A Structural Estimation of the Cost of Suboptimal Matching in the CEO Labor Market

Jordan Nickerson*
Boston College

December 26, 2014

*The author can be reached at Jordan.Nickerson@phd.mccombs.utexas.edu. I thank Andres Almazan, Aydoğan Alti, Jonathan Cohn, John Griffin, Jay Hartzell, Ari Kang, Richard Lowery, Robert Parrino, Sheridan Titman, and Toni Whited for their helpful discussion.

A Structural Estimation of the Cost of Suboptimal Matching in the CEO Labor Market

Abstract

Using a structural model, I examine the distortionary effects of frictions in the CEO labor market. Firms experience productivity shocks over time and either outgrow or underutilize their incumbent CEO's talent, but keep their manager to avoid a switching cost. The decision to replace a manager depends on the magnitude of the cost and dispersion of CEO talent. I find CEO talent to be quite heterogeneous. Additionally, I estimate the switching cost to be 20% of the median firm's annual earnings. While reduced-form estimates of the switching cost serve as a lower bound on the reduction in firm value, they underestimate the overall effect which also includes the resulting inefficient firm-CEO matches. Using counterfactual analysis, the switching cost is estimated to decrease the median firm's value by 4.8%, four times larger than the reduced-form estimate. While firms experience an observable decrease in earnings when finally replacing CEOs, I find evidence of a considerable unobservable cost associated with the inability of firms and managers to be optimally matched in the cross-section.
Keywords: CEO Turnover, Executive Compensation, Structural Estimation

1 Introduction

There is an extensive literature focusing on the importance of CEOs and their impact on firm value. Gabaix and Landier (2008) and Terviö (2008) estimate the importance and differentiation among CEOs based on the compensation they receive. However, an assumption common to both models is socially optimal matching between firms and managers in a frictionless labor market, where matches are formed in a positive assortative nature. In reality, a firm faces an explicit cost when switching managers, which includes severance pay and lost productivity while the new CEO learns the company. Thus, changes to a firm's value and overall productivity due to technology shocks, exogenous industry contractions or expansions, and other factors beyond the control of the firm will result in a distortion to positive assortative matching in the cross-section. It is unclear how implications from a model of efficient matching are affected by such a friction. Furthermore, reduced-form methods can establish a lower bound for the economic magnitude of this friction by measuring the decrease in a firm's earnings around the turnover of its CEO. However, this estimate understates the true reduction in firm value which also encompasses the value destroyed by an inefficient match between firms and managers as a result of the switching cost.

In this paper, I estimate a dynamic matching model of firms and CEOs with an embedded switching cost that results in inefficient matching in the cross-section. This allows me to re-evaluate the importance of CEO talent while simultaneously estimating the reduction in overall firm value relative to a frictionless economy where firms and CEOs are always optimally matched. In the model, a firm experiences a series of shocks to asset productivity that result in an inefficient match with the incumbent CEO. The firm can elect to replace the existing manager with an optimal replacement from the CEO labor pool but must pay a fixed cost to do so. The decision to retain or replace the incumbent CEO is the outcome of a dynamic programming problem based on the current productivity of the firm's assets in place as well as the existing manager's talent level.

I estimate this structural model using data on executive compensation and tenure along with firm profitability for large U.S. firms from 1992-2011, gauging the impact of inefficient matching on a firm's overall value. I find that the cost of switching managers constitutes a non-negligible percent of a firm's annual profits. Using counterfactual analysis, I find that this cost results in a 4.8% decrease in median firm value. While this is partially attributable to the explicit cost paid by the firm when switching managers, I find that the implicit cost from the resulting suboptimal matching scheme constitutes 76.2% of the total reduction in firm value.

The dynamic model presented here is built on the groundwork set forth in Gabaix and Landier (2008). They present a static model in which heterogeneous managers with publicly observable talent and heterogeneous firms which differ in the amount of their productive assets meet in a frictionless labor market. The positive assortative matches formed in equilibrium represent the value optimizing assignment of CEOs to firms. In contrast, this paper's model departs from this static framework by allowing the productivity of a firm's assets to vary over time. Each period, a firm receives a positive or negative shock to the productivity of its assets. These shocks reduce the efficiency of the match with the incumbent CEO. Thus, the firm has an incentive to enter the labor market to be optimally matched with a new CEO.

