) - p_c(y(\theta_\Delta)) = \mathbb{E}[f(\tilde{x}(\theta_\Delta), t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] - \bar{p}_c. \quad (92)$$

(92) implies that

$$\sup_{x < \tilde{x}(\theta_\Delta)} \pi_o(x) = \lim_{x \nearrow \tilde{x}(\theta_\Delta)} \pi_o(x) = \lim_{x \nearrow \tilde{x}(\theta_\Delta)} \pi_d(x) = \sup_{x < \tilde{x}(\theta_\Delta)} \pi_d(x). \quad (93)$$

To see whether firm $x < \tilde{x}(\theta_\Delta)$ has an incentive to hire an executive with y_Δ , take the derivative of $\pi_d(x)$ and $\pi_o(x)$ respectively:

$$\begin{aligned} \pi'_d(x) &= \mathbb{E}[f_x(x, t(y_\Delta, \theta)) | \theta \geq \theta_\Delta] > 0, \\ \pi'_o(x) &= f_x(x, t(\tilde{y}_c(x), \mu_c(\tilde{y}_c(x)))) > 0. \end{aligned}$$

The expression of $\pi'_o(x)$ above is due to the envelop theorem. Because $t(y_\Delta, \theta) > t(\tilde{y}_c(x), \mu_c(\tilde{y}_c(x)))$ given $y_\Delta > \tilde{y}_c(x)$ and $\theta > \mu_c(\tilde{y}_c(x))$ for all $\theta \geq \theta_\Delta$, the super-modular property of f implies that

$$f_x(x, t(y_\Delta, \theta)) > f_x(x, t(\tilde{y}_c(x), \mu_c(\tilde{y}_c(x)))) \text{ for all } \theta \geq \theta_\Delta,$$

which leads to

$$\pi'_d(x) > \pi'_o(x) \text{ at every } x < \tilde{x}(\theta_\Delta). \quad (94)$$

Because $\pi_d(x)$ has a steeper slope than $\pi_o(x)$ does at every point x (according to (94)) and two functions have the same supremum at the right open end $\tilde{x}(\theta_\Delta)$ of the interval (according to (93)), we have that

$$\pi_d(x) < \pi_o(x) \text{ at every } x < \tilde{x}(\theta_\Delta).$$

Therefore, firm $x < \tilde{x}(\theta_\Delta)$ has no incentive to hire an executive with y_Δ as its CEO. This shows that no firm has incentives to change its CEO.

References

- [1] Baker, G. P., R. Gibbons, and K. J. Murphy (1994): “Subjective Performance Measures in Optimal Incentive Contracts,” *Quarterly Journal of Economics*, 109(4), 1125-1156.
- [2] Bebchuk, L., and J. Fried (2004): *Pay Without Performance: The Unfulfilled Promise of Executive Remuneration*. Cambridge, MA., Harvard University Press.
- [3] Cebon, P. and B. E. Hermalin (2015): “When Less Is More: The Benefits of Limits on Executive Pay,” *Review of Financial Studies*, 28 (6): 1667-1700
- [4] Fong, Y.-F. and J. Li (2015): “Relational Contracts, Limited Liability, and Employment Dynamics,” mimeo, Kellogg School of Management, Northwestern University.
- [5] Frydman, C. (2014): “Rising through the Ranks. The Evolution of the Market for Corporate Executives, 1936-2003,” mimeo, Boston University.
- [6] Gabaix, X. and A. Landier (2008): “Why has CEO pay increased so much?” *The Quarterly Journal of Economics* 123(1), 49-100.
- [7] Gayle, G.-L., Golan, L., and A. M. Robert (2015): “Promotion, Turnover, and Compensation in the Executive Labor Market,” *Econometrica*, 83(6), 2293-2369.

- [8] Holmström, B. (1979): “Moral Hazard and Observability,” *Bell Journal of Economics*, 10, 74 -91.
- [9] Jewitt, I., Kadan, O., and J. M. Swinkels (2008): “Moral Hazard with Bounded Payments,” *Journal of Economic Theory*, 143, 59-82.
- [10] Mirrlees, J. (1976): “The Optimal Structure of Incentives and Authority within an Organization,” *Bell Journal of Economics*, 7, 105-131.
- [11] Mirrlees, J. (1999): “The Theory of Moral Harzard and Unobservable Behavior: Part I,” *Review of Economic Studies*, 66, 3-21.
- [12] Mishel, A. and L. Davis (2014): “CEO Pay Continues to Rise as Workers Are Paid Less,” Report, Economic Policy Institute, Washington DC., U.S.
- [13] Kaplan, S. N. (2013): “Executive Compensation and Corporate Governance in the U.S.: Perceptions, Facts and Challenges,” *Journal of Applied Corporate Finance*, 25(2), 8-25.
- [14] Levin, J. (2003): “Relational Incentive Contracts,” *American Economic Review*, 93(3): 835-857.
- [15] Malcomson, J. M. (2013): Relational Incentive Contracts. In R. S. Gibbons and J. Roberts (Eds.). *The Handbook of Organizational Economics*. Princeton, NJ: Princeton University Press.
- [16] Murphy, K. J. and J. Zabojsnik (2004): “CEO Pay and Appointments: A Market-Based Explanation for Recent Trends,” *American Economic Review Papers and Proceedings*, 94, 192-196.
- [17] Pearce, D. and E. Stacchetti (1998): “The Interaction of Implicit and Explicit Contracts in Repeated Agency,” *Games and Economic Behavior*, 23, 75-96.
- [18] Sattinger, M. (1993): “Assignment Models of the Distribution of Earnings,” *Journal of Economic Literature*, 31(2), 831-80
- [19] Teulings, C. N. (1995): “The pay Distribution in a Model of the Assignment of Skills to Jobs,” *Journal of Political Economy*, 103(2), 280-315.
- [20] Tervio, M. (2008): “The Difference that CEOs Make: An Assignment Model Approach,” *American Economic Review*, 98(3), 642-668.
- [21] Teschl, G. (2012): *Ordinary Differential Equations and Dynamical Systems*, American Mathematical Society, Graduate Studies in Mathematics, Volume 140