Executive Job Matching: Estimates from a Dynamic Model

Ruoyan Huang* †

December 12, 2016

Abstract

I evaluate how firms' and CEOs' learning about their fit with one another affects CEO turnover and compensation decisions. Building and estimating a dynamic model of the executive labor market, I find that learning and selection eliminate low-quality matches and provide explanatory power for the excess skewness of CEO compensation in the data after controlling for firm and CEO characteristics. I further establish that learning generates a hump-shaped hazard rate curve of CEO turnover conditional on CEO tenure. Using a hand-collected dataset of CEO turnover, I discover that the speed and precision of learning determine the level and length of the "discovery phase" of the conditional hazard rate curve. I also find that CEO compensation demonstrates a firm's evaluation of the match quality and is predictive of the expected future tenure of its CEO. In short, I demonstrate the importance of learning and selection in explaining the relations between CEO compensation, CEO turnover, and firm performance.

Keywords: Learning, CEO compensation, CEO turnover, Structural Estimation

JEL Classification: G30, G34, J30

---

*School of Economics and Finance, the University of Hong Kong. Email: ryhuang@hku.hk

†I would like to express my special appreciation to my advisors, Toni Whited and Ron Kaniel, for their continuous support and encouragement. I thank my thesis committee members, Olga Itenberg and Erin Smith, for their invaluable guidance and tremendous help. I would also like to thank all seminar participants at Hong Kong Joint Finance Research Workshop, University of Rochester, Univeristy of Toronto, University of Minnesota, Hong Kong University of Science and Technology, University of Hong Kong, University of Melbourne for helpful discussions and comments. All remaining errors are my own.

1. Introduction

A recent survey of the AFL-CIO[1] finds that the average CEO-to-worker pay ratio increased 91% from 1993 to 2014. Strikingly, the survey also demonstrates a significant dispersion in the CEO-to-worker pay ratios. For example, the CEO of PepsiCo, Inc, Indra K. Nooyi, is paid twice as much as the CEO of the Coca-Cola Co, Muhtar Kent, in 2013,[2] after controlling for the average compensation of their workers. The high level and dispersion of the CEO-to-worker pay ratio motivated the U.S. Securities and Exchange Commission (SEC) to release final rules of the pay ratio provision of Section 953(b) of the Dodd-Frank Wall Street Reform and Consumer Protection Act ("Dodd-Frank"), mandating a public company to "disclose the ratio of the compensation of its chief executive officer (CEO) to the median compensation of its employees."[3] Despite public and political concerns, most of the literature focuses on the absolute increase in CEO compensation and its relation to firm performance. Still, the question of why some CEOs are highly paid, after controlling for observable firm and CEO characteristics, remains unclear and needs further investigation.

In this paper, I depart from traditional studies of CEO compensation and focus on how firms and CEOs learn about their fit with one another. In particular, using a CEO-firm sample from 1994 to 2013, I first observe that, after controlling for observable firm and CEO characteristics, the excess skewness accounts for 65% of the skewness of CEO compensation. Then, I develop a model of the executive labor market in which learning and selection eliminate low-quality matches and generate a right-skewed distribution of CEO compensation without introducing any heterogeneity to firm and CEO characteristics. Bringing the model to the data, I find that, in addition to matching the empirical excess skewness of CEO compensation, I can quantify the relations between CEO compensation, turnover, and tenure that are consistent with their data counterparts. Moreover, the learning mechanism I propose is robust to a CEO moving between companies and explains 25% -

[1] "Executive Paywatch 2014" at http://www.aflcio.org/Corporate-Watch/Paywatch-2014.

[2] The AFL-CIO "Executive Paywatch 2013" reports that the hourly compensation of the PepsiCo. Inc CEO, Indra K. Nooyi, equals 638 hours of the average worker's pay at PepsiCo, and the hourly compensation of the Coca-Cola Co. CEO, Muhtar Kent, equals 320 hours of Coca-Cola's average worker's pay.

[3] http://www.sec.gov/news/statement/statement-on-open-meeting-on-pay-ratio-aguilar.html.

62% of the skewness found by Gabaix and Landier (2008). Therefore, learning plays an important role in determining CEO turnover decisions and shaping CEO compensation arrangements.

To explain how learning affects CEO turnover decisions and the consequent distribution of CEO compensation, I model the executive labor market where CEOs and firms are matched. Upon matching, firm productivity is determined by the unobservable match quality and an idiosyncratic productivity shock. Then, the matched firm and CEO update their beliefs about their match quality, and decide when to separate. At the same time, the firm also compensates its CEO conditional on the belief.

Even though firm productivity fluctuates in the model, the belief about the match quality represents the learning history and can be used to determine the future prospect of a match. Comparing the continuing payoffs of staying and separating, a CEO-firm match only dissolves when their beliefs decline to a cutoff. Reflected in the equilibrium, only CEOs and firms with low beliefs separate, skewing the belief distribution to the right. More importantly, the equilibrium CEO compensation is also belief-contingent, weighing the CEO's opportunity cost with her claim on the expected firm productivity and the option value of the match. As a result, the distribution of CEO compensation inherits the shape properties of the belief distribution and is, thus, right-skewed. In other words, highly paid CEOs emerge due to belief-contingent turnover and compensation decisions that are endogenously determined in the equilibrium.

To validate the model, I develop two testable hypotheses that are unique to the learning mechanism. In the first hypothesis, I posit that the hazard rate curve of CEO turnover increases in the first few years of CEO tenure and then gradually decays, featuring a "discovery phase" and a "decaying phase." The hump shape of the conditional hazard rate curve arises because CEOs and firms steadily reveal their "true" match quality from productivity realizations. Equally important, the learning mechanism also indicates that the signal quality of the productivity measure affects the speed and precision of learning. Therefore, the noisier the firm productivity measure, the slower and less accurately the firm and CEO learn, the longer the "discovery phase," and the less skewed the distribution of CEO compensation.

In the second hypothesis, I relate CEO compensation to CEO tenure continuation. Precisely, I posit that higher CEO compensation reflects a stronger belief that the CEO-firm match is a "good" fit and increases the likelihood of the CEO remaining with the firm. Additionally, the marginal effect of learning diminishes as low-quality matches are gradually eliminated. Therefore, an increasing and concave correlation is identified between CEO compensation and the probability of the CEO staying in the current match.

To quantify the model, I estimate the structural parameters of the model to match the key properties of CEO compensation, firm profitability, and CEO turnover rate in the data. Parameter estimates are used to characterize the executive labor market. For example, consistent with Gabaix and Landier (2008) and Pan (2010), I find that an average CEO only captures 1.25% of the firm value in each period when splitting the total matching payoffs with her matched firm. Then, I quantify both model hypotheses with parameter estimates and find strong empirical support without directly matching to their data counterparts. For example, both the simulated and actual hazard rate curves of CEO turnover peak at 3-4 years of CEO tenure and then steadily decay. Similarly, I also find that CEO compensation is predictive of the likelihood of tenure continuation in an increasing and concave way in both cases. Coefficient estimates from the actual and simulated data imply that the likelihood of a CEO remaining with a firm rises by 1.2% - 1.3% when CEO compensation increases by 1%. Therefore, the learning mechanism I propose can be used to explain the empirical relations between CEO compensation, turnover, and tenure.

Estimating model parameters also allows me to quantify the impacts of noise in the firm performance measure on CEO turnover and its sensitivity to firm performance. Specifically, I find that a 100% increase in the productivity shock decreases the CEO turnover rate by 65%, increases the length of the "discovery phase" of the hazard rate curve by two years, and lowers the turnover-performance sensitivity by 45%. These results imply that the signal quality of the firm performance measure affects the speed and precision of learning. Moreover, I also observe that 46% - 61% of the skewness in CEO compensation and 16% - 35% of the total welfare of the executive labor market are directly attributable to learning and selection. At the same time, without introducing any skewness to CEO or firm characteristics, learning and selection still explain 25% - 62% of the skewness in CEO compensation generated from Gabaix and Landier's (2008) competitive assignment model. These results further imply that belief-contingent turnover and compensation are critical in shaping the distribution of CEO compensation and improving the efficiency of the executive labor market.

Departing from conventional methods,[4] this paper supplements the CEO compensation literature attributing CEO and firm characteristics to the right-skewed distribution of CEO compensation.[5] In particular, I demonstrate that the learning-induced skewness is widespread and reconcilable to the executive matching process. Two papers, Taylor (2013) and Nickerson (2014), are closely related to mine as the authors quantify the implications of executive matching on CEO compensation. However, my paper diverges from each in several key aspects. First, Nickerson (2014) focuses on the pay-size elasticity derived from an inelastic supply of CEO labor and ignores the learning channel. In sharp contrast, in my paper, learning about the CEO-firm fit has a first-order influence in shaping CEO compensation and turnover. Second, Taylor (2013) emphasizes learning as an essential element in determining CEO compensation. However, without introducing the initial matching process in the executive labor market, he assumes an exogenous process for CEO compensation. Aiming to explain the excess skewness of CEO compensation, my model originates from the executive labor market, preserving the learning implications from Taylor (2013) but featuring the endogenous separation and the consequent cross-sectional distribution of CEO compensation. Therefore, my investigation extends the perspectives of the previous literature.

This paper also contributes to the empirical literature on CEO turnover. First, Jenter and Kanaan (2015) criticize the literature for investigating CEO compensation and turnover using different economic mechanisms. Notably, I directly model CEO turnover as "an extreme case of pay-for-performance" and explain the relation between CEO turnover and compensation. Second, Gao

[4]Theoretically, I apply the labor search literature (Mortensen and Pissarides, 1994, 1999; Jovanovic, 1979, 1984; Moscarini, 2003, 2005) to a CEO compensation setup. Empirically, I estimate model parameters via the simulated method of moments (SMM) following Hennessy and Whited (2005, 2007).

[5]Researchers provide numerous explanations of highly-paid CEOs in relation to firm size (Gabaix and Landier, 2008; Terviö, 2009), demand for general managerial skills in large corporations (Murphy and Zabojnik, 2004; Frydman, 2005), managerial entrenchment (Bebchuk and Fried, 2004; Kuhnen and Zwiebel, 2008), risks associated with incentive compensation (Hall and Liebman, 1998; Hermalin, 2005), etc.

et al. (2014) question the literature on CEO entrenchment[6] and show that public firms have higher turnover rates and greater turnover-performance sensitivities compared to private firms in the U.S. In this paper, I quantify the magnitude of CEO turnover through the cross-sectional properties of CEO compensation and investigate the variations of turnover-performance sensitivities with respect to changes in market conditions. As a result, my results provide new insights into CEO turnover and relate it to CEO compensation and firm performance.

The rest of the paper is organized as follows: I present the model in Section 2 and describe the data and the identification strategy in Section 3. In Section 4, I discuss the estimation results, investigate model hypotheses, and conduct counterfactual analyses. In Section 5, I extend the model to subsamples and provide robustness checks. I conclude in Section 6.

2. The Dynamic Model of the Executive Labor Market

In this section, I introduce the dynamic model featuring the executive matching process and its consequences on CEO turnover and compensation. Under a discrete time and infinite horizon scheme, I first define the matching environment and its outcome, as well as their joint impacts on firm productivity. Second, I characterize the learning process of a CEO-firm pair and outline the choice set and value function of each agent. Last, I define a unique stationary general equilibrium of the executive labor market.

2.1. The Matching Environment

In this paper, the central focus is the labor market for CEOs. I assume that, on the demand side, there is a continuum of ex ante homogeneous risk-neutral firms of measure F, actively hiring CEOs. Meanwhile, on the supply side, there is a continuum of ex ante homogeneous risk-neutral individuals,[7] normalized to a measure of 1, searching for CEO positions. In each period, new

[6] Bebchuk and Fried (2004) promote the prevalent view that CEO turnover is insensitive to firm performance. Taylor (2010) decomposes the total CEO turnover cost to the real cost of shareholders and the effective personal cost to the board. He concludes that the board "behaves as if firing the CEO costs shareholders 5.9% of assets, whereas it really only costs shareholders 1.3%."

[7] The risk-neutrality assumption simplifies the model. It also captures the idea that individuals who have the expertise to be executives are wealthy individuals in general. They have access to multiple saving and investment opportunities

matches emerge from the market. While being matched, the match-specific productivity of CEO-firm pair i, μ_i,[8] summarizes the match quality and only takes two discrete values $\{\mu_H, \mu_L\}$ in the space. Particularly, μ_H indicates that the match between the CEO and the firm is of high quality, while μ_L ($\mu_L < \mu_H$) suggests that the match is of low quality.

Although in the model there is no information asymmetry between the CEO and the firm inside a match, the quality of the match is unknown to both sides. Consequently, the CEO, along with the matched firm, infers their match type through the direct impact on firm productivity. However, two sources of idiosyncratic shock prevent both the firm and the CEO from perfectly inferring their match type. First, a normally distributed idiosyncratic shock, $z_{it} \sim^{i.i.d} N(0, \sigma_z^2)$, affects firm productivity in each period. If I denote firm i's productivity in period t as x_{it}, then:

$$x_{it} = \mu_i + z_{it}. \tag{1}$$

Second, a separation shock, arriving at an exogenous rate of $\delta \in (0,1)$, randomly dissolves existing CEO-firm matches. In reality, the exogenous turnover shock could be from the demand side and be due to merges and acquisitions, bankruptcies, buyouts, spinoffs, etc. The shock could also emerge from the supply side, as CEOs may leave their positions due to spousal relocations, human capital shocks, health reasons, retirements, deaths, etc. Accordingly, I model the executive labor market conditional on these observations.

2.2. *The Evolution of Belief*

Although the performance of a CEO-firm match is driven by idiosyncratic uncertainties, the CEO and the firm form and update their beliefs about their fit. In the model, immediately upon matching, I assume that the CEO and the firm share a common prior belief, $p_0 = \Pr(\mu = \mu_H) = 1 - \Pr(\mu = \mu_L)$, independent of their employment histories. Then, as the match continues, they

equivalent to high outside options. Without the risk-neutrality assumption, the model needs to distinguish alternative compensation components, which is beyond the scope of this paper.

[8] The match-specific productivity μ_i does not exclude the innate ability of a CEO. In the general case, I can re-define $\mu_i = \theta_i + a_i$, where θ_i is the *ex-ante* unobservable part attributable to match, and a_i is the *ex-ante* observable individual innate ability transferable among positions.

update their posterior beliefs following Bayes' rule. Specifically, suppose that the ith CEO-firm pair enters period t with a posterior belief $p_{it} = \Pr(\mu_i = \mu_H | \mathscr{F}_{t-1})$. After observing the productivity realization x_{it}, they update their beliefs following:

$$p_{i,t+1} = \Pr(\mu_i = \mu_H | \mathscr{F}_t) = \frac{p_{it}\Phi(x_{it}|\mu_H)}{p_{it}\Phi(x_{it}|\mu_H) + (1-p_{it})\Phi(x_{it}|\mu_L)}, \tag{2}$$

where $\Phi(x_{it}|\cdot)$ is the distribution function of firm productivity. Then, both sides choose whether to leave or stay conditional on the updated belief $p_{i,t+1}$. Therefore, the posterior belief about the match quality indicates the productivity history and determines whether the match will continue.

2.3. *Value Functions*

In this section, I define the optimization problem of each agent in the executive labor market. The posterior belief p_{it}, along with firm productivity x_{it}, defines the state of a CEO-firm pair and characterizes their optimization problems. From here forward, I drop the subscript "*it*" for simplication.

2.3.1. *The Supply Side*

Let $W(p,x)$ be the value function of a matched CEO, with posterior belief p and productivity realization x. Also, let U be the value function of an unmatched individual. In each period, while the matched CEO receives $w(p,x)$ as compensation, the unmatched individual only holds her outside wealth b. To make the match quality non-trivial, I further assume that the outside wealth b satisfies the following condition: $\mu_L < b < p_0\mu_H + (1-p_0)\mu_L$, indicating that a CEO should always accept a new match and terminate a "bad" one.

In each period, the matched CEO decides whether to leave ($d^{CEO} = 1$) or stay ($d^{CEO} = 0$) in the current match conditional on the updated belief p'. The Bellman equation of a utility-maximizing CEO equals:

$$W(p,x) = w(p,x) + \frac{1-\delta}{1+r} \max_{d^{CEO} \in \{0,1\}} \left\{ U, \mathbb{E}_{x'}[W(p',x')] \right\} + \frac{\delta}{1+r} U, \tag{3}$$

where $\mathbb{E}_{x'}[W(p',x')]$ is the expected payoff of the matched CEO in the following period if she remains in the current match. Then, equation (3) implies that the matched CEO will separate ($d^{CEO} = 1$) with her current firm if and only if the expected payoff of leaving is no less than that of staying ($U \geq \mathbb{E}_{x'}[W(p',x')]$). Otherwise ($U < \mathbb{E}_{x'}[W(p',x')]$), the CEO stays ($d^{CEO} = 0$) and carries the updated belief p' to the next period. The risk-free rate r, together with the exogenous turnover rate δ, discounts the continuing part of the value function.

The value function of the unmatched individual is state independent and equals:

$$U = b + \frac{1}{1+r}\{\lambda \mathbb{E}_{x'}[W(p_0,x')] + (1-\lambda)U\}, \tag{4}$$

where $\lambda \in (0,1)$ is the job-finding rate determined by the equilibrium matching technology. Equation (4) indicates that the belief about the CEO-firm fit will be reset to its prior value p_0 when a new match emerges.