I embed a switching cost in the model that the firm must pay in order to reenter the CEO labor market, discouraging the firm from continually seeking an optimal replacement. This switching cost is meant to represent the search cost to find an adequate replacement, a period of reduced productivity during the transition period, severance pay, and possible firm-specific knowledge that must be acquired by the new manager, among other factors.[1]

[1] Almazan and Suarez (2003) look specifically at severance pay and entrenchment to prevent the replacement of an incumbent CEO with a superior replacement.

Therefore, if a firm chooses to retain its incumbent CEO following a shock to productivity due to the switching cost, the resulting match will be an inefficient one relative to a frictionless economy. Furthermore, the inefficiency of this match increases following a series of positive or negative shocks which aggregate together.

The firm effectively chooses how much the productivity of its assets can change before replacing its manager with a more suitable one. This decision, and hence the degree of inefficiency that can be maintained in equilibrium, is also influenced by the dispersion in CEO talent. If managers are virtually homogeneous, a firm has little incentive to replace its existing manager and suffer a switching cost. It is also influenced by the current matches of CEOs as well the anticipated competition from other firms in the labor market, which are endogenously determined in equilibrium.

I estimate this dynamic model to gleam some insight into the importance of CEOs and their contribution to firm value. This also allows me to generate counterfactual scenarios to gauge the destruction in firm value as a result of such frictions in the CEO labor market. I estimate parameters related to the dispersion of managerial talent, the cost associated with changing managers, and the volatility of a firm's productive assets using the Simulated Method of Moments (SMM) approach. I identify these parameters using data on the persistence and cross-sectional properties of earnings, earning patterns around CEO turnover, the frequency of these turnover events, relative changes in firm values and executive compensation.

Empirically, I find that CEOs are quite heterogeneous in their ability, with the most talented manager able to generate gross profits greater than the average CEO of the 500 largest firms by a factor of 1.78. I estimate the tail thickness of this managerial distribution to be consistent with the findings of Gabaix and Landier (2008). The cost a firm experiences when switching managers is estimated to be 2.18% of its assets in place, or 20.0% (23.2%) of the median (mean) firm's yearly return on assets. However, this decrease in a firm's earnings only accounts for one-fourth of the reduction in its value. The remaining reduction in firm value is attributed to the inefficient CEO match the firm tolerates in order to avoid suffering the switching cost on a more frequent basis. Using counterfactual analysis, it is estimated that the median (mean) firm's value would increase by 5.06% (4.21%) were it able to replace managers in the absence of a switching cost.

While the structural model used here has as its basis the model of Gabaix and Landier, the two also differ along many avenues. My model endogenizes the distribution of available managers in the labor pool each period. Also, while they take firm value as exogenously given when identifying the talent distribution, in this paper it is endogenously determined as a function of both manager talent and wages paid. Finally, the two works differ in their scope and focus. Gabaix and Landier explain the rise in CEO compensation over the recent decades using a competitive equilibrium framework. In addition to extending this framework to a dynamic setting, I also seek to examine the implications that inefficient matching between firms and managers has on a firm's value.

The model and results in this paper have ties to three distinct strands of literature, the first of which looks at the value generated by a CEO. Rosen's (1981) superstar effect provides the foundation on which this model rests. An adaptation of the assignment model derived in Sattinger (1993) is crucial when solving the model in equilibrium, which both Terviö (2008) and Gabaix and Landier (2008) apply to CEOs to estimate talent levels based on wages and relative firm sizes. Nickerson (2013) also examines the relationship between firm size and wages set in competitive equilibrium, focusing on time-varying demand for CEO talent. Bertrand and Schoar (2003) and Graham, Li and Qiu (2012) examine the CEO's effect on firm performance across multiple firms.[2] Bennedsen, Perez-Gonzales and

[2]Kaplan, Klebanov and Sorensen (2012) map this effect on performance to observable characteristics using evaluations of potential CEOs, identifying particular CEO traits that are positively related to future performance. Additionally, Adams, Alemida and Ferreira (2005) find that more powerful CEOs have a stronger impact, measured by the volatility of stock returns, and Carter, Franco and Tuna (2010) and Falito, Li and Milbourn (2013) find that executive compensation is related to observable CEO characteristics.