2.3.2. *The Demand Side*

Analogously, let $J(p,x)$ be the value function of a matched firm, with posterior belief p and productivity realization x. Also, I define V as the value function of an idle firm. In each period, after compensating its CEO, the matched firm obtains $r(p,x)$ as profitability. Meanwhile, the idle firm, receiving no profit in period t, incurs a search cost $\kappa > 0$. Loosely interpreting, κ could be the direct cost to search and groom the successor, or be the difference in profitabilities between the matched and the idle firms.

The matched firm also decides whether to turn over ($d^F = 1$) its current paired CEO or not ($d^F = 0$) conditional on the updated belief p'. The Bellman equation of a profit-maximizing firm equals:

$$J(p,x) = r(p,x) + \frac{1-\delta}{1+r} \max_{d^F \in \{0,1\}} \left\{V, \mathbb{E}_{x'}[J(p',x')]\right\} + \frac{\delta}{1+r}V, \tag{5}$$

where $\mathbb{E}_{x'}[J(p',x')]$ is the expected payoff of the matched firm in the next period if the firm retains its current CEO. Then, equation (5) implies that the matched firm will only turn over ($d^F = 1$) its current CEO if and only if the continuing value of being idle is no less than that of staying in the current match ($V \geq \mathbb{E}_{x'}[J(p',x')]$). Otherwise ($V < \mathbb{E}_{x'}[J(p',x')]$), the firm retains its current CEO ($d^F = 0$) and carries the updated belief p' to the next period.

The value function of an idle firm is state independent and equals:

$$V = -\kappa + \frac{1}{1+r}\{q\mathbb{E}_{x'}[J(p_0,x')] + (1-q)V\}, \tag{6}$$

where $q \in (0,1)$ is the vacancy-filling rate in the equilibrium. To deter entries, I further assume that the executive labor market satisfies the "free-entry" condition $V = 0$. Under this assumption, no other firm will find it profitable to enter in any case.

2.3.3. The Nash Bargaining

CEO compensation and firm profitability are determined by a Nash bargaining over the matching surplus. Suppose that, in each period, the CEO takes a constant proportion $b^{NB} \in (0,1)$ of the surplus. I also assume that the firm retains μ_z from the expected matching surplus before paying its CEO.[9] Then, the Nash bargaining solves the following problem in each period:

$$w(p,x) \in \arg\max_{w(p,x)} [W(p,x) - U]^{b^{NB}} [J(p,x) - V]^{1-b^{NB}}, \tag{7}$$

where the necessary and sufficient first-order condition equals:

$$b^{NB}[J(p,x) - V] = (1 - b^{NB})[W(p,x) - U]. \tag{8}$$

Equivalently, let $S(p,x) = W(p,x) - U + J(p,x) - V$ be the net matching surplus. Then equation (8) shows that, in each period, the matched CEO receives a b^{NB} fraction of the matching rent, namely $W(p,x) - U = b^{NB}S(p,x)$. The matched firm claims the rest $J(p,x) - V = (1 - b^{NB})S(p,x)$. Thus, equation (8) specifies a linear rent-sharing rule between the CEO and the firm inside a match. More importantly, it derives functional forms of CEO compensation and firm profitability in Propo-

[9]The average retaining profitability μ_z enters into the realized profitability of the firm, but is excluded from the net matching surplus when the CEO bargains for compensation. In reality, it could be the retaining profitability for dividend payouts, internal cash flow, or any fixed cost to compensation the CEO.

sition 1:

Proposition 1. *The value function of the matched CEO $W(p,x)$ and the value function of the matched firm $J(p,x)$ are strictly increasing in the belief p. As a result, the Nash bargaining solves CEO compensation as:*

$$w(p,x) = (1-b^{NB})b + b^{NB}\{p\mu_H + (1-p)\mu_L + \frac{\lambda}{1+r}\mathbb{E}_{x'}[J(p_0,x')]\} \tag{9}$$

and firm profitability as:

$$r(p,x) = (1-b^{NB})[p\mu_H + (1-p)\mu_L - b] + \mu_z - b^{NB}\frac{\lambda}{1+r}\mathbb{E}_{x'}[J(p_0,x')]. \tag{10}$$

The implication of Proposition 1 is threefold. First, equilibrium turnover is a threshold decision, depending on the posterior belief p'. Therefore, belief-contingent turnover weeds out low-quality matches and skews the distribution of beliefs to the right. Second, in Proposition 1, I derive a belief-contingent CEO compensation scheme. In particular, equation (9) indicates that the compensation a CEO receives is an affine function of the belief p, weighing the CEO's outside wealth b with her bargaining share in the expected firm productivity and the option value of the match. Last, belief-contingent turnover and compensation decisions also indicate that the dispersion of beliefs uniquely determines the excess skewness of CEO compensation. Meanwhile, the observed CEO compensation provides a novel angle to quantify the influence of learning in the executive labor market.

2.4. The Stationary General Equilibrium

In this section, I specify a matching technology, a "free-entry" condition, and a measure over the state pair (p,x). Then I close the model with the definition of the stationary general equilibrium.

In the model, the matching process determines an individual's job-finding rate λ and a firm's vacancy-filling rate q. Without loss of generality, I assume that the matching technology is an increasing, concave, and linear homogeneous matching function $g(m, f)$, satisfying the Inada conditions such that:

$$g(m, f) = m^{\eta} f^{1-\eta}, \tag{11}$$

where m is the number of unmatched individuals, f is the number of idle firms, and $\eta \in (0, 1)$ is the matching elasticity with respect to unmatched individuals. Particularly, the "linear homogeneous" assumption suggests that the "tightness" of the executive labor market, defined as the number of vacant CEO positions per unmatched individual, $\theta = \frac{f}{m}$, is a sufficient statistic in determining the equilibrium. Henceforth, the job-finding rate λ of an unmatched individual is:

$$\lambda = \frac{g(m, f)}{m} = g(1, \frac{f}{m}) = g(1, \theta) = \theta^{1-\eta} \tag{12}$$

and the vacancy-filling rate q of an idle firm equals:

$$q = \frac{g(m, f)}{f} = g(\frac{1}{\theta}, 1) = \theta^{-\eta}. \tag{13}$$

Coupled with the "free-entry" condition $V = 0$, the matching technology also links the market "tightness" θ to the expected payoff of a new match:

$$\mathbb{E}_{x'}[J(p_0, x')] = (1+r)\kappa\lambda^{\frac{\eta}{1-\eta}} = (1+r)\kappa\theta^{\eta}. \tag{14}$$

Finally, let $h(p, x)$ be the measure of the mass over CEO-firm pairs with posterior belief p and productivity x, then there exits a general equilibrium in the executive labor market:

A Stationary General Equilibrium is a vector of scalars $\{\lambda^{\star}, q^{\star}, \theta^{\star}, m^{\star}, f^{\star}, F^{\star}\}$ and a tuple of functions $\{J^{\star}, W^{\star}, w^{\star}, r^{\star}, \underline{p}^{\star}, h^{\star}\}$ such that:

1. The tightness of the labor market equals:

$$\theta^{\star} = \frac{f^{\star}}{m^{\star}}.$$

2. $h^\star$ is an invariant distribution over the state space of the belief and the productivity (p,x).

3. $W^\star$ and $J^\star$ characterize the optimization problems of the matched CEO in equation (3) and the matched firm in equation (5). $w^\star$ and $r^\star$ solves the Nash bargaining problem in equation (8).

4. The separation policy $\underline{p}^\star(p,x)$ solves both the CEO's optimization problem $W^\star$ in equation (3) and the firm's optimization problem $J^\star$ in equation (5).

5. The "free-entry" condition in equation (14) uniquely specifies the relations between the total number of firms $F^\star$, the invariant distribution $h^\star$, and an individual's job finding rate $\lambda^\star$ in the equilibrium.

By a detailed proof in Appendix A.3, I show that:

Proposition 2. *The stationary general equilibrium of the executive labor market uniquely exists.*

3. Data and Identification

In this section, I briefly discuss the datasets and identification strategies I use to quantify the model.

3.1. Data

To quantify the dynamic model of the executive labor market, I create a sample from several sources. First, I collect financial statements from the Compustat North America Fundamental Annual files to merge with managerial compensation data from Execucomp. Following Erickson et al. (2014), I delete firms with fewer than \$2 million in total assets and require non-missing data for the main variables in Table 1. Likewise, I eliminate observations outside the $[0,1]$ interval for leverage, tangibility, and all compensation measures. Additionally, I winsorize compensation at its top and bottom 5%, along with firm profitability at the top and the bottom 2%, to eliminate sample outliers. Last, I require at least three consecutive observations for each CEO-firm pair. After all these steps, the Compustat-Execucomp sample contains $19,600$ CEO-year observations, ranging from fiscal year 1994 to 2013. It includes $2,714$ firms and $4,286$ CEOs, representing $4,353$ unique matching pairs.

[Insert Table 1]

Based on the Compustat-Execucomp sample, I calculate CEO tenure by subtracting the date the CEO left her position from the date she became the CEO. To further classify job flows in the executive labor market, I construct a unique CEO turnover sample with precise separating reasons. Specifically, I categorize each turnover case in the Compustat-Execucomp sample into exogenous or endogenous turnover by searching news and tracking the CEO's career path after her departure.[10] Appendix C lists the classification criteria, as well as the sample characteristics of the novel dataset.

CEO characteristics also affect the quantitative analyses. Accordingly, I construct another characteristics sample by collecting the education and external/internal/founder status for each CEO in the Compustat-Execucomp sample. Specifically, I first record the highest degree each CEO receives and assign numerical scores representing one of the six categories in Table 1. Likewise, I gather the external/internal/founder status following Graefe-Anderson (2014) for each CEO in my sample. First, I let the dummy variable $Founder = 1$ for each founder or co-founder CEO. Second, the dummy variable $Internal = 1$ if the CEO has already worked at the firm for more than three years or is a family member of the founder/owner of the firm at her promotion. Third, the dummy variable $External = 1$ if the CEO is from another firm or has worked less than three years in succession. The criterion of internal CEOs rules out "recent outsiders" who are groomed to be the successor of the ongoing CEO. Taking the education and status variables together, I form the characteristics sample of CEOs.

[10] The main data source for CEO departures and career news is Factiva dataset. I also use Google search, BusinessWeek, NNDB people search, Marquis who's who, Equilar Atlas, Finding universe, and other news sources to supplement the case-by-case analysis.

In the last step, I merge both the turnover and characteristics samples back into the Compustat-Execucomp sample to form the main sample. I also provide summary statistics of the main variables in Table 3. To quantify the excess skewness of CEO compensation, I control for data variations unattained in the model using the following regression:

$$y_{it} = \beta \times Control\ Variables_{it} + Year_t + Industry_i + \varepsilon_{it}, \tag{15}$$

where y_{it} equals firm profitability and CEO total compensation respectively, and $Control\ Variables_{it}$ includes:

$$\{Leverage_{it-1}, Tangibility_{it-1}, Market\ Capitalization_{it-1}, Gender_{it}, Interlock_{it}, External_{it}, Internal_{it}, Founder_{it}, Education_{it}\}.$$

I also control for the fiscal year ($Year_t$) and industry ($Industry_i$) fixed effects to eliminate institutional changes over time and variations across industries. Then, I construct two main variables, $Profit\_Residual$ and $Comp\_Residual$, by adding the population mean of firm profitability and CEO total compensation to the regression residuals of equation (15). These two variables are used to estimate the model parameters in Section 4.1. I also report their summary statistics in Panel B of Table 4.

[Insert Table 3]

[Insert Table 4]

Panel A of Table 3 and Panel B of Table 4 present the results of a characterization of the main sample from different aspects. First, after controlling for firm and CEO characteristics in Table 4, CEO compensation is still right-skewed. Moreover, the skewness of firm profitability is not fully explanatory of the skewness of CEO compensation. Second, the majority of CEO compensation comes from stock and option grants. Third, the firms in the sample are relatively large. For example, the average firm size is \$7 billion in total assets and \$6 billion in market capitalization. Fourth, on average, a CEO-firm match lasts for $9\frac{1}{2}$ years while half of the CEOs leave their positions within 7 years. Last, the decomposition infers that, on average, the exogenous turnover rate ($\approx 8\%$) is twice the magnitude of the endogenous turnover rate ($\approx 4\%$).

Panel B of Table 3 presents novel results of a characterization of the main sample from its compositions. First, among all the CEOs in my sample from 1994 to 2013, 30% hold bachelors degrees while 42% have Masters or other professional degrees. These results are consistent with Graham and Harvey (2001), suggesting that U.S. CEOs are well-educated. Second, 59% of the CEOs are corporate insiders at the time of succession, compared to 27% of external hires and 10% of founders. Last, 98% of the CEOs are male, having little influence over their compensation. Boards of directors are of moderate qualities represented by the G-index and E-index in my sample.

3.2. Identification

Key parameters of the model are estimated via simulated method of moments (SMM). However, I also estimate other parameters outside the model. First, I use the average 3-month Treasury bill rate, $r = 2.86\%$, to proxy for the risk-free rate. Second, I set the exogenous turnover rate at its sample average. For example, it equals 8.68% in the main sample. Third, I cannot separately estimate a firm's search cost κ and the matching elasticity η without the search lengths of firms and CEOs. Consequently, I assume $\eta = 0.5$ based on the $[0.5, 0.7]$ range from Petrongolo and Pissarides (2001). Last, following Hall and Murphy (2002), I assume that the average outside wealth of CEOs is \$5 million for the entire life span. As a result, the discounted annual value, scaled by the average of firm total assets, is approximately 0.

Table 2 lists the seven remaining parameters $\{\mu_H, \mu_L, p_0, \mu_z, \sigma_z, b^{NB}, \kappa\}$. The validity of the SMM estimation critically depends on the choice of the moments that are sensitive to changes in the underlying parameters. I select 11 moments that feature the key properties of firm profitability, CEO compensation, and CEO turnover in the model. Although the parameters have intertwining impacts, I categorize them into three groups to emphasize the economic mechanisms they represent.

[Insert Table 2]

The first group of moments includes the means and standard deviations of firm profitability and CEO compensation. These moments impose constraints on the match-specific productivities

μ_H and μ_L, as well as their composition p_0. Also, the mean and standard deviation ratios between CEO compensation and firm profitability distinguish the average bargaining power b^{NB} of a CEO. Most importantly, the degree of skewness of CEO compensation identifies the informativeness of firm performance, represented by the signal-to-noise ratio $\frac{\mu_H - \mu_L}{\sigma_z}$. Figure 1 shows that a higher signal-to-noise ratio improves the proportion of high-quality matches in the equilibrium. Reflected in the compensation distribution, the degree of skewness also rises as learning weeds out a larger proportion of low-quality matches.

[Insert Figure 1]

I then use the AR(1) regression coefficient ρ and the residual deviation $\mathbb{E}[\varepsilon_{it}^2]$ in:

$$y_{it} = \rho y_{it-1} + \varepsilon_{it} \tag{16}$$

to distinguish the mean of the retaining productivity μ_z and the standard deviation of the firm's productivity shock σ_z, where y_{it} is firm i's profitability in period t. These two moments separate the invariant and the changing parts in firm productivity and monotonically map onto μ_z and σ_z. When estimating equation (16), I follow Han and Phillips (2010) to treat firm fixed effects with the differencing-based estimator.

The last set of moments relates to the endogenous separation between CEOs and firms. First, I use the endogenous turnover rate to assess the firm's search cost κ. Then, I estimate a linear probability model (LPM):

$$End\_Turnover_{it} = \alpha + \beta_1 \Delta y_{it} + \beta_2 log(Tenure_{it}) + \beta_3 \big(\Delta y_{it} \times log(Tenure_{it})\big) + \varepsilon_{it}, \tag{17}$$

where Δy_{it} is the absolute change in firm profitability from period $t-1$ to t. Precisely, β_1 defines a firm's willingness to replace its current CEO per unit of profitability change and isolates the search cost κ and the bargaining power $1 - b^{NB}$ from other parameters. Then, both β_2 and β_3 identify the dynamic feature of learning through the signal-to-noise ratio $\frac{\mu_H - \mu_L}{\sigma_z}$ along CEO tenure.

4. Results

In this section, I bring the model to the main sample constructed in Section 3.1 and present the quantitative results. I first discuss the parameter estimates of the model and their implications for the executive labor market. Then, I develop two learning hypotheses from the model and test them in the data. Finally, I perform two types of counterfactual exercises to quantify the impacts of the learning mechanism on CEO turnover and the importance of the learning mechanism in shaping CEO compensation.

4.1. Benchmark Result

I first estimate the benchmark model and evaluate its implications for the executive labor market. Panel A of Table 5 presents the results from moment matching. On the positive side, the model provides a good fit to the actual CEO compensation, persistence in firm profitability, and the CEO turnover rate. In particular, the simulated skewness agrees with the actual dispersion in CEO compensation. Similarly, simulated coefficient estimators $\{\beta_1, \beta_2, \beta_3\}$ of the LPM in equation (17) capture the empirical observation that a CEO is less likely to leave with higher firm profitability and longer CEO tenure. On the negative side, the model cannot match the mean and the standard deviation of firm profitability. One possible explanation is that I fail to control for some unobservable factors in the data that introduce noises to firm performance. Overall, an over-identification test fails to reject the model with a p-value of 0.5923.

[Insert Table 5]

Panel B of Table 5 lists the structural parameter estimates. Three parameters are important in describing the executive labor market. First, an average CEO only obtains 1.23% of the total payoffs from matching, equivalent to 1.25% of the average firm value,[11] when bargaining with the matched firm. The estimator is close to the 2% of firm value in Gabaix and Landier (2008) but differs from others such as the 20% in Terviö (2009). The comparison implies that market imper-

[11] The calculation follows: $1.23\% \times S(p,x) = \frac{1.23\%}{1-1.23\%} \times J(p,x) = 1.25\% \times J(p,x)$.

fection in this model reduces CEO bargaining power to a similar extent as by the extreme value theory in Gabaix and Landier (2008). Second, it costs a firm 23.20% of the average profitability, equaling 1.27% of the average firm's total assets,[12] to replace its CEO. Even though there is no direct evidence from the data, Taylor (2010) does find that it costs the shareholder 1.3% of firm assets to dismiss a CEO. This estimated $70 million could include fees paid to a third-party recruiter[13] and losses to firm value during the transition period. Last, only 29% of the matches are of high quality. This observation suggests that a CEO's innate ability, such as education, would be a poor proxy for the match-specific quality in the empirical research.

4.2. A Hump-shaped Hazard Rate Curve of CEO Turnover Conditional on CEO Tenure

In this section, I develop and test the first learning hypothesis on the relation between CEO turnover and tenure. Jovanovic (1979) proposes a single peak in the hazard rate curve conditional on tenure for the general labor market. Jenter and Lewellen (2014) observe that the performance-driven CEO turnover rate initially increases and then gradually decays with CEO tenure. The learning mechanism in my model also provides explanatory power for a hump-shaped hazard rate curve of CEO turnover conditional on CEO tenure. Specifically, in the first few years of CEO tenure, the turnover rate increases because CEO-firm pairs learn their match quality through productivity realizations. Then, the turnover rate declines when the selection effect dominates for longer CEO tenure. Therefore, I define the initial increasing part of the conditional hazard rate curve as its "discovery phase," and the declining part of the curve as its "decaying phase."

To test this hypothesis in the actual data, I plot the actual versus the simulated endogenous turnover rates for seven CEO tenure intervals (in years): 1-2, 3-4, 5-6, 7-8, 9-11, 12-16, and 17+ defined by Jenter and Lewellen (2014) in Panel A of Figure 2. The plot shows that my model generates a simulated hazard rate curve that follows the pattern of the actual curve along CEO tenure. In particular, without matching the turnover rates, both hazard rate curves first increase,

[12]The estimate of κ is the difference between the profitability of the matched and idle firms. As a result, it equals to $0.2320 - 0.1438 = 0.0882$ of firm profitability, equalling to $0.1438 \times 0.0882 = 0.0127$ of the average firm's total assets of the sample.

[13]For example, when Microsoft searched for its recent CEO, the market estimated that Microsoft paid \$0.4 million to Heidrick & Struggles

peak at 3-4 years of CEO tenure, and gradually decline. Notably, Jenter and Lewellen (2014) find it hard to reconcile a hump-shaped conditional hazard rate curve with a learning model of constant match quality. Yet, I demonstrate that, by modeling the matching process of the executive labor market with search friction, the speed of learning in the equilibrium provides explanatory power to the delayed response of a firm in evaluating CEO performance and making separation decisions. Therefore, I find strong empirical evidence on the relation between CEO turnover and tenure that is consistent with the learning hypothesis inferred from my model.

[Insert Figure 2]

4.3. Predicting CEO Tenure Continuation

The second learning hypothesis focuses on the relation between CEO compensation and the likelihood of CEO tenure continuation. Precisely, the model implies that higher CEO compensation represents a stronger belief that the CEO-firm match is of high quality and raises the likelihood of the CEO remaining in the current match. Furthermore, the marginal effect of learning is diminishing as low-quality matches are gradually eliminated. Hence, I posit that, unconditional on CEO tenure, the likelihood of a CEO remaining with her matched firm is increasing and concave in her current compensation.

To test this hypothesis, columns (1) - (4) of Table 6 report coefficient estimates from a LPM on the actual data:

$$1 - END\_Turnover_{it} = \alpha + \beta_1 Comp_{it} + \beta_2 Comp_{it}^2 + \gamma Controls + \varepsilon_{it}, \tag{18}$$

where $Controls$ includes subsets of $\{Leverage_{it-1}, Tangibility_{it-1}, Internal_{it}, External_{it}, Founder_{it}, Gender_{it}, Interlock_{it}, Education\ Dummies_{it}$, fiscal year and industry fixed effects$\}$. Comparatively, I also estimate the same LPM without $Controls$ on the simulated data in column (5) of Table 6.

[Insert Table 6]

The results in Table 6 confirm that CEO compensation significantly relates to the probability of tenure continuation in an increasing and concave way after controlling for firm and CEO characteristics in the actual data. Quantitatively, a 1% increase in CEO compensation raises the likelihood of the CEO remaining with the firm by 1.2% - 1.3% in both the simulated and the actual data. Remarkably, although I do not include these regression coefficients in my moment matching process, the estimated regression coefficients from the simulated sample still closely follow their counterparts in the actual data. Therefore, the results in Table 6 emphasize that CEO compensation reflects a firm's evaluation of its fit with the CEO and indicates the likelihood of match continuation.

4.4. Counterfactual Analyses

Above I develop two testable hypotheses from the model and find strong empirical support in the actual data. Therefore, the learning mechanism in my model provides explanatory power for the empirical relations between CEO compensation, turnover, and tenure. In this section, I conduct counterfactual analyses to further explore their relations. In particular, I address two questions: How does noise in the firm performance measure affect CEO turnover decisions through learning? And how important is the learning mechanism in shaping CEO compensation?

To answer the first question, I evaluate the influence of the idiosyncratic productivity shock σ_z on CEO turnover rate. In the model, I conjecture that the noisier the firm performance measure is, the slower and less accurately the CEO and the firm learn their match quality, and the less efficient the executive labor market in dissolving low-quality CEO-firm matches. To quantify the learning mechanism above, I first plot the hazard rate curves of CEO turnover conditional on CEO tenure by increasing the value of the noise parameter σ_z in Panel B of Figure 2. Consistent with Engel et al. (2003) and Farrell and Whidbee (2003), I also find that the endogenous turnover rate of CEOs decreases at all CEO tenure intervals when firm performance becomes noisier. On average, CEO turnover rate reduces by 43% when increasing σ_z by 50%, and decreases by 65% with a 100% increase in σ_z. Crucially, I also identify a positive correlation between the noise measure σ_z and the length of the "discovery phase" of the conditional hazard rate curve. For example, the hump of the hazard rate curve tilts from 3-4 years of CEO tenure to 5-6 years and then 9-11 years when increasing σ_z by 50% and 100% respectively. Consequently, I find that the length of the “discovery phase” of a conditional hazard rate curve represents the speed of learning and, thus, contains information about the effectiveness of the executive labor market in dissolving low-quality CEO-firm matches.

Next, I evaluate the impact of the noise in firm performance on turnover-performance (TP) sensitivity. Recently, Bushman et al. (2010) examine the conjecture that risks unrelated to the CEO decrease a firm’s ability to fire its CEO efficiently through TP sensitivity measures. Similarly, I also conduct a counterfactual analysis to estimate the TP sensitivities from the following model:

$$END\_Turnover_{it} = \alpha + \beta_1 Profitability_{it} + \beta_2 log(Tenure_{it}) + \beta_3 \left(Profitability_{it} \times log(Tenure_{it})\right) + \varepsilon_{it}. \tag{19}$$

Columns (1) - (3) in Table 7 present the estimation results at the benchmark value of σ_z. I then repeat the analysis by increasing σ_z by 50% (columns (4) - (6)) and 100% (columns (7) - (9)). The results in Table 7 verify a negative association between firm profitability and the likelihood of CEO turnover. More importantly, a comparison across columns suggests that the TP sensitivity coefficient β_1 is strictly decreasing in σ_z for all specifications. For example, controlling for CEO tenure and an interaction term, the TP sensitivity drops by 45% when increasing σ_z by 100%. Therefore, the results in Table 7 emphasize the learning implication from my model and provide evidence consistent with Bushman et al.’s conjecture that "the probability of CEO turnover is decreasing in the variance unrelated to CEO talent, holding firm performance and variation over CEO talent constant.” The results also provides a clean quantification of the learning impact on the effectiveness of separation between CEOs and firms without endogeneity or measurement error concerns.

[Insert Table 7]

I conduct two additional counterfactual analyses to address the second question of the importance of the learning mechanism. In the first counterfactual analysis, I shutdown the belief-contingent selection channel and directly assess its consequences on the skewness of CEO compensation and the total welfare of the executive labor market. Particularly, I fix the equilibrium separation rate and re-estimate the model by randomizing the endogenous separation policy. Columns

(1) and (2) of Table 8 show that the skewness of CEO compensation decreases by 46% while the total welfare of the executive labor market lowers by 16%. These results indicate that the belief-contingent selection is non-trivial in providing explanatory power to CEO compensation dispersion and improving the efficiency of the executive labor market.

[Insert Table 8]

In the second counterfactual analysis, I evaluate the relative contribution of learning and selection as it relates to CEO compensation skewness compared to a competitive assignment model. This exercise is motivated by Gabaix and Landier (2008), who emphasize that firm size, when assortatively matched with CEO talent, plays a key role in determining the cross-sectional CEO compensation arrangement.

To provide a fair comparison, I replicate Gabaix and Landier's pay-size regression using my sample[14] and find a consistent elasticity estimator of 0.3896. I then use the actual firm size in my sample, along with the elasticity estimator, to generate the compensation of the top 1,000 CEOs, denoted as *Comp_GL*, applying Gabaix and Landier's formula.[15] The skewness of *Comp_GL* represents the dispersion in CEO compensation driven by the competitive assignment model proposed by Gabaix and Landier (2008). I then simulate CEO compensation from my model as defined by the benchmark parameter estimates in Table 5. I assume that each CEO is assigned to a firm with the same size and take the first 1,000 highest compensation as the simulated CEO compensation, denoted as *Comp_Simul*. The skewness of *Comp_Simul* is the dispersion in CEO compensation purely driven by the learning mechanism in my model without introducing any heterogeneity to firm and CEO characteristics.

[14] To replicate, I first run the pay-size regression in Gabaix and Landier's equation (18):

$$\ln(w_{i,t+1}) = d + e \times \ln(S_{n_\star,t}) + f \times \ln(S_{i,t}) + \varepsilon_{it}$$

with my sample, where $w_{i,t+1}$ is the total compensation of CEO i in period $t+1$, $S_{i,t}$ is the size of firm i in period t, and $S_{n_\star,t}$ is the size of the reference firm in period t. In this case, f is the pay-size elasiticity defined by Gabaix and Landier (2008). I find the estimator equals to 0.3896, compared to a $0.37-0.39$ range from their regression specifications.

[15] In Proposition 2 (Level of CEO Pay in the Market Equilibrium), Gabaix and Landier posit that the compensation of the manager of index n runs a firm of size $S(n)$ equals $w(n) = D(n_\star)S(n_\star)^{\beta/\alpha}S(n)^{\gamma-\beta/\alpha}$, where $S(n_\star)$ is the size of the reference firm ($n_\star = 250$ in their specification), and $D(n_\star) = \frac{w(n_\star)}{S(n_\star)^\gamma}$. To generate *Comp_GL*, I follow their assortative matching assumption and use the top 1,000 CEOs by firm size in my sample as $S(n)$. Then I use the estimated pay-size elasticity and firm size and compensation of the 250th firm to calculate each $w(n)$ as *Comp_GL*.

I provide two cases to evaluate the relative contribution of the learning mechanism. In the first case, I plot a histogram of the skewness ratio, defined as $skewness(Comp\_Simul)/skewness(Comp\_GL)$, in Panel A of Figure 3 by repeating the simulation for $5,000$ times. The histogram shows that learning and selection still account for 25% of the skewness in CEO compensation compared to the mechasnim of matching different CEO talent levels to heterogeneous firm sizes in Gabaix and Landier (2008). Moreover, the 25% is only a lower bound because the learning mechanism could be amplified by heterogeneities in firm and CEO characteristics. Therefore, the learning mechanism in my paper is non-trivial in magnitude compared to other channels.

[Insert Figure 3]

To further emphasize the importance of the belief-contingent selection in shaping CEO compensation skewness, I extend the same analysis by allowing CEOs to perform on-the-job searches in the second case. In Appendix A.1, I present a full version of the benchmark model with on-the-job search. Qualitatively, on-the-job searches provide another mechanism through which CEOs can actively dissolve low-quality matches in the executive labor market. Reflected in the simulated moments in Table 9, I find that the mean of firm profitability significantly increases when a higher proportion of low-quality matches are eliminated. More importantly, column (3) and (4) of Table 8 show that the skewness of CEO compensation decreases by an additional of 15% when shutting down the learning channel to randomize CEOs' on-the-job search decisions. Under this circumstance, I conjecture that the skewness ratio will be higher when comparing the simulated compensation ($Comp\_Simul\_Otjs$) from a model with CEOs performing on-the-job searches to the compensation ($Comp\_GL$) generated from Gabaix and Landier (2008).

[Insert Table 9]

Panel B of Figure 3 shows that the relative explanatory power increases from 25% to 62% when allowing CEOs to actively dissolve their matches with firms. Remarkably, this 37% jump in the skewness ratio is entirely due to an extra selection channel based on the same information accumulated through learning from productivity realizations in the first case. Consequently, combining the first and the second cases, in the second counterfactual analysis, I compare the relative contribution to CEO compensation skewness from my model to that from Gabaix and Landier's (2008) competitive assignment model and demonstrate the importance of learning and selection in shaping CEO compensation.

5. Subsample and Robustness

In this section, I evaluate model implications across subsamples and discuss robustness tests.

5.1. Subsample Analyses

Although the learning mechanism is independent of firm and CEO characteristics in the model, there are wide cross-sectional and time series heterogeneities affecting the speed of learning and the strength of selection. In this section, I examine four subsamples: external CEOs, internal CEOs, pre-2006 and post-2006, to see if they provide explanatory power.

In the first analysis, I split the main sample into external and internal CEOs after excluding founder CEOs. The external subsample contains $5,288$ CEO-year observations, compared to $11,434$ observations for the internal subsample. If learning about the CEO-firm fit plays an essential role in shaping CEO turnover decisions, I hypothesize that the turnover rate will be lower for the internal CEOs since firms have already observed their matching outcomes before promotion. Additionally, I also conjecture that, on average, it takes a firm longer to discover its match quality with an outside CEO before the selection effect dominates. In Panel C of Figure 2, I plot the hazard rate curves of CEO turnover conditional on CEO tenure for both the external and internal subsamples and observe consistent evidence for both learning implications above. To illustrate, I show that the turnover rate in Panel C of Figure 2 is twice the magnitude for the external CEO subsample compared to the internal one at all CEO tenure intervals. In addition, the external CEO subsample exhibits a steeper and longer "discovery phase," indicating that it takes an average firm with an outside CEO two more years before the selection effect takes over.

Mapped onto parameter estimates in Table 10, I find that a slower learning speed partially cancels out the level effect of turnover in the external CEO subsample as the estimated signal-to-noise ratios $\frac{\mu_H - \mu_L}{\sigma_z}$ are extremely close for both subsamples. I also discover that the proportion of high-quality matches is 24% lower for firms with outside replacements. Therefore, learning about CEO-firm fit plays a key role in shaping CEO turnover decisions.

[Insert Table 10]

I also use the external/internal subsamples to evaluate whether greater CEO bargaining power leads to higher compensation to external CEOs. In line with Graefe-Anderson (2014), Figure 4 shows that the compensation of external CEOs is significantly higher compared to insider CEOs in the data. Reflected in parameter estimates, The results in Table 10 show that the external CEOs claim 2.31% of the total payoffs, compared to 1.27% for internal CEOs. Similarly, it also costs a firm an extra 46% of its profitability to replace the outside hire. As a result, the parameter estimates show that managerial power is crucial in determining the high compensation of external CEOs.

[Insert Figure 4]

In the second analysis, I divide the main sample into pre- and post-2006 periods. I use the fiscal year 2006 as a cutoff for three reasons. First, this cutoff generates two subsamples containing similar numbers of observations.[16] Second, Execucomp changed the definition of total compensation (TDC1) in 2006. Although the model is insensitive to different compensation measures, this analysis serves as a robustness check. Last and most important, the volatility of firm profitability is significantly higher for the post-2006 period, mainly due to the subprime mortgage crisis from 2007 to 2009. Therefore, consistent with the counterfactual analysis on CEO turnover in Section 4.4, I hypothesize that the turnover rate and skewness of CEO compensation will be lower for the post-2006 period.

I plot six moments: mean and standard deviation of firm profitability, mean and skewness of CEO compensation, endogenous turnover rate, and on-the-job search rate for both the pre- and

[16] The pre-2006 subsample contains $9,244$ CEO-year observations and the post-2006 subsample includes $10,356$ CEO-year observations

post-2006 subsamples in Figure 4. I find that the mean of CEO compensation remains the same for both the pre- and post-2006. However, the turnover rate and compensation skewness are significantly lower for the post-2006 period in both the actual and simulated data. This demonstrates that the model is able to explain the empirical associations between the volatility of firm performance, CEO turnover rate, and CEO compensation skewness. Therefore, I find consistent empirical evidence that a noisier firm performance measure decreases the effectiveness of the executive labor market in dissolving low-quality matches. Parameter estimates in Table 10 further confirm that firm productivity is less informative in the post-2006 period as the signal-to-noise ratio $\frac{\mu_H - \mu_L}{\sigma_z}$ is 9% lower. These estimates also suggest that during post-2006 period, CEOs have a 22% lower bargaining power and a 13% higher proportion of high-quality matches. The cost of replacing a CEO also decreases by 19% after 2006, along with a 32% drop in the relative switching rate. As a result, the overall effect on the efficiency of the executive labor market is mixed.