Wolfenzon (2012) measure a CEO's impact by examining productivity downtime resulting from CEO hospitalization. This paper extends this literature by focusing on a dynamic setting in which the productivity of a firm's assets, and thus the value created by its CEO, vary over time and the resulting matching outcome between firms and managers.

This paper is also related to the determinants of CEO turnover. Parrino (1997) finds an industry component to a CEO's job fragility. Huson, Parrino and Starks (2001) find an increase in turnover over time without an increased sensitivity of turnover to firm performance. Murphy and Zabojnik (2007) propose a model that explains recent trends in CEO wages and turnover by an increase in transferable, cross-company skills. Finally, Eisfeldt and Kuhnen (2013) propose a competitive assignment model, matching firms and managers on multiple dimensions to explain the relationship between CEO turnover and both relative and absolute performance measures. I contribute to this literature by examining the tradeoff a firm faces between an incumbent CEO who is an inefficient match and an optimal replacement who is accompanied by a switching cost.

Finally, the choice of structural estimation in this paper is related to a large strand of literature examining numerous firm decisions. Structural modeling has been previously used to study other firm choices involving investment financing (Bond and Meghir (1994), Hennessy and Whited (2005, 2007)), investment decisions (Kang, Liu and Qi (2010)), optimal managerial ownership (Coles, Lemmon and Meschke (2012)) and compensation (Taylor (2013)), and market participation (Roberts and Tybout (1997)). Within this strand of literature, Taylor (2010) presents a structural learning model used to measure the switching cost associated with CEO turnover. While I also study a firm's switching cost, the two works differ on their primary focus. Fundamentally, while Taylor's model measures the cost to change managers in a learning environment, the focal point of this paper is on the magnitude that such a switching cost would distort the optimal matching of firms and CEOs and the overall reduction in firm value. Additionally, while Taylor focuses on the firm's turnover decision in a partial equilibrium, the implications in this paper are built upon a general equilibrium model that includes competition for CEOs and competitive wages.

The rest of the paper is organized as follows. The basic model and equilibrium solution are presented in Section 2, while the data and details of the estimation process are outlined in Section 3. Section 4 is reserved for a discussion of the estimation results. Section 5 presents a counterfactual outcome measuring the overall distortionary effect. Section 6 concludes.

2 The Model

I begin by briefly presenting the static matching model of both Gabaix and Landier (2008) and Terviö (2008) and the resulting equilibrium wage function to show the intuition of the equilibrium wages and matching procedure. I will then extend the model into a dynamic framework and discuss the resulting departures from the static equilibrium.

2.1 Static Model and Equilibrium

The model environment consists of a continuum of firms with heterogeneous levels of productive assets $A \in \left[\underline{A}, \overline{A}\right]$, who's distribution is characterized by $f_s(A)$. Additionally, the economy is populated with a continuum of managers with fully observable, heterogeneous talent levels, $\theta \in [\underline{\theta}, \theta_{max}]$, who's distribution is governed by $f_\theta(\theta)$. All individuals share a common reservation wage, $\underline{w}$.