5.2. *Robustness Analyses*

Tests of robustness are discussed in this section. I first examine whether the results are sensitive to different components of CEO compensation. The first three rows of Table 11 indicate that the model fits alternative compensation measures equally well. Also, parameter estimators are insensitive to changes in compensation measures except for the search cost κ and the relative contact rate ψ. Interestingly, I find a surge in a firm's search cost by excluding incentive compensation. Additionally, other incentive compensation, such as retirement plan, deferred compensation, and golden parachutes, significantly lowers a CEO's willingness to change jobs by 70%.

[Insert Table 11]

Second, I split the main sample into four education subsamples based on the highest degree a CEO receives - High School Diploma/No School, Bachelor Degree, Master Degree, and Doctorate Degree. This analysis supplements the benchmark results by varying a CEO's innate ability.[17] The

[17] In the model, the innate ability of CEOs is defined as the *ex-ante* observable productivity of CEOs that is transferable and orthogonal to matching productivities. There is no direct measure of this variable. Following Spence's (1973)

results in Table 11 suggest that the model provides decent fit to each subsample using the over-identification test $J-$statistics. I also characterize the executive labor markets defined by CEO education levels from parameter estimates. For example, I find that CEOs with doctorate degrees have the highest bargaining power of 3%. Yet, those with master degrees most frequently perform on-the-job searches for CEO positions in other firms. Interestingly, the proportion of high-quality matches is negatively correlated with the education level defined for each subsample. This result indicates that caution should be used when using education to proxy for CEO-firm match quality.

In addition, I create MBA and JD subsamples to test whether the model holds for professional managers. The results in the last two rows of Table 11 imply that the model provides a decent description for these two subsamples. Parameter estimates also attribute low search cost and high contact rate to frequent separation and switching observed in the MBA subsample. The results reveal that the average bargaining power of CEOs with MBA degrees is 20% higher than the main sample.

6. Conclusion

In this paper, I examine the excess skewness of CEO compensation after controlling for firm and CEO characteristics. I emphasize the learning mechanism and the consequent belief-contingent selection in shaping CEO compensation. Quantitative analyses of the model generate five distinct results. First, learning and selection generate a right-skewed distribution of CEO compensation, which is robust to a CEO's capacity to move between companies. Second, the learning mechanism is non-trivial as it provides an explanatory power of 25% - 62% of the skewness in CEO compensation generated from Gabaix and Landier's (2008) competitive assignment model. It also improves the welfare of the executive labor market by 16% - 35%. Third, learning provides explanatory power for the empirical observation of a hump-shaped hazard rate curve of CEO turnover conditional on CEO tenure. Importantly, I can quantify the inference of learning speed on

signaling theory, I assume that the education level be the dominating factor, or at least the primary signal a firm choose to evaluate the CEO. As a result, variations in CEOs' highest degrees help to identify the impact of innate ability.

the CEO turnover decision through the length of the "discovery phase." Fourth, unconditional on CEO tenure, the likelihood of the CEO remaining with her matched firm is increasing and concave in CEO compensation. Fifth, I provide a clean quantification of the impact of noise in the firm performance measure on the CEO turnover rate and the turnover-performance sensitivity. With different compensation determinants, I take a similar stand on the matching process between CEOs and firms in the executive labor market as in Pan (2010) and Nickerson (2014) but instead employ the learning mechanism in Taylor (2013). In like manner, I rationalize the relation between CEO turnover and tenure in Jenter and Lewellen (2014) and quantify the influence of signal quality on turnover-performance sensitivity in Bushman et al. (2010). Most importantly, I address the concern in Jenter and Kanaan (2015) by interpreting turnover as "an extreme case of pay-for-performance," providing quantitative guidance as urged by Gao et al. (2014).

My study is particularly important with the SEC implementing the pay ratio provision of the "Dodd-Frank" Act. First, I directly address Bebchuk and Fried's (2004) critique by showing that economic models can reconcile CEO compensation arrangement to firm productivity even after controlling for observable firm and CEO characteristics. Second, I assess the efficiency of the labor market for executives, characterize its unique features, and demonstrate its sensitivity to the ability of executives to move between companies. To this end, I quantify the extent to which corporate governance matters.

One question remains unanswered in this paper. Without information asymmetry, I cannot address the learning implications on a CEO's risk-taking behaviors, nor on the compensation design in a principal-agent framework. Therefore, future studies of CEO compensation contract design with frictional searching and matching in the executive labor market will advance the understanding of corporate governance.

References

Aiyagari, S. R., 1994. Uninsured idiosyncratic risk and aggregate saving. The Quarterly Journal of Economics pp. 659–684.

Bebchuk, L. A., Fried, J. M., 2004. Pay without performance.

Bushman, R., Dai, Z., Wang, X., 2010. Risk and ceo turnover. Journal of Financial Economics 96, 381–398.

Engel, E., Hayes, R. M., Wang, X., 2003. Ceo turnover and properties of accounting information. Journal of Accounting and Economics 36, 197–226.

Erickson, T., Jiang, C. H., Whited, T. M., 2014. Minimum distance estimation of the errors-in-variables model using linear cumulant equations. Journal of Econometrics 183, 211–221.

Erickson, T., Whited, T. M., 2000. Measurement error and the relationship between investment and q. Journal of Political Economy 108, 1027–1057.

Farrell, K. A., Whidbee, D. A., 2003. Impact of firm performance expectations on ceo turnover and replacement decisions. Journal of Accounting and Economics 36, 165–196.

Frydman, C., 2005. Rising through the ranks. the evolution of the market for corporate executives, 1936-2003. V Columbia University .

Gabaix, X., Landier, A., 2008. Why has ceo pay increased so much? The Quarterly Journal of Economics 123, 49–100.

Gao, H., Harford, J., Li, K., 2014. Ceo turnover-performance sensitivities in private firms. Nanyang Technological University Working Paper .

Graefe-Anderson, R., 2014. Ceo turnover and compensation: An empirical investigation. The Quarterly Journal of Finance 4, 1450008.

Graham, J. R., Harvey, C. R., 2001. The theory and practice of corporate finance: Evidence from the field. Journal of Financial Economics 60, 187–243.

Hall, B. J., Liebman, J. B., 1998. Are ceos really paid like bureaucrats? The Quarterly Journal of Economics 113, 653–691.

Hall, B. J., Murphy, K. J., 2002. Stock options for undiversified executives. Journal of Accounting and Economics 33, 3–42.

Han, C., Phillips, P. C., 2010. Gmm estimation for dynamic panels with fixed effects and strong instruments at unity. Econometric Theory 26, 119–151.

Hennessy, C. A., Whited, T. M., 2005. Debt dynamics. The Journal of Finance 60, 1129–1165.

Hennessy, C. A., Whited, T. M., 2007. How costly is external financing? evidence from a structural estimation. The Journal of Finance 62, 1705–1745.

Hermalin, B. E., 2005. Trends in corporate governance. The Journal of Finance 60, 2351–2384.

Jenter, D., Kanaan, F., 2015. Ceo turnover and relative performance evaluation. The Journal of Finance .

Jenter, D., Lewellen, K., 2014. Performance-induced ceo turnover. Working Paper .

Jovanovic, B., 1979. Firm-specific capital and turnover. The Journal of Political Economy pp. 1246–1260.

Jovanovic, B., 1984. Matching, turnover, and unemployment. The Journal of Political Economy 92, 108.

Kuhnen, C. M., Zwiebel, J., 2008. Executive pay, hidden compensation and managerial entrenchment. Rock Center for Corporate Governance Working Paper 16.

Michaelides, A., Ng, S., 2000. Estimating the rational expectations model of speculative storage: A monte carlo comparison of three simulation estimators. Journal of Econometrics 96, 231–266.

Mortensen, D. T., Pissarides, C. A., 1994. Job creation and job destruction in the theory of unemployment. The Review of Economic Studies 61, 397–415.

Mortensen, D. T., Pissarides, C. A., 1999. Unemployment responses to 'skill-biased'technology shocks: The role of labour market policy. The Economic Journal 109, 242–265.

Moscarini, G., 2003. Skill and luck in the theory of turnover. Working paper.

Moscarini, G., 2005. Job matching and the wage distribution. Econometrica 73, 481–516.

Murphy, K. J., Zabojnik, J., 2004. Ceo pay and appointments: A market-based explanation for recent trends. American Economic Review pp. 192–196.

Nickerson, J., 2014. A structural estimation of the cost of suboptimal matching in the ceo labor market. Working paper.

Nikolov, B., Whited, T. M., 2014. Agency conflicts and cash: Estimates from a dynamic model. The Journal of Finance 69, 1883–1921.

Pan, Y., 2010. The determinants and impact of executive-firm matches. Unpublished Working Paper. University of Minnesota .

Petrongolo, B., Pissarides, C. A., 2001. Looking into the black box: A survey of the matching function. Journal of Economic Literature pp. 390–431.

Spence, M., 1973. Job market signaling. The Quarterly Journal of Economics pp. 355–374.

Tauchen, G., 1986. Finite state markov-chain approximations to univariate and vector autoregressions. Economics Letters 20, 177–181.

Taylor, L. A., 2010. Why are ceos rarely fired? evidence from structural estimation. The Journal of Finance 65, 2051–2087.

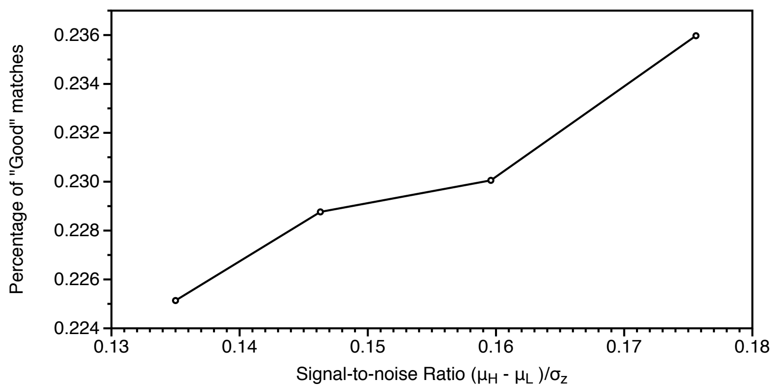
Taylor, L. A., 2013. Ceo wage dynamics: Estimates from a learning model. Journal of Financial Economics 108, 79–98.

Terviö, M., 2009. Superstars and mediocrities: Market failure in the discovery of talent. The Review of Economic Studies 76, 829–850.

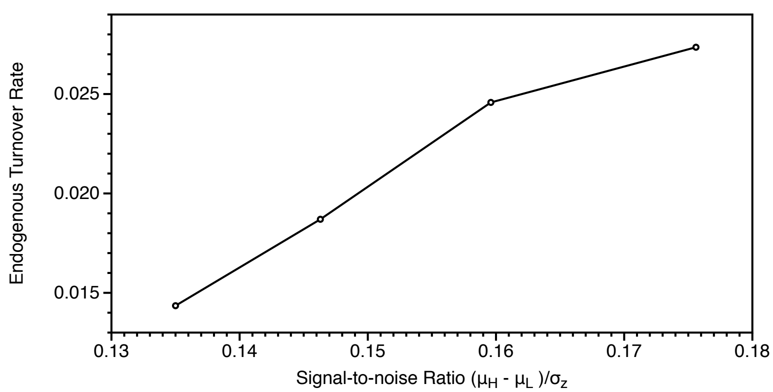
Figure 1. Comparative Analysis: Signal-to-Noise Ratio

Figure 1 depicts changes in a) the percentage of "good" matches, b) endogenous turnover rate, and c) the degree of skewness of CEO compensation in response to changes in the signal-to-noise ratio $\frac{\mu_H - \mu_L}{\sigma_z}$. All plots are based on a CEO-firm panel simulated from parameter estimates in Table 9.

Panel A: Percentage of "Good" Matches



Panel B: Endogenous Turnover Rate



Panel C: Degree of Skewness of CEO Compensation

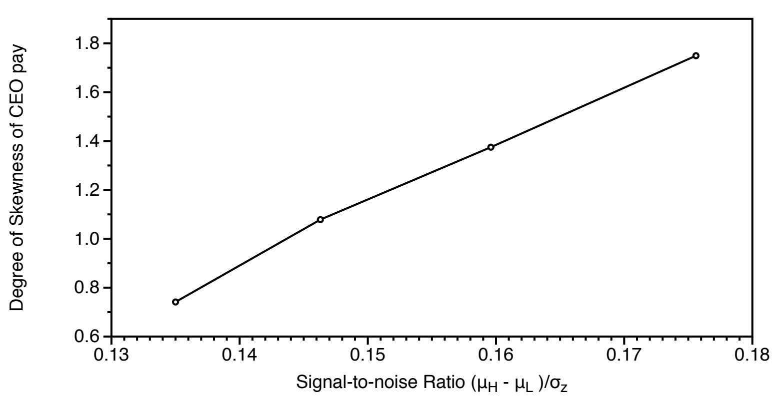
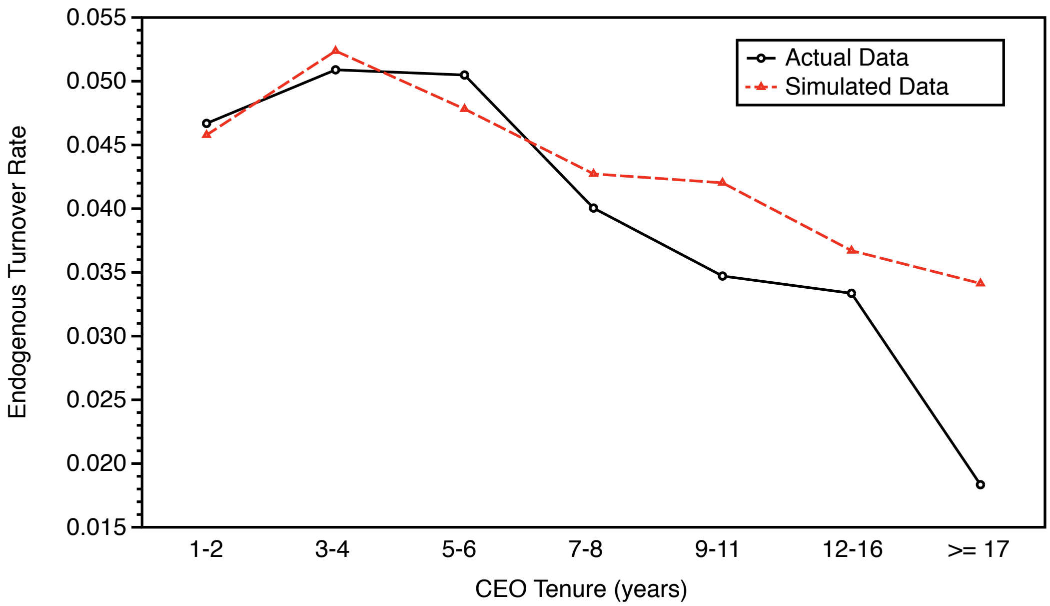


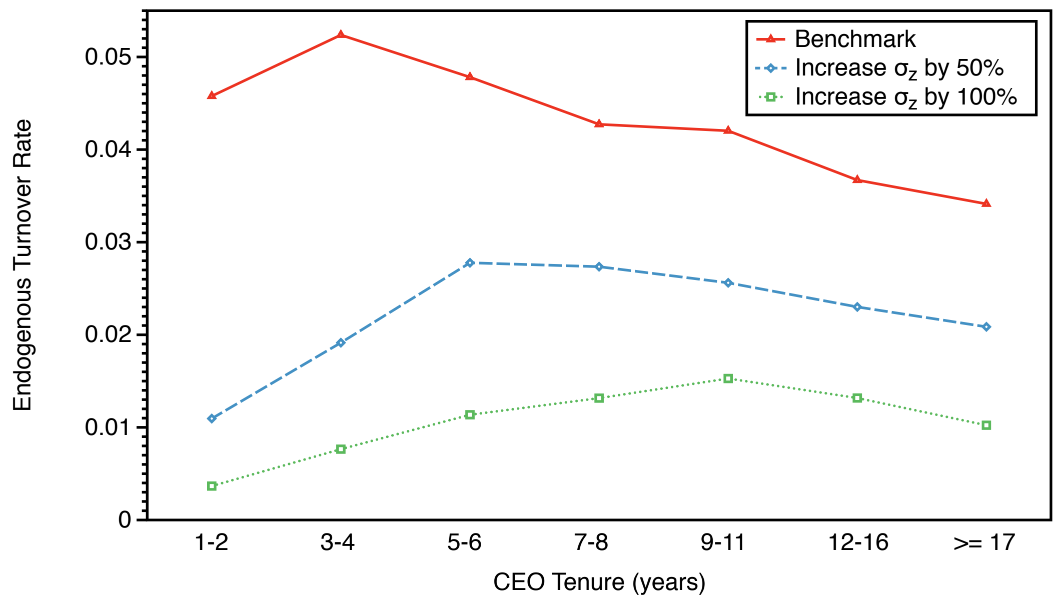
Figure 2. The Hazard Rate Curve of CEO Turnover Conditional on CEO Tenure

In Figure 2, I plot the hazard rate curves of CEO turnover conditional on seven CEO tenure intervals (in years): 1-2, 3-4, 5-6, 7-8, 9-11, 12-16, >=17 defined by Jenter and Lewellen (2014). Panel A shows the actual versus the simulated hazard rate curves of CEO turnover. The black solid line represents the endogenous turnover rates calculated from the actual data; the red dotted line depicts the endogenous turnover rates of a simulated CEO-firm panel from parameter estimates in Table 5. Panel B shows the hazard rate curves from the counterfactual analysis in Section 4.4. The red line represents endogenous turnover rates from the simulated data at the benchmark estimation in Table 5. The blue and green lines represent the conditional hazard rate curves with 50% and 100% increases in the idiosyncratic productivity shock σ_z. Panel C shows the hazard rate curves from the actual data in two subsamples. The black line depicts turnover rates of external CEOs and the red line shows turnover rates of CEOs who are internally promoted. The criteria to distinguish external CEOs from internal ones are defined in Section 3.1.

Panel A: Actual vs. Simulated



Panel B: Counterfactual Analysis: Increasing σ_z



Panel C: Actual Hazard Rate Curve: External v.s. Internal

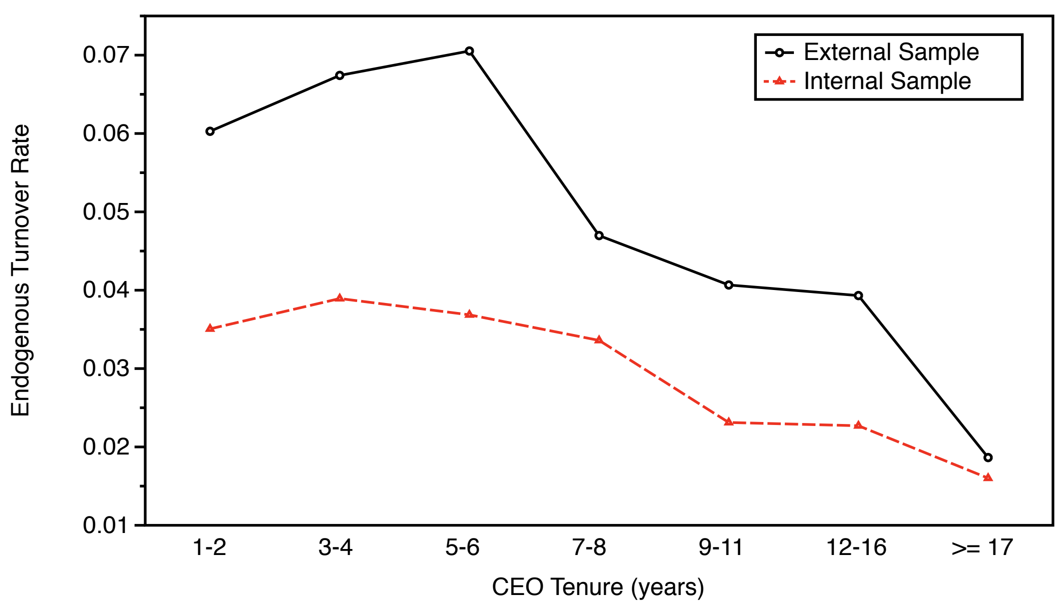
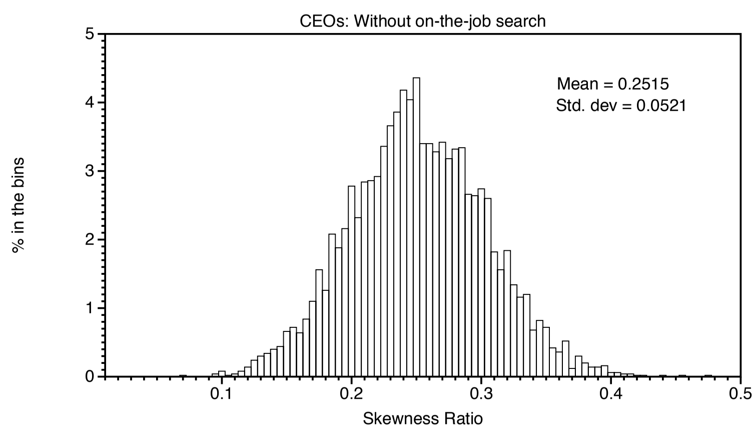


Figure 3. Counterfactual Analyses: Comparing to Gabaix and Landier (2008)

Figure 3 depicts the skewness ratios of the simulated CEO compensation to the CEO compensation derived from actual firm size in Gabaix and Landier (2008) for the 1,000 highest paid CEOs. In each case, I estimate the pay-size elasticity $\widehat{\gamma - \frac{\beta}{\alpha}}$ in the Gabaix and Landier (2008) regression $\ln(w_{it+1}) = d + \frac{\beta}{\alpha} \times \ln(S_{n_\star,t}) + (\gamma - \frac{\beta}{\alpha}) \times \ln(S_{i,t}) + \varepsilon_{it}$ with my sample, generate compensation for the 1,000 highest paid CEOs $Comp\_GL$ following Gabaix and Landier's (2008) Proposition 2. Their Proposition 2 defines the compensation the $n-th$ CEO receives as $w(n) = D(n_\star)S(n_\star)^{\frac{\beta}{\alpha}}S(n)^{\gamma-\frac{\beta}{\alpha}}$, where $S(n)$ is the actual size observed in my sample and $n_\star = 250th$ is the reference firm. I then calculate the skewness of their compensation measure $skewness(Comp\_GL)$ as the denominator for both histograms. In Panel A, the skewness ratio is defined as $skewness(Comp\_Simul)/skewness(Comp\_GL)$, where $Comp\_Simul$ is the simulated compensation for the 1,000 highest paid CEOs based on the simulated CEO-firm sample from parameter estimations of the benchmark model in Table 5. In Panel B, the skewness ratio is defined as $skewness(Comp\_Simul\_Otjs)/skewness(Comp\_GL)$, where $Comp\_Simul\_Otjs$ is the simulated compensation for the 1,000 highest paid CEOs based on the simulated CEO-firm sample from parameter estimations of the model with on-the-job search in Table 9. In each panel, the histogram represents 5,000 simulations.

Panel A



Panel B

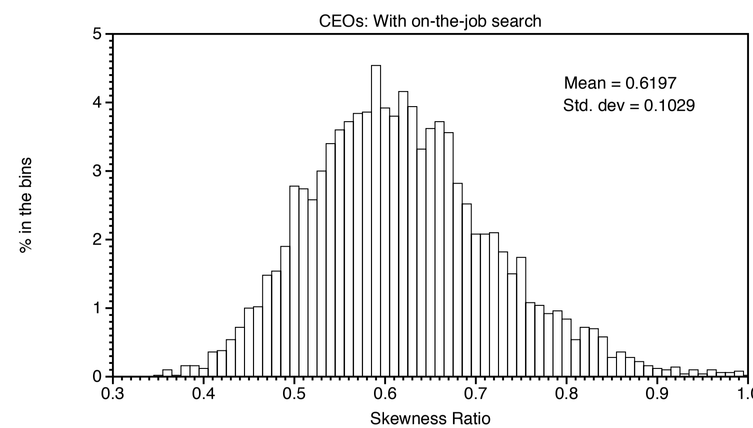


Figure 4. Subsample Analysis

In Figure 4, I plot the actual versus the simulated moments by applying the model with on-the-job search to four subsamples, including 1) external CEOs, 2) internal CEOs, 3) pre-2006 and 4) post-2006. Moments include: a) mean of firm profitability, b) standard deviation of firm profitability, c) mean of CEO compensation, d) skewness of CEO compensation, e) endogenous turnover rate, and f) on-the-job search rate. All the actual moments are calculated from a sample of 19,600 CEO-year observations, ranging from the fiscal year of 1994 to 2013.

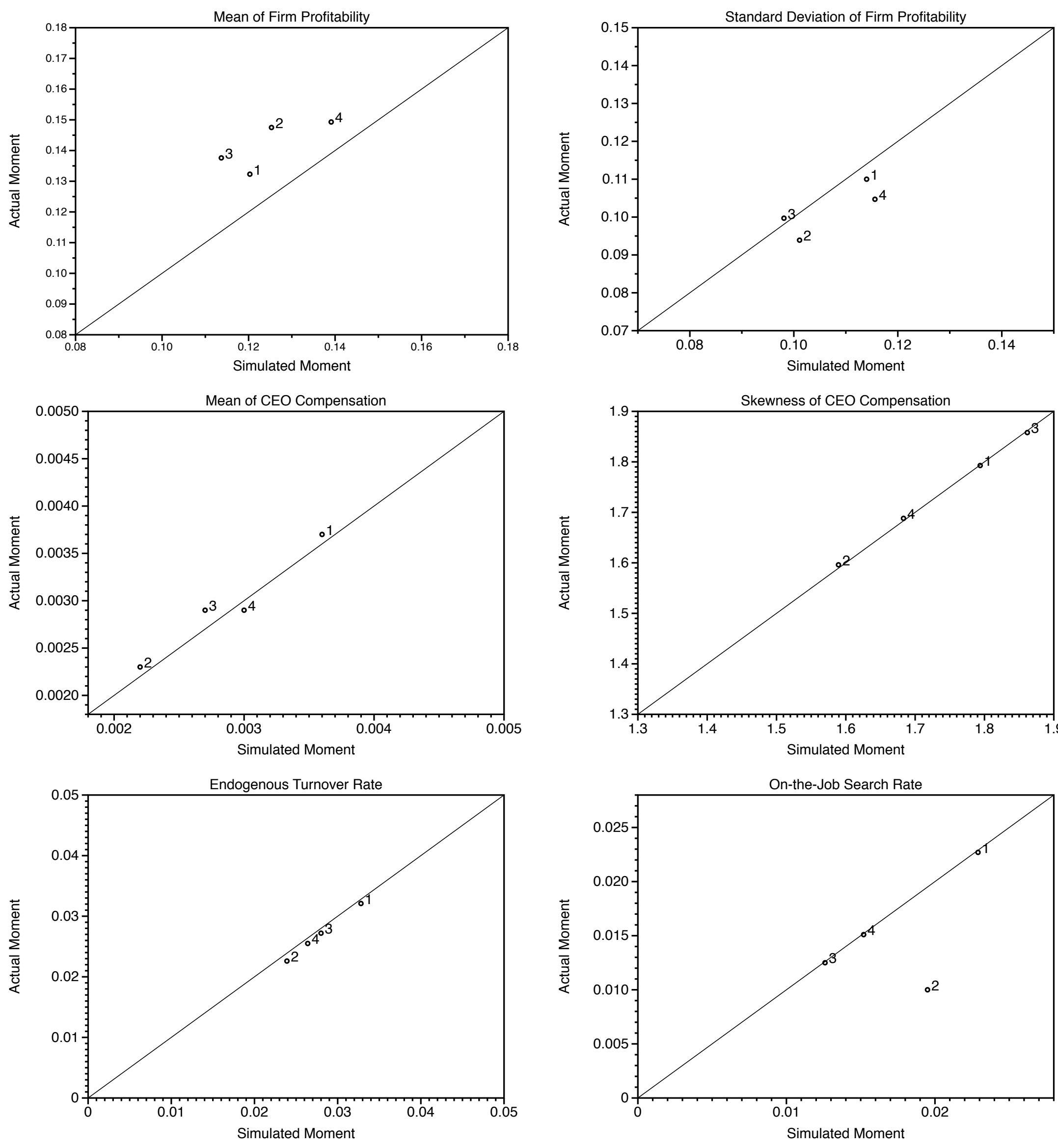


Table 1: Variable Definitions

In Table 1, I define all variables used in this paper.

Variable	**Definition**
Financial Characteristics	
Profitability	Operating Income Before Depreciation (OIBDP) / lagged Assets - Total (AT)
Leverage	(Debt in Current Liabilities - Total (DLC) + Long-Term Debt - Total (DLTT)) / Assets - Total (AT)
Tangibility	Property, Plant and Equipment - Total: Net (PPENT) / Assets - Total (AT)
Total Assets	Assets - Total (AT)
MarketCap	Price Close - fiscal (PRCC_F) × Common Shares Outstanding (CSHO)
Industry Dummy Variable	Dummy variable defined at the first two-digit of Standard Industrial Classification code (SIC)
Fiscal Year Dummy Variable	Dummy variable defined for each distinct fiscal year (Fyear)
Executive Compensation	
Salary + Bonus	(Salary (SALARY) + Bonus (BONUS)) / lagged Assets - Total (AT)
Salary + Bonus + Stock + Option	Before 2006: (Salary (SALARY) + Bonus (BONUS) + Stock (RSTKGRNT) + Option (Option_awards_BLK_value))/ lagged Assets - Total (AT) After 2006: (Salary (SALARY) + Bonus (BONUS) + Stock (stock_award_fv) + Option (option_award_fv))/ lagged Assets - Total (AT)
Salary + Bonus + Stock + Option + Other Incentive pay	Before 2006: (Salary (SALARY) + Bonus (BONUS) + Stock (RSTKGRNT) + Option (Option_awards_BLK_value) + Other Incentive pay (LTIP))/ lagged Assets - Total (AT) After 2006: (Salary (SALARY) + Bonus (BONUS) + Stock (stock_award_fv) + Option (option_award_fv) + Other Incentive pay (NONEQ_INCENT))/ lagged Assets - Total (AT)
Total Compensation	Total Compensation (TDC1)/ lagged Assets - Total (AT)

Variable	Definition
Executive Characteristics	
Gender	Dummy Variable = 1 if a CEO's gender (GENDER) = "FEMALE"
Tenure	Date Left as CEO (LEFTOFC) - Date Became CEO (BECAMECEO), rounded to its closest integer
Education	Indicating the highest degree the CEO receives: = 0 if unobservable, =1 if no school, = 2 if with high school diploma, = 3 if with Bachelors Degrees, = 4 if with Master Degrees or other professional degrees (JD and MD included), = 5 if graduates with a Doctoral degree
MBA	Dummy variable = 1 if the CEO receives a MBA degree
JD	Dummy variable = 1 if the CEO receives a JD degree
Turnover	
EXO_Turnover	Dummy variable = 1 if the CEO is exogenously turned over in the given fiscal year (details refers to Appendix C)
END_Turnover	Dummy variable = 1 if the CEO is endogenously turned over in the given fiscal year (details refers to Appendix C)
OTJS	Dummy variable = 1 if the CEO accepts an equivalent offer from other firms in the given fiscal year (details refers to Appendix C)
Insider/Outsider Status	
External	Dummy variable = 1 if the CEO is an outsider of the firm at the time being selected as the CEO or CEO successor
Internal	Dummy variable = 1 if the CEO is an insider of the firm at the time being selected as the CEO or CEO successor
Founder	Dummy variable = 1 if the CEO is the founder or co-founder of the firm
Governance Characteristics	
Interlock	Dummy variable = 1 if the CEO is listed in the Compensation Committee Interlocks section of the proxy
G-index	Governance index constructed by Gompers *et al*. (2003), ranging from 0 to 24 to indicate the quality of corporate governance
E-index	Governance index constructed by Bebchuk and Fried (2004), ranging from 0 to 6 to indicate the quality of corporate governance
Macroeconomic Variables	
Risk-free rate	Average 3-month T-bill rate over sample period

Table 2: Parameter Definitions

In Table 2, I define all parameters used in this paper.

Parameter	Parameter Definition
Estimated outside the model	
r	Risk-free rate
b	CEO outside wealth
δ	Exogenous turnover rate of CEOs
$\eta \in (0,1)$	Matching elasticity with respect to unmatched individuals
Estimated inside the model	
$p_0 \in (0,1)$	Proportion of "high" quality matches
	also serves as the prior belief of the match quality
κ	Average search cost of firms
μ_H	Average productivity of "high" quality matches
μ_L	Average productivity of "low" quality matches
μ_z	Average retaining profitability of firms
σ_z	Standard deviation of the idiosyncratic productivity shock
$b^{NB} \in (0,1)$	Average bargaining power of CEOs
$\psi \in (0,1)$	Relative job finding rate/ contact rate of on-the-job search

Table 3: Summary Statistics: Main Variables

Table 3 presents summary statistics of the main variables. Panel A focuses on the mean, standard deviation, 25% quantile, median, and 75% quantile of each variable. Panel B provides statistics on sample composition. The main sample ranges from the fiscal year of 1994 to 2013, with 19,600 CEO-year observations.