Profits are a function of a firm's productive assets, the CEO's talent level and wages paid:

$$\pi = A \cdot \theta - w \tag{1}$$

Managerial talent is modeled as a multiplicative effect on a firm's gross profits. Thus, a superior CEO can make every dollar under her control more productive, relative to a less talented manager. In equilibrium, given observable talent, managers and firms are matched in accordance with the assignment equation of Sattinger (1979), as a result of the following equilibrium wage condition. Define the mapping of a manager in the pth percentile of the talent distribution as $T(p) \equiv f_{\theta}^{-1}(p)$. Similarly, let $S(p) \equiv f_{s}^{-1}(p)$, represent the productive assets of a firm in the pth percentile. Then, the equilibrium wages paid to a manager in the pth percentile with talent $T(p)$ have the following property:

$$\frac{\partial W(T(p))}{\partial p} = S(p) \cdot \frac{\partial T(p)}{\partial p} \tag{2}$$

As a result, each firm is exactly indifferent between the manager it is matched with in equilibrium and a manager with a talent level one epsilon greater, thus preserving positive assortative matching between firms and managers. Furthermore, the smallest firm in the economy with assets $\underline{A}$ will be matched with a manager of talent $\underline{\theta}$, who's wage will be set at her reservation level $\underline{w}$. From this fixed point and (2), competitive wages follow from the following equation:

$$W(\theta) = \underline{w} + \int_{0}^{F_{\theta}(\theta)} S(u) \cdot T(u)' \, du \tag{3}$$

However, while the static model gives insight into the formation of wages and the matches between firms and managers in equilibrium, it assumes that there are no frictions preventing the positive assortative matching between CEOs and firms. Therefore, I now present the dynamic model with an embedded switching cost examined in this paper.

2.2 Dynamic Model

Similar to the static model, the dynamic framework consists of a measure one continuum of heterogeneous firms and managers. In contrast to the previous model, firms are modeled as being infinitely lived while managers possess a finite lifespan. Additionally, each firm experiences a shock to its assets, possibly has its CEO retire or fired, and realizes profits within each period. Profits are modeled in a similar fashion to (1) with a few exceptions. Each period's profits remain a function of the productivity of a firm's assets, the talent of the manager, and equilibrium wages paid. However, these profits are also a function of an unobservable idiosyncratic noise component (ε_t), and if applicable a switching cost imposed if a firm's incumbent manager is replaced ($c_{replace}$) or retires (c_{retire}):

$$\pi_t = A_t \cdot (\theta_t + \varepsilon_t - 1\,(replace) \cdot c_{replace} - 1\,(retire) \cdot c_{retire}) - w_t \tag{4}$$
$$\varepsilon \sim N\left(0, \sigma_\varepsilon^2\right)$$

The cost of managerial replacement is modeled as being linear in a firm's productive assets. This single factor represents the search cost to find a suitable replacement, the decrease in productivity during the interim period, and other losses incurred by a firm, which likely increase with a firm's size.[3] I do not restrict the cost imposed on a firm for managerial retirement, which is also linear in productive assets, to be equivalent to managerial replacement. As a CEO approaches retirement age, the firm is able to seek out a successor in advance to avoid times of reduced productivity during the transition period. In the model, retirement follows an exponential arrival time with probability of δ for estimation feasibility. Finally, an additional idiosyncratic noise component is considered in the dynamic model. While not necessary when estimating a static model, when considering the estimation of a dynamic model this noise component allows for the inclusion of additional time-series moments useful in identification of the model's parameters.

While firm profits are modeled in a slightly different form in the model considered in this paper relative to the static model of Gabaix and Landier (2008), the dynamic nature of the model stems from changes to a firm's productive assets. These fluctuations encompass changes to a firm's assets in place, contractions or expansions to a firm's industry, or technol-

[3] Yermack (2006) finds that "golden handshakes" or severance pay for CEOs increases with firm value.

ogy shocks, all of which affect a firm's profits. All of these contributing factors are captured in the model by a firm's productive assets. However, in reality a firm can choose the level of assets in place. Therefore, the productive assets considered here can be interpreted as the optimal size of the firm given industry and technology conditions, relative to the other firms in the economy. To capture these changes, each period a firm receives a shock to its productive assets, governed by the following processes:

$$A_t = A_{t-1} \cdot x_t \tag{5}$$
$$x_t \sim U\left(1-\gamma, 1+\gamma\right)$$

The choice that shocks be drawn from a uniform distribution is made for computational convenience, although the consideration of other distributions is easily done with a concession to the estimation time.