Panel A: Characteristics	Mean	S.D.	25%	50%	75%
Financial Characteristics					
Profitability	0.14	0.11	0.08	0.14	0.20
Leverage	0.23	0.57	0.07	0.21	0.34
Tangibility	0.28	0.24	0.08	0.21	0.42
Total Assets ($ billion)	7.99	17.39	0.59	1.76	6.25
MarketCap ($ billion)	5.78	11.52	0.58	1.59	4.84
Executive Compensation ($\times$100)					
Salary + Bonus	0.10	0.12	0.02	0.06	0.13
Salary + Bonus + Stock + Option	0.25	0.30	0.05	0.13	0.31
Salary + Bonus + Stock + Option + Other	0.28	0.32	0.06	0.16	0.36
Incentive pay					
Total Compensation	0.29	0.34	0.06	0.17	0.38
Executive Characteristics					
Tenure Length	9.43	6.59	5.00	7.00	12.00
Age	56.46	6.83	52.00	57.00	61.00

Panel B: Sample Compositions (in %)				
Education				
	No School/ High School	Bachelors Degree	Masters Degree	Doctorate Degree
	3.23	30.00	42.04	5.79
	MBA	JD		
	26.63	5.05		
Turnover				
	Exogenous Turnover	Endogenous Turnover	On-the-job search	
	8.48	4.02	1.39	
Insider/Outsider				
	External	Internal	Founder	
	27.26	58.70	10.62	
Gender				
	Female	Male		
	2.04	97.96		
Governance				
	Interlock = 1	Interlock = 0		
	3.88	96.12		
	$G-index \leq 7$	$7 < G-index \leq 9$	$9 < G-index \leq 11$	$G-index > 11$
	14.43	15.72	16.30	12.66
	$E-index \leq 2$	$2 < E-index \leq 4$	$E-index > 4$	
	47.02	42.60	6.50	

Table 4: First-stage regression of Main Variables

Panel A of Table 4 reports the regression results of firm profitability (column (1)) and CEO total compensation (column (2)). $T-$statistics clustered at two-digit SIC code level are reported in the parentheses. Panel B reports summary statistics of the main variables $Profit\_Residual$ and $Comp\_Residual$ used to estimate the model. $Profit\_Residual$ equals the regression residual of column (1) plus the population mean of firm profitability; $Comp\_Residual$ sums regression residuals of column (2) with the population mean of CEO total compensation. All regressions are based on a main sample containing 19,600 observations ranging from the fiscal year of 1994 to 2013.

Panel A: Regression

Dependent Variable	Firm Profitability ($OIBDP_t/AT_{t-1}$)	CEO Total Compensation ($TDC1_t/AT_{t-1}$)
	(1)	(2)
$Leverage_{it-1}$	$-0.0153^{\star}$	-0.0004
	(-1.68)	(-0.93)
$Tangibility_{it-1}$	$0.0793^{\star\star\star}$	$-0.0024^{\star\star\star}$
	(25.52)	(-20.67)
$MarketCap_{it-1}$	$4.1103E-07^{\star\star\star}$	$-2.77E-08^{\star\star\star}$
	(13.16)	(-18.57)
$Gender_{it}$	0.0060	$0.0007^{\star\star\star}$
	(1.46)	(4.29)
$Interlock_{it}$	0.0036	$-0.0002^{\star\star}$
	(1.05)	(-2.21)
$External_{it}$	$0.0192^{\star\star\star}$	$0.0005^{\star\star\star}$
	(5.94)	(4.52)
$Internal_{it}$	$0.0288^{\star\star\star}$	$-0.0006^{\star\star\star}$
	(9.35)	(-5.28)
$Founder_{it}$	$0.0518^{\star\star\star}$	$0.0012^{\star\star\star}$
	(13.72)	(9.37)
Fiscal Year FE	x	x
Industry Clustering	x	x
Education Dummy	x	x
Adjusted R^2	0.6578	0.4884

$\star$, $\star\star$ and $\star\star\star$ indicate statistical significance at the 10%, 5%, and 1% levels.

Panel B: Summary Statistics

	Min	25%	50%	75%	Max	Mean	S.D	Skewness
$Profit\_Residual$	-0.19	0.08	0.14	0.20	0.99	0.14	0.10	0.50
$Comp\_Residual$ ($\times$ 100)	-0.27	0.09	0.21	0.37	4.34	0.29	0.32	1.75

Table 5: Simulated Moments Estimation: Full Sample

Table 5 provides the estimation results of the benchmark model from simulated method of moments (SMM). Panel A reports the actual versus the simulated moments, as well as the $t-$statistics for the difference. Panel B reports parameter estimations with standard errors in parantheses. The estimation is based on a sample of 19,600 CEO-year observations ranging from the fiscal year of 1994 to 2013.

Panel A: Moments

	Actual Moments	Simulated Moments	T-test for Difference
Mean of Compensation	0.0029	0.0030	-0.1393
Std of Compensation	0.0032	0.0015	0.0725
Skewness of Compensation	1.7546	1.7538	0.0004
Mean of Profitability	0.1438	0.1180	2.6489
Std of Profitability	0.1024	0.1237	-3.7063
AR(1) Coefficient ρ of Profitability	0.7263	0.8630	-1.8954
AR (1) Std of Residual of Profitability	0.1302	0.1179	0.9208
Endogeneous Turnover Rate	0.0402	0.0369	0.2299
$\hat{\beta}_{\Delta Profitability}(Endogenous\ Turnover\ Rate)$	-0.5737	-1.1994	0.8072
$\hat{\beta}_{log(Tenure)}(Endogenous\ Turnover\ Rate)$	-0.0158	-0.0078	-0.7508
$\hat{\beta}_{Inter}(Endogenous\ Turnover\ Rate)$	0.1865	0.2965	-0.2908
		J-statistics: 3.9876	P-val: 0.5923

Panel B: Parameters

μ_H	μ_L	p_0	μ_z	σ_z	b^{NB}	κ
0.7846	-0.2548	0.2844	0.0253	4.6164	0.0123	0.2320
(0.0187)	(0.0165)	(0.0132)	(0.0182)	(0.0198)	(0.0029)	(0.0189)

Table 6: CEO Tenure Continuation

Table 6 reports coefficient estimates from a linear probability model (LPM): $1 - End\ Turnover_{it} = \alpha + \beta_1 Comp_{it} + \beta_2 Comp_{it}^2 + Controls + \varepsilon_{it}$. The results in columns (1) - (4) are coefficient estimates using the actual data, consisting of 19,600 CEO-year observations ranging from the fiscal year of 1994 to 2013. In the regression of each column, *Comp* is the total compensation $TDC1_{it}/AT_{it-1}$ of the i-th CEO in period t from the actual data. Control variables include different firm and CEO characteristics. $T-$statistics clustered at fiscal year and industry level (defined as the first two-digit of the SIC code) are in parentheses. The results in column (5) are coefficient estimates of the LPM model using the simulated data generated from parameter estimation in Table 5, with bootstrapped $t-$statistics in the parentheses.

	Actual Data				Simulated Data
	(1)	(2)	(3)	(4)	(5)
Comp	3.0519**	3.1355**	3.4457***	3.4180***	6.8962***
	(2.35)	(2.41)	(2.63)	(2.61)	(23.39)
$Comp^2$	−199.8926**	−201.5638**	−215.4190**	−211.4544**	−562.5264***
	(-2.08)	(-2.09)	(-2.23)	(-2.19)	(-15.32)
lag(Leverage)		−0.0001	−0.0001	−0.0001	
		(-0.20)	(-0.20)	(-0.15)	
lag(Tangibility)		0.0123	0.0121	0.0111	
		(1.35)	(1.33)	(1.21)	
Internal			0.0187**	0.0183**	
			(2.19)	(2.17)	
External			0.0092	0.0088	
			(1.05)	(1.01)	
Founder			0.0202**	0.0193**	
			(2.17)	(2.08)	
Gender				−0.0300**	
				(-2.56)	
Interlock				0.0068	
				(1.30)	
Education	Y	Y	Y	Y	N
FiscalYear FE	Y	Y	Y	Y	N
Industry FE	Y	Y	Y	Y	N
R^2	0.9740	0.9740	0.9748	0.9756	\
#obs	19,600	19,600	19,600	19,600	100,000

⋆, ⋆⋆ and ⋆⋆⋆ indicate statistical significance at the 10%, 5%, and 1% levels.

Table 7: Counterfactual Analysis: Turnover - Performance Sensitivity

Table 7 presents the results of counterfactual analyses on endogenous turnover rate and turnover-performance (TP) sensitivity with firm productivity shock σ_z at its benchmark value from Table 5 (columns (1) - (3)), with a 50% increase (columns (4) - (6)) and a 100% increase in σ_z (columns (7) - (9)). TP sensitivities $\hat{\beta}$ are estimated by the following linear probability model (LPM): $END\_Turnover_{it} = \alpha + \beta Profitability_{it} + \gamma Controls_{it} + \varepsilon_{it}$. In columns (2), (5), and (8), $Controls$ equals $log(Tenure)$; in columns (3), (6), and (9), $Controls$ includes $log(Tenure)$ and an interaction term of $Profitability \times log(Tenure)$. In each case, bootstrapped $t-$statistics are in parentheses.

σ_z		*Benchmark*			+50%			+100%	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.0711***	0.0531***	0.1067***	0.0463***	0.0271***	0.0624***	0.0350***	0.0185***	0.0365***
	(64.74)	(30.21)	(27.80)	(47.75)	(25.35)	(18.35)	(37.59)	(21.13)	(9.82)
Profitability	−0.2936***	−0.3084***	−1.0080***	−0.2903***	−0.3012***	−0.8342***	−0.2731***	−0.2834***	−0.5517***
	(-45.77)	(-45.79)	(-32.36)	(-60.65)	(-59.60)	(-22.04)	(-36.78)	(-36.88)	(-12.22)
$log(Tenure)$		0.0075***	−0.0110***		0.0081***	−0.0042***		0.0065***	0.0004
		(11.39)	(-8.37)		(19.47)	(-3.60)		(19.66)	(0.31)
Inter			0.2364***			0.1794***			0.0896***
			(23.79)			(14.62)			(6.13)
		$END\_Turnover\ Rate = 0.0369$			$END\_Turnover\ Rate = 0.0209$			$END\_Turnover\ Rate = 0.0128$	

$\star$, $\star\star$ and $\star\star\star$ indicate statistical significance at the 10%, 5%, and 1% levels.

Table 8: Counterfactual Analysis: Randomizing Seperation Decisions

Table 8 reports the results of counterfactual analyses on 1) total welfare of the executive labor market in columns (1) and (3) $TotalWelfare = \sum_{\Omega(p.x)} S(p,x) \times h(p,x)$, where $S(p,x) = J(p,x) - V + W(p,x) - U - \mu_z$ is the net matching surplus and $h(p,x)$ is the stationary distribution over the state space (p,x); 2) the degree of skewness of the simulated CEO compensation in columns (2) and (4). Columns (1) and (2) are based on parameter estimates from a model without on-the-job search in Table 5; the results in columns (3) and (4) are from parameter estimates from a model with on-the-job search in Table 9. The counterfactual analyses randomize endogenous separation (Row 2), on-the-job search (Row 3) and both (Row 4), and compare the results with belief-contingent turnover and on-the-job search decisions derived from both a model with and without on-the-job search (Row 1).

	Without On-the-Job Search		*With On-the-Job Search*	
	Total Welfare	*Compensation Skewness*	*Total Welfare*	*Compensation Skewness*
	(1)	(2)	(3)	(4)
1. Benchmark Model	1.2337	1.7538	1.3247	1.7492
2. Randomizing Endogenous Separation	1.0360	0.9557	0.9445	0.7709
3. Randomizing On-the-job Search	\	\	1.1981	1.6143
4. Randomizing Both	\	\	0.8837	0.6819

Table 9: Simulated Moments Estimation: On-the-Job Search

Table 9 summarizes estimation results of the benchmark model from simulated method of moments (SMM). Panel A reports the actual versus the simulated moments, as well as the t-statistics for the difference. Panel B reports parameter estimation with standard errors in parantheses. The estimation is based on a sample of 19,600 CEO-year observations ranging from the fiscal year of 1994 to 2013.

Panel A: Moments

	Actual Moments	Simulated Moments	T-test for Difference
Mean of Compensation	0.0029	0.0030	-0.1611
Std of Compensation	0.0032	0.0019	0.0578
Skewness of Compensation	1.7546	1.7490	0.0022
Mean of Profitability	0.1438	0.1325	0.9973
Std of Profitability	0.1024	0.1114	-1.0415
AR(1) Coefficient ρ of Profitability	0.7263	0.8719	-2.1679
AR (1) Std of Residual of Profitability	0.1302	0.1000	2.0600
Endogeneous Turnover Rate	0.0263	0.0274	-0.1471
$\hat{\beta}_{\Delta Profitability}(Endogenous\ Turnover\ Rate)$	-0.4234	-0.8746	0.9513
$\hat{\beta}_{log(Tenure)}(Endogenous\ Turnover\ Rate)$	-0.0103	-0.0022	-0.3240
$\hat{\beta}_{Inter}(Endogenous\ Turnover\ Rate)$	0.1291	0.2092	-0.3100
On-the-job Search Rate	0.0139	0.0141	-0.0271
$\hat{\beta}_{\Delta Profitability}(On-the-job\ Search\ Rate)$	-0.1504	-0.1324	-0.0376
$\hat{\beta}_{log(Tenure)}(On-the-job\ Search\ Rate)$	-0.0055	-0.0030	-0.1397
$\hat{\beta}_{Inter}(On-the-job\ Search\ Rate)$	0.0574	0.0274	0.1608
		J-statistics: 4.0099	P-val: 0.2214

Panel B: Parameters

μ_H	μ_L	p_0	μ_z	σ_z	b^{NB}	κ	ψ
0.9184	-0.3335	0.2905	0.0710	7.1308	0.0163	0.8692	0.1395
(0.0409)	(0.0053)	(0.0400)	(0.0380)	(1.2294)	(0.0034)	(0.0189)	(0.0229)

Table 10: Parameter Estimations: Subsamples

Table 10 reports the results of paramter estimations, along with $J-$statistics and $P-$values of over-identification tests, by applying the full model to four subsamples, including: 1) External, 2) Internal, 3) Pre-2006, 4) Post-2006 period.

	Parameters: Subsamples										
	μ_H	μ_L	p_0	μ_z	σ_z	b^{NB}	κ	ψ	# obs	J-stat	p-val
1. External Subsample	0.9440	-0.2224	0.2240	0.0301	5.1194	0.0231	0.9937	0.3915			
	(0.0045)	(0.0558)	(0.0047)	(0.0591)	(0.6019)	(0.0046)	(0.2781)	(0.0903)	5,288	4.7795	0.3131
2. Internal Subsample	0.6962	-0.2201	0.2915	0.0476	4.0478	0.0127	0.6837	0.1861			
	(0.0109)	(0.0077)	(0.0071)	(0.0116)	(0.0061)	(0.0024)	(0.1068)	(0.0611)	11,434	4.8222	0.3183
3. Pre-2006	0.8249	-0.2667	0.2834	0.0377	6.3420	0.0179	0.8732	0.1549			
	(0.0086)	(0.0186)	(0.0041)	(0.0239)	(0.9498)	(0.0030)	(0.0121)	(0.0625)	9,244	2.8856	0.1046
4. Post-2006	0.9865	-0.4038	0.3204	0.0707	8.7674	0.0140	0.7070	0.1051			
	(0.0978)	(0.0230)	(0.0863)	(0.0239)	(1.1230)	(0.0018)	(0.0716)	(0.0423)	10,356	5.6616	0.4202

Table 11: Parameter Estimations: Robustness

Table 11 reports the results of paramter estimations, along with $J-$statistics and $P-$values of over-identification tests, by applying the full model to nine robustness samples. The first three rows provide estimation results of the sample with three different measures of CEO compensation 1) Salary and Bonus, 2) Salary, Bonus, Stock, and Option, and 3) Salary, Bonus, Stock, Option, and Other Incentives. Then, I estimate the model with respect to different CEO education subsamples to test the validation of the model, including: 4) High School Diploma/ No School, 5) Bachelors Degree, 6) Masters Degree, 7) Doctorate Degree, 8) MBA, 9) JD. Standard errors are in parentheses.

	Parameters: Robustness										
	μ_H	μ_L	p_0	μ_z	σ_z	b^{NB}	κ	ψ	# obs	J-stat	p-val
1. Salary and Bonus	0.7335	-0.1922	0.2835	0.0122	4.0071	0.0072	0.8786	0.1799			
	(0.0399)	(0.0404)	(0.1106)	(0.0234)	(0.7974)	(0.0062)	(0.3156)	(1.3671)	19,600	3.4551	0.1600
2. Salary, Bonus, Stock and Option	0.7254	-0.2211	0.2924	0.0170	4.9236	0.0138	0.3132	0.2924			
	(0.0467)	(0.0578)	(0.0274)	(0.0509)	(0.9409)	(0.0034)	(0.0026)	(0.0274)	19,600	3.7258	0.1892
3. Salary, Bonus, Stock, Option,	0.7503	-0.2121	0.2550	0.0509	3.8832	0.0123	0.3124	0.0682			
and Other Incentives	(0.0451)	(0.0449)	(0.0578)	(0.0012)	(0.7449)	(0.0030)	(0.0372)	(0.0374)	19,600	4.2024	0.2438
4. High School Diploma/ No	0.7784	-0.2778	0.3453	0.0541	6.5730	0.0108	0.7771	0.0004			
School	(0.0367)	(0.0003)	(0.0718)	(0.0013)	(0.5650)	(0.0022)	(0.0242)	(1.5067)	664	5.9488	0.4543
5. Bachelors Degree	0.8724	-0.3194	0.3182	0.0309	6.5179	0.0204	0.8287	0.2040			
	(0.0235)	(0.0007)	(0.0417)	(0.0235)	(0.6350)	(0.0036)	(0.1339)	(0.0508)	5,869	7.0279	0.5740
6. Masters Degree	0.7239	-0.2610	0.2993	0.0274	6.2912	0.0187	0.8981	0.4142			
	(0.0342)	(0.0189)	(0.0596)	(0.0005)	(1.4706)	(0.0034)	(0.1763)	(0.0401)	8,183	3.3928	0.1536
7. Doctorate Degree	0.7387	-0.2781	0.2943	0.0519	4.3406	0.0272	0.8862	0.2702			
	(0.0279)	(0.0017)	(0.0113)	(0.0089)	(0.1546)	(0.0058)	(0.0145)	(0.1145)	1,144	5.4446	0.3941
8. MBA Subsample	0.7881	-0.2876	0.2947	0.0277	5.7717	0.0195	0.4833	0.2527			
	(0.0352)	(0.0139)	(0.0324)	(0.0161)	(0.6838)	(0.0113)	(0.0362)	(3.0794)	5,219	4.1301	0.2353
9. JD Subsample	0.7480	-0.2879	0.3146	0.0396	6.1014	0.0103	0.2898	0.1004			
	(0.0006)	(0.0148)	(0.0360)	(0.0146)	(1.0916)	(0.0012)	(0.0456)	(0.0059)	989	4.7678	0.3117

Appendix A. Mathematical Appendix

A.1. A Full Model with On-the-Job Search

To extend the benchmark model to a full model with CEOs performing on-the-job searches, I assume that an incumbent CEO contacts open vacancies at a rate of $\psi\lambda$, where $\psi \in (0,1)$ is the relative contact rate and λ is the job-finding rate defined by the equilibrium matching technology. Then, I assume that a "poaching auction" in Moscarini (2005) takes place. Precisely, the contacted firm first makes an offer, followed by sequential bids between the contacted firm and the current matched firm within that period. The auction settles when one firm declines to submit a new bid over the CEO. In the end, the CEO receives an offer from the highest bid in the form of a lump-sum transfer. If the CEO accepts the offer from the contacted firm, the belief of the match quality of the CEO with the newly matched firm resets to its prior value p_0. Then, the value functions of an unmatched individual U and of an idle firm V still equal to equation (4) and (6). The value functions of a matched CEO and a matched firm become:

$$W(p,x) = w(p,x) + \frac{1-\delta}{1+r} \max_{d_1^{CEO}\in\{0,1\}} \Big\{ U, \psi\lambda \max_{d_2^{CEO}\in\{0,1\}} \Big\{ \mathbb{E}_{x'}[W(p_0,x')], \mathbb{E}_{x'}[W(p',x')] \Big\} + (1-\psi\lambda)\mathbb{E}_{x'}[W(p',x')] \Big\} + \frac{\delta}{1+r} U \tag{A.1}$$

and

$$J(p,x) = r(p,x) + \frac{1-\delta}{1+r} \max_{d^F\in\{0,1\}} \Big\{ V, \psi\lambda \Big[V \times \mathbb{I}_{\{d_2^{CEO}=1\}} + \mathbb{E}_{x'}[J(p',x')] \times \mathbb{I}_{\{d_2^{CEO}=0\}} \Big] (1-\psi\lambda)\mathbb{E}_{x'}[J(p',x')] \Big\} + \frac{\delta}{1+r} V. \tag{A.2}$$

where $\mathbb{I}_{\{d_2^{CEO}=1\}}$ equals to 1 if the matched CEO accepts the outside offer. The Nash bargaining solves CEO compensation as:

$$w(p,x) = (1-b^{NB})b + b^{NB}[p\mu_H + (1-p)\mu_L] + b^{NB}\kappa\theta \times (1-\psi\mathbb{I}_{\{p'<p_0\}}), \tag{A.3}$$

firm profitability excluding the retaining part μ_z as:

$$r_{exclude}(p,x) = (1-b^{NB})[p\mu_H + (1-p)\mu_L - b] - b^{NB}\kappa\theta \times (1-\psi\mathbb{I}_{\{p'<p_0\}}), \tag{A.4}$$

and firm profitability equals:

$$r(p,x) = (1-b^{NB})[p\mu_H + (1-p)\mu_L - b] + \mu_z - b^{NB}\kappa\theta \times (1-\psi\mathbb{I}_{\{p'<p_0\}}). \tag{A.5}$$

In Sections A.2 and A.3, I prove propositions with respect to the full model with on-the-job search. The benchmark proof is equivalent by setting $\psi = 0$.

A.2. *Proof of Proposition 1*

Proof. To prove Proposition1, I first re-write the value function of the CEO in equation (A.1) into threshold decisions:

$$W(p,x)=w(p,x)+\frac{1-\delta}{1+r}\Big\{U\times\mathbb{I}_{\{d_1^{CEO}=1\}}+\Big[\psi\lambda\Big\{\mathbb{E}_{x'}[W(p_0,x')]\times\mathbb{I}_{\{d_2^{CEO}=1\}}$$
$$+\mathbb{E}_{x'}[W(p',x')]\times\mathbb{I}_{\{d_2^{CEO}=0\}}\Big\}+(1-\psi\lambda)\mathbb{E}_{x'}[W(p',x')]\Big]\times\mathbb{I}_{\{d_1^{CEO}=0\}}\Big\}+\frac{\delta}{1+r}U. \quad \text{(A.6)}$$

Coupled with the "free-entry" condition $V=0$, the value function of the matched firm in equation (A.2) is:

$$J(p,x)=r(p,x)+\frac{1-\delta}{1+r}\Big\{\psi\lambda\mathbb{E}_{x'}[J(p',x')]\times\mathbb{I}_{\{d_2^{CEO}=0\}}+(1-\psi\lambda)\mathbb{E}_{x'}[J(p',x')]\Big\}\times\mathbb{I}_{\{d^F=0\}}. \quad \text{(A.7)}$$

Then, I plug equation (A.6) and equation (A.7) into equation (8). Notice that the Nash bargaining takes place before firm productivity x is realized. Thus, I substitute x with $\mathbb{E}(x|\mathscr{F}_{t-1})$ in the following equation:

$$b^{NB}\Big\{p\mu_H+(1-p)\mu_L-w(p,x)+$$
$$\frac{1-\delta}{1+r}\Big\{\psi\lambda\mathbb{E}_{x'}[J(p',x')]\times\mathbb{I}_{\{d_2^{CEO}=0\}}+(1-\psi\lambda)\mathbb{E}_{x'}[J(p',x')]\Big\}\times\mathbb{I}_{\{d^F=0\}}\Big\}$$
$$=(1-b^{NB})\Big\{w(p,x)+\frac{1-\delta}{1+r}\Big\{U\times\mathbb{I}_{\{d_1^{CEO}=1\}}+\Big[\psi\lambda\Big\{\mathbb{E}_{x'}[W(p_0,x')]\times\mathbb{I}_{\{d_2^{CEO}=1\}}$$
$$+\mathbb{E}_{x'}[W(p',x')]\times\mathbb{I}_{\{d_2^{CEO}=0\}}\Big\}+(1-\psi\lambda)\mathbb{E}_{x'}[W(p',x')]\Big]\times\mathbb{I}_{\{d_1^{CEO}=0\}}\Big\}+\frac{\delta}{1+r}U\Big\}. \quad \text{(A.8)}$$

Next, I show $\mathbb{I}_{\{d^F=0\}}=\mathbb{I}_{\{d_1^{CEO}=0\}}$ from equation (8). Denote $\mathbb{E}^{CEO}(p',x')=\psi\lambda\Big\{\mathbb{E}_{x'}[W(p_0,x')]\times\mathbb{I}_{\{d_2^{CEO}=1\}}+\mathbb{E}_{x'}[W(p',x')]\times\mathbb{I}_{\{d_2^{CEO}=0\}}\Big\}+(1-\psi\lambda)\mathbb{E}_{x'}[W(p',x')]$ and $\mathbb{E}^F(p',x')=\psi\lambda\mathbb{E}_{x'}[J(p',x')]\times\mathbb{I}_{\{d_2^{CEO}=0\}}+(1-\psi\lambda)\mathbb{E}_{x'}[J(p',x')]$, then $\mathbb{I}_{\{d_1^{CEO}=0\}}=\mathbb{I}_{\{U<\mathbb{E}^{CEO}(p',x')\}}$ and $\mathbb{I}_{\{d^F=0\}}=\mathbb{I}_{\{\mathbb{E}^F(p',x')>0\}}$. Substitute the linear rent-sharing rule:

$$\mathbb{E}_{x'}[W(p',x')]-U=\frac{b^{NB}}{1-b^{NB}}\mathbb{E}_{x'}[J(p',x')] \quad \text{(A.9)}$$

and

$$\mathbb{E}_{x'}[W(p_0,x')]-U=\frac{b^{NB}}{1-b^{NB}}\mathbb{E}_{x'}[J(p_0,x')] \quad \text{(A.10)}$$

into equation (A.8), and it becomes:

$$\begin{aligned} w(p,x) &= (1-b^{NB})b + b^{NB}\Big\{(p\mu_H + (1-p)\mu_L) \\ &+ \frac{\lambda}{1+r}\mathbb{E}_{x'}[J(p_0,x')] \times (1-\psi\mathbb{I}_{\{d_2^{CEO}=1\}})\Big\} \\ &= (1-b^{NB})b + b^{NB}(p\mu_H + (1-p)\mu_L) \\ &+ b^{NB}\kappa\theta \times (1-\psi\mathbb{I}_{\{\mathbb{E}_{x'}[W(p',x')]<\mathbb{E}_{x'}[W(p_0,x')]\}}) \end{aligned} \tag{A.11}$$

and firm profitability yields:

$$\begin{aligned} r(p,x) &= (1-b^{NB})[p\mu_H + (1-p)\mu_L - b] + \mu_z \\ &- b^{NB}\kappa\theta \times (1-\psi\mathbb{I}_{\{\mathbb{E}_{x'}[W(p',x')]<\mathbb{E}_{x'}[W(p_0,x')]\}}). \end{aligned} \tag{A.12}$$

Given the functional forms of CEO compensation and firm profitability, I further show that both value functions in equations (A.6) and (A.7) are strictly increasing in the posterior belief p. First, for the matched CEO:

$$\begin{aligned} \frac{\partial W(p,x)}{\partial p} &= \frac{\partial w(p,x)}{\partial p} + \frac{1-\delta}{1+r}\mathbb{E}_{x'}[\frac{\partial W(p',x')}{\partial p'}\frac{\partial p'}{\partial p}] \\ &= b^{NB}(\mu_H - \mu_L) + \frac{1-\delta}{1+r}\mathbb{E}_{x'}[b^{NB}(\mu_H-\mu_L)\frac{\partial p'}{\partial p}]. \end{aligned} \tag{A.13}$$

Equation (2) suggests:

$$\begin{aligned} \frac{\partial p'}{\partial p} &= \frac{\phi(x|\mu_H)\cdot[p\phi(x|\mu_H)+(1-p)\phi(x|\mu_L)] - [\phi(x|\mu_H)-\phi(x|\mu_L)]\cdot p\phi(x|\mu_H)}{[p\phi(x|\mu_H)+(1-p)\phi(x|\mu_L)]^2} \\ &= \frac{\phi(x|\mu_H)\phi(x|\mu_L)}{[p\phi(x|\mu_H)+(1-p)\phi(x|\mu_L)]^2} > 0. \end{aligned} \tag{A.14}$$

Substituting equation (A.14) into equation (A.13), and combining with the the assumption that $\mu_H > \mu_L$, it yields:

$$\frac{\partial W(p,x)}{\partial p} > 0. \tag{A.15}$$

Thus, the value function of the matched CEO is strictly increasing in the posterior belief p. Similarly, by the linear rent-sharing rule in Equation (8), it must follow that:

$$\frac{\partial J(p,x)}{\partial p} > 0 \tag{A.16}$$

Last, I use equation (A.15) to simplify the decision rule of on-the-job searches. Since the value function of the matched CEO is strictly increasing in the posterior belief p, then $\mathbb{E}_{x'}[W(p',x')] < \mathbb{E}_{x'}[W(p_0,x')]$ if and only if $p' < p_0$. As a result, $\mathbb{I}_{\{d_2^{CEO}=1\}} = \mathbb{I}_{\{\mathbb{E}_{x'}[W(p',x')]<\mathbb{E}_{x'}[W(p_0,x')]\}} = \mathbb{I}_{\{p'<p_0\}}$, indicating that the CEO will leave her current match if and only if $p' < p_0$. At the same time, equation (A.11) and equation (A.12) could be simplified to:

$$w(p,x) = (1-b^{NB})b + b^{NB}[p\mu_H + (1-p)\mu_L] + b^{NB}\kappa\theta \times (1-\psi\mathbb{I}_{\{p'<p_0\}}) \tag{A.17}$$

and

$$r(p,x) = (1-b^{NB})[p\mu_H + (1-p)\mu_L - b] + \mu_z - b^{NB}\kappa\theta \times (1-\psi\mathbb{I}_{\{p'<p_0\}}). \tag{A.18}$$

□

A.3. *Proof of Proposition 2*

Proof. In this appendix, I prove the existence, the uniqueness, and the stationarity of the equilibrium.

First, I prove that the general equilibrium in Proposition 2 uniquely exists.

I denote the domain of firm productivity x as Ω_x. In equilibrium, the "free-entry" condition of Equation (14) is:

$$J^\star(p_0) = \sum_{\Omega_x} J(p_0,x) \times h(p_0,x) = \frac{1}{1+r}\kappa\lambda^{\frac{\eta}{1-\eta}} \tag{A.19}$$

defining a continuous and strictly increasing relation between $J^\star(p_0)$ and the job-finding rate λ over the domain Ω_x. It follows that $\frac{\partial J^\star(p_0)}{\partial\lambda} > 0$ and $J^\star(p_0) > 0$ for a given λ. The second relation between the equilibrium entry value $J^\star(p_0)$ and λ comes from the value function of the matched firm $J(p,x)$. Setting $p = p_0$ and converting to a function of the job-finding rate λ, it follows that:

$$\begin{aligned} J(p_0,x) = &(1-b^{NB})[p_0\mu_H + (1-p_0)\mu_L - b] + \mu_z - b^{NB}\kappa\lambda^{\frac{1}{1-\eta}} \\ &\frac{1-\delta}{1+r}\mathbb{E}^F(p',x') \times \mathbb{I}_{\{\mathbb{E}^F(p',x')>0\}}, \end{aligned}$$

where $\mathbb{E}^F(p',x') = \psi\lambda\mathbb{E}_{x'}[J(p',x')] \times \mathbb{I}_{\{d_2^{CEO}=0\}} + (1-\psi\lambda)\mathbb{E}_{x'}[J(p',x')]$. Using the chain rule, for each given productivity realization x, it follows that:

$$\frac{\partial J(p_0,x)}{\partial\lambda}|_x = -\frac{b^{NB}}{1-\eta}\kappa\lambda^{\frac{-\eta}{1-\eta}} \leq 0.$$

As a result, summing over the whole domain of x, the equilibrium entry value $J^\star(p_0)$ is a decreasing function of the job-finding rate λ. Moreover, as long as the CEO has a positive bargaining power $b^{NB} > 0$, $J^\star(p_0)$ is a strictly decreasing function of λ. Similarly, I can also show that the decreasing relation is bounded below from 0. Taking these two functions together, the continuity and bounded conditions insure that the intersection exists. In addition, from monotonicity, it is unique. As a result, there exits a $\lambda^\star$ such that the equilibrium in Proposition 2 uniquely exists.

Second, I prove that the general equilibrium in Proposition 2 is stationary

I denote the stationary distribution mass at $x = x_i$ and $p = p_j$ as $h^\star_{ij}$. Also I denote the transition matrix of firm productivity as $\pi(a(s),i) = \Pr(x' = i|x = a(s))$, and the optimal policy by solving the value function of the matched firm as *Policy**.

Then the invariant distribution $h^\star$ solves the following two equations:

If $p_j \neq p_0$, then:

$$h^\star_{ij} = h^\star_{ij} + \sum_{a(s)} \pi(a(s),i) \times h^\star_{(a(s),b(s))} - \delta h^\star_{ij} \\ - (1-\delta)\left\{h^\star_{ij} \times \mathbb{I}_{\{p_j \leq \underline{p}^\star\}} - \psi\lambda h^\star_{ij} \times \mathbb{I}_{\{\underline{p}^\star < p_j < p_0\}}\right\}, \tag{A.20}$$

where $Policy^\star(x_{a(s)}, p_{b(s)}) = p_j$.

If $p_j = p_0$, then:

$$h^\star_{ij} = h^\star_{ij} + \sum_{a(s)} \pi(a(s),i) \times h^\star_{(a(s),b(s))} - \delta h^\star_{ij} \\ + \psi\lambda \sum_{\underline{p}^\star < p_{b(s)} < p_0} \pi(a(s),i) h^\star_{(a(s),b(s))} + \lambda \sum_{p_{b(s)} \leq \underline{p}^\star} \pi(a(s),i) h^\star_{(a(s),b(s))}, \tag{A.21}$$

where $Policy^\star(x_{a(s)}, p_{b(s)}) = p_j$.

After solving the invariant distribution $h^\star$, the transition between state pair (x_i, p_j) and $(x_{a(s)}, p_{b(s)})$ also includes two cases:

If $p_j \neq p_0$, then:

$$\Pr\left((x_i,p_j)|(x_{a(s)},p_{b(s)})\right) = \pi(a(s),i) \times h^\star(a(s),b(s)) \times \mathbb{I}_{\{Policy^\star(x_{a(s)},p_{b(s)})=p_j\}}. \tag{A.22}$$

If $p_j = p_0$, then:

$$\Pr\left((x_i,p_j)|(x_{a(s)},p_{b(s)})\right) = \pi(a(s),i) \times h^\star(a(s),b(s)) \times \mathbb{I}_{\{Policy^\star(x_{a(s)},p_{b(s)})=p_j\}} \\ + \psi\lambda \sum_{\underline{p}^\star < p_{b(s)} < p_0} \pi(a(s),i) h^\star_{(a(s),b(s))} + \lambda \sum_{p_{b(s)} \leq \underline{p}^\star} \pi(a(s),i) h^\star_{(a(s),b(s))}. \tag{A.23}$$

To show that the distribution $h^\star$ at equilibrium is stationary, I need to show that the transition between any two states is both irreducible and positive recurrent. To prove that the transition between state pairs is irreducible, it is equivalent to show that $\forall (i,j),\ (a(s),b(s) \in \Omega$, $\Pr(x_{n+m} = x_i, p_{n+m} = p_j | x_n = x_{a(s)}, p_n = p_{b(s)}) > 0$, where Ω is the state-pair space. First, $\pi(a(s),i) > 0$ given that the idiosyncratic part of firm productivity is i.i.d. Second, $h^\star(a(s),b(s)) > 0$ by equations (A.20) and (A.21). Third, the Bayes' rule in equation (2) suggests that $\mathbb{I}_{\{Policy^\star(x_{a(s)},p_{b(s)})=p_j\}}$ could adjust both upwards and downwards. As long as the idiosyncratic shock is extremely volatile, namely σ_z is large, then $\Pr(Policy^\star(x_{a(s)}, p_{b(s)}) = p_j) > 0$ even j and $b(s)$ are sufficiently far away. Lastly, the exogenous turnover shock dissolves matches in each state pair at a rate of $\delta > 0$. Hence, there is no absorbing state in the equilibrium. Overall, based on the four equations above and the four arguments, the transition between state pairs is irreducible. Moreover, given that the state space is finite in the model, a finite irreducible Markov chain is automatically positive recurrent.

□

Appendix B. Estimation

Appendix B presents the algorithm used to solve the model, along with the estimation procedure.

B.1. Model Solution

Given the heterogeneity of CEO-firm matches in my model, I need four components to solve the general equilibrium: the value function of the matched firm $J(p,x)$, the separation policy $d^F(p,x)$ and the switching policy $d_2^{CEO}(p,x)$, the invariant distribution $h(p,x)$, and the equilibrium "free-entry condition" $\mathbb{E}_{x'}[J(p_0,x')]$.

The first step is to simplify the double-sided matching problem. The Nash bargaining in equation (8) implies that separation thresholds of the CEO and the firm coincide in the equilibrium. As a result, by solving the value function of the matched firm $J(p,x)$, I equivalently solve the value function of the CEO as $W(p,x) - U = \frac{b^{NB}}{1-b^{NB}} J(p,x)$.

Second, the difficulty in finding a numerical solution lies in the aggregate market condition. Consequently, I iterate over the aggregate market condition in spirit of Aiyagari (1994) in the following steps:

1. **Step 0: Discretize the state space**
   I let the posterior belief p evenly spaced with 40 points in an interval of (0,1). To discretize firm profitability, I first transform a firm's productivity shock into 20 discrete states in $[-3\sigma_z, +3\sigma_z]$ using the discretization method of Tauchen (1986). Then, I let the first 20 grid points of firm productivity representing "low" quality matches $x = \mu_L + z$, while the last 20 ones represent "high" quality matches $x = \mu_H + z$. Although agents in the model cannot observe the distinction, the "true" transition matrix of firm productivities repeats the transition matrix of z in the upper-left and lower-right corner.

2. **Step 1: Start with an initial guess of the aggregate market condition**
   I start with an initial guess of the expected entry value of firm $\mathbb{E}_{x'}[J(p_0,x)]^{(0)}$, calculate the job-finding rate $\lambda^{(0)}$ and the initial market "tightness" $\theta^{(0)}$ from equation (14).

3. **Step 2: Given the initial market condition, I solve the value function of the firm**
   Coupled with the initial guess of $\mathbb{E}_{x'}[J(p_0,x)]^{(0)}$, I solve the value function of firm $J(p,x)$ in equation (A.2) via value function iteration. At convergence, it produces the value function $J^{(0)}(p,x)$, its optimal separation policy $d^{F(0)}(p,x)$, and switching policy $d_2^{CEO(0)}(p,x)$.

4. **Step 3: Given optimal policies in Step 2, I solve the invariant distribution**
   After obtaining the optimal turnover policy $d^{F(0)}(p,x)$ and switching policy $d_2^{CEO(0)}(p,x)$ from Step 2, along with the law of motion of belief update $p'(p,x)$, I iterate over the invariant distribution $h(p,x)$. At convergence, it produces the stationary invariant distribution $h^{(0)}(p,x)$ given the initial guess of the market condition.

5. **Step 4: Update the aggregate market condition**
   In this step, I use firm's value function $J^{(0)}(p,x)$ and the invariant distribution $h^{(0)}(p,x)$ to

update the expected entry value of the firm $\mathbb{E}_{x'}[J(p_0,x)]^{(1)} = \sum_{i=1}^{\#xgrid} J(p_0,x_i)^{(0)} \times h^{(0)}(p_0,x_i)$. Then I update the job-finding rate $\lambda^{(1)}$and the market tightness $\theta^{(1)}$ from equation (14).

6. **Step 5: Repeat until convergence**
I repeat Step 2 - Step 4 utill the expected entry value converges $\mathbb{E}_{x'}[J(p_0,x)]^{(n+1)} = \mathbb{E}_{x'}[J(p_0,x)]^{(n)}$. Then $J(p,x)^{(n+1)}$ solves the firm's optimal problem, along with $d^{F(n+1)}(p,x)$ and $d_2^{CEO(n+1)}(p,x)$. At the same time, $h^{(n+1)}(p,x)$ characterizes the equilibrium distribution over the state space.

B.2. Estimation

To describe the estimation procedure, I denote x_i be an i.i.d. data vector, and $y_{ik}(b)$ as an i.i.d. simulated vector, where i takes value from 1 to NT (N is the cross-sectional length and T is the time-series length of the simulated sample), and k takes value from 1 to K (the number of times the model is simulated). In my simulation, I roughly follow Michaelides and Ng's (2000) "ten times rule" and generate 10,000 matched and unmatched firms with respect to the equilibrium market tightness and the stationary invariant distribution. Specifically, I multiply the invariant distribution matrix $h^\star(p,x) \times 10,000 \times \lambda^\star$, and then round each number to its closest integer, to represent matched firms. At the same time, the number of idle firms equals to $10,000 \times (1-\lambda^\star)$. I simulate the panel for the T period, equaling to the maximum tenure length of each sample under estimation. Then I drop the first $T-20$ period to replicate the 20-year sample length and to avoid the impact from initial conditions. Last, I maintain the free-entry condition in each simulated period t. Specifically, I let the equilibrium "tightness" $\theta^\star$ to hold. In another word, I allow firm entry and exit for the entire estimation spell.

In order to identify the set of parameters $b = [\mu_H, \mu_L, \mu_z, p_0, \kappa, \sigma_z, b^{NB}, \psi]$, I match a set of simulated moments $h(y_{ik}(b))$ with data moments $h(x_i)$ defined in Section 3. The identification of the parameter vector is from minimizing the quadratic distance between the data moments and the simulated moments, which is defined as:

$$g_{NT}(b) = \frac{1}{NT}\sum_{i=1}^{NT}[h(x_i) - \frac{1}{K}\sum_{k=1}^{K} h(y_{ik}(b))] \tag{B.1}$$

and

$$\hat{b} = \arg\min_b g_{NT}(b)'\hat{W} g_{NT}(b) \tag{B.2}$$

in which $\hat{W}$ is the inverse of sample covariance matrix of moments. I use Erickson and Whited's (2000) influence function approach to get $\hat{W}$. There are two types of variables in the estimation. I use variables after regression (15) to estimate the model. In contrast, I use variables before regressions to calculate influence functions of mean moments, those after regressions to calculate the influence functions of other moments except for the AR(1) moments. Lastly, I use an analogous regression in Han and Phillips (2010) to calculate influence functions of AR(1) moments.

To calculate the standard error of the estimated parameter vector $\hat{b}$, I use the following asympotOtic distribution properties:

$$\sqrt{n}(\hat{b} - b) \to^d N(0, avar(\hat{b})), \tag{B.3}$$

where the sample approximation of the variance is:

$$avar(\hat{b}) = (1+\frac{1}{K})[\frac{\partial g_n(b)}{\partial b}\hat{W}\frac{\partial g_n(b)}{\partial b'}]^{-1}, \tag{B.4}$$

where the optimal weighting matrix $\hat{W}$ is the same one in equation(B.2). I cluster all standard error at the two-digit SIC code as in Nikolov and Whited (2014).

Additionally, to calculate the standard error of the simulated moments, the covariance matrix of the two-step estimator is:

$$\hat{W}_{new} = (\frac{\partial g_n(b)}{\partial b}\Omega^{-1}\frac{\partial g_n(b)}{\partial b'})^{-1}, \tag{B.5}$$

where:

$$\hat{\Omega} = \frac{1}{NT}\sum_{i=1}^{NT}[g_n(b) - \frac{\partial g_n(\hat{b})}{\hat{b}}\phi^b(b)]\sum_{i=1}^{NT}[g_n(b) - \frac{\partial g_n(\hat{b})}{\hat{b}}\phi^b(b)]' \tag{B.6}$$

and $\phi^b(b)$ is the influence function for b.

Appendix C. Data

Turnover Classification

This paper features a unique turnover sample based on case-by-case analyses. In the first step, I categorize turnovers into exogenous and endogenous turnover by searching news and tracking career paths around and after CEO departure. The exogenous turnover include the following cases, with a finer classification labelled in parentheses:

1. The CEO position is on an interim basis (A1), co-CEO (A2), reappointment (A3), or due to the mis-recording when co-CEO/interim CEO is promoted to the sole-CEO position (A4).

2. The turnover is at the time one of the following events happens: company M&As (B1), bankruptcies (B2), spinoffs (B3), buyouts (B4), selling companies (B5), re-organizations (B6), privatizations (B7), liquidations (B8), regulatory bans (B9), or takeovers by the parent firm or government (B10).

3. The CEO reaches her mandatory retirement age and the firm explicitly quotes this reason in the public annoucement (C1), leaves the firm inherent to a pre-annoucend transition plan at least 6-month ahead (C2), is more than 70 years old (C3), or does not take an equivalent position utill to the end of the sample period after her departure (C4).

4. The CEO quits due to identified personal hobbies and interests unrelated to business (D1), political appointment (D2), quits to academic positions (D3), or other activities with tractable records (D4).

5. The CEO quits due to identified health reason (E1), death (E2), explicitly-cited family reasons (E3), or other reasons related to personal and family well-beings (E4).

6. The turnover case is an internal transfer (T).

7. The CEO is a turnaround specialist (TA).

In any of the above case, I let the exogenous turnover indicator $EXO\_Turnover = 1$. After deleting ambiguous cases, the rest are endogenous turnovers $END\_Turnover = 1$. In the next step, I track the career path of each CEO after her departure to select a subsample of CEOs performing on-the-job search ($OTJS = 1$) from endogenous turnovers if the CEO is able to launch onto an equivalent position within one year of departure. I also tag on-the-job searches for a future robustness check:

1. The departing CEO finds an equivalent position in another publicly-traded firm listed on major exchanges within one year of her departure (F1).

2. The departing CEO launched onto an equivalent position (i.e. CEO, President, Principal, major Partner, etc.) in a private-equity owned firm (G1).

3. The CEO founds her own firm:

   (a) Consulting firm (H1).
   (b) Not consulting, but similar to a post-retirement position (H2).
   (c) In the same industry as the firm where she had been employed as CEO (H3).
   (d) Private equity (H4).

In the last step, I categorize all the rest of endogenous turnovers ($END\_Turnover = 1$) into the following classes. They are:

1. The departing CEO finds another equivalent job or founds her own firm after one year of her departure (J1).

2. The company goes bankruptcy (J2), M&A or restruction (J3) after her departure.[18]

3. The company is in financial difficulties at the time CEO leaves the firm (J4).

4. The departing CEO is the CEO of another company at the time of turnover (J5).

5. The company or the CEO is under legal or financial investigation, inducing the turnover event (J6).

[18] The difference between the "B" category and here lies in the timing. If the CEO separation is simultaneously annoucend, or happens exactly at the time the event happens, it belongs to the "B" category as exogenous turnover. Otherwise, if the turnover and the event do not happen simultaneously, yet it may be the poor performance of the CEO induces the consequent event, I classify the turnover event in the "J" category as endogenous turnover. A robustness check will relax the assumption here.

6. The firm explicitly cites bad performance as a reason of separation (J7).

7. There is a cited conflict between CEO and the board, or with the major shareholders, founders, labor unions, etc. of the firm. Or the board is experiencing a proxy fight at the time the turnover occurs (J8).

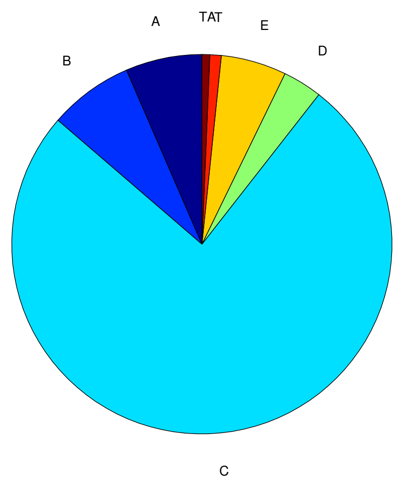
8. The turnover is to separate the CEO from the chairman role (J9).

Figure C1 depicts pie graphs of the composition of each classification.

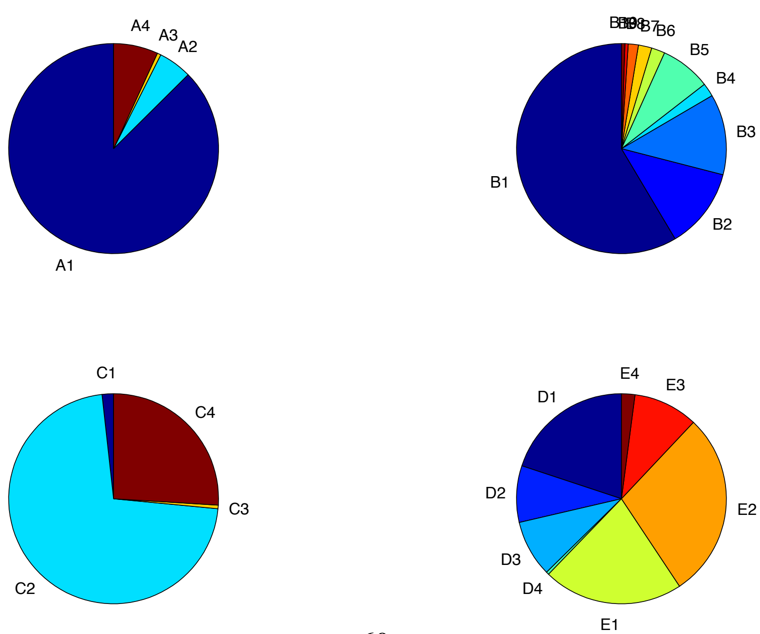
Figure C1. Turnover Classification

Figure C1 shows the turnover classifications. In Panel A, all exogenous turnover is decomposed into 7 categories, and Panel B further provides the composition within each category. In Panel C all endogenous turnover excluding on-the-job searches is further decompsed into 10 categories, and Panel D shows the classification under on-the-job search. In each case, the symbol represents the classification criteria defined in Appendix C.

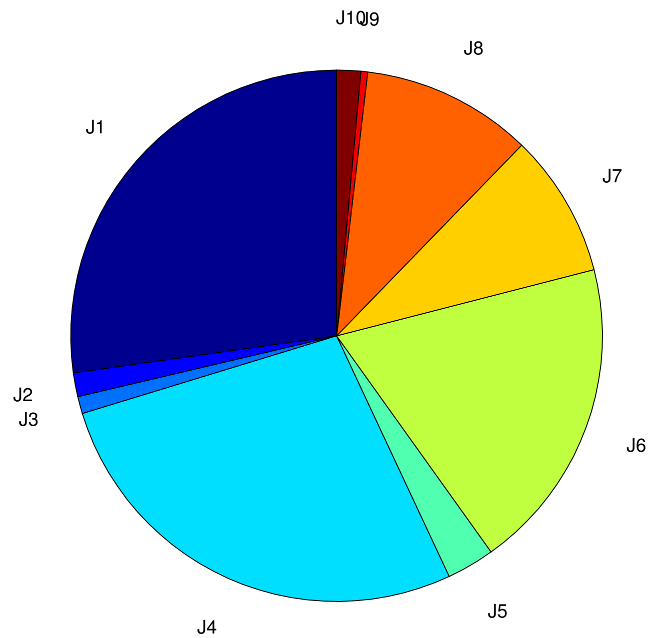
Panel A: Composition of Exogenous Turnover



Panel B: Composition of Exogenous Turnover - 2



Panel C: Composition of Endogenous Turnover (excluding on-the-job search)



Panel D: Composition of On-the-Job Search