Figure 1 illustrates the timing of the model within each time period. At the onset of the period, the firm experiences a shock to productive assets. While the timing of the shock is modeled at the beginning of the period, it can also be thought of as gradually occurring throughout the length of the previous period, and will now be acted upon. Following this shock, each firm faces the possibility of its CEO retiring, which will occur with probability δ. Immediately following the realization of retirements, all remaining firms must choose to retain the incumbent manager or enter the labor market and be matched with a replacement. If a manager is fired from their current firm they re-enter the labor pool. There, they will be joined by a cohort of newly born CEOs of measure Ît' who take the place of the individuals who retired. At this point all firms in need of a CEO, by choice or because their incumbent CEO retired, enter the labor market and are simultaneously matched with a new manager. While broken down into three different steps, these events can be thought of as occurring almost simultaneously at the beginning of the period. Following this matching process, every firm commences production. Finally, profits are realized at the end of the period at which point the process repeats itself.

For simplicity, each periods profits are immediately paid out as a dividend to shareholders. Thus, there are only two sources of variation in the expected value of a firm: changes in a firm's productive assets and changes in firm management. There are no other factors that are material to a firm's expected dividend stream. While asset productivity and CEO retirement are exogenous in the model, the decision to replace managers is not. Therefore, each period a firm whose manager does not retire chooses $d_t \in \{retain, replace\}$ in a manner that maximizes the total firm value in expectation, V_t:

$$\max_{\{d_s\}_{s=t}^{\infty}} V_t = E_t \left[\sum_{s=t}^{\infty} \beta^{s-t} \pi_s \right] \tag{6}$$

where π_s is defined in equation (4) and β is the firm's per-period discount rate. The series of the firm's decisions that maximizes expected value is arrived at by solving the dynamic programming problem discussed next.

2.3 Dynamic Programming Problem

As a firm experiences shocks to its productivity, it must continually decide to either retain its existing manager or replace the CEO with another individual from the labor pool. If the firm has experienced a series of positive shocks, the firm has an incentive to reenter the labor market and be matched with a new manager who has a higher talent level. Conversely, if the firm has experienced a series of negative shocks, the existing manager is more talented than the replacement obtained from the labor market. However, as discussed below, equilibrium wages must be set in such a way that firms who commit to participate in the labor market are matched to managers in a positive assortative manner.[4] This implies that the current

[4]Note, this does not imply that firm and managers in the economy are positive assortatively matched. However, once a firm pays the switching cost to replace their existing manager and enters the labor market,

manager is relatively too costly relative to a replacement, giving the firm an incentive to switch managers. However, while a firm has the incentive to switch managers, this decision is also accompanied by a switching cost.

Therefore, the firm's decision to retain or replace the existing manager can be represented by the following Bellman equation, which is solved for in Appendix I:

$$\begin{aligned} V(\theta, A_t) &= \max_{d_t} \left\{ V(\theta, A_t)^{retain}, V(\theta, A_t)^{replace} \right\} \\ V(\theta, A_t)^{retain} &\equiv A_t \cdot \theta - W(\theta) + \beta \cdot (1-\delta) \cdot E_t \left[V(\theta, A_{t+1}) \right] \\ &\quad + \beta \cdot \delta \cdot E_t \left[V(\theta(A_{t+1}), A_{t+1}) - A_{t+1} \cdot c_{retire} \right] \\ V(\theta, A_t)^{replace} &\equiv V(\theta(A_t), A_t) - A_t \cdot c_{replace} \end{aligned} \tag{7}$$

where $W(\theta)$ is the equilibrium wages paid to a manger with talent level θ, and $\theta(A)$ is the talent level of a manager matched to a firm with productive assets A in the labor market. Therefore, a firm's decision to retain or replace the incumbent manager depends on the distribution of managers available in the talent pool in addition to the distribution of competing firms who are also seeking a new manger from the talent pool.

If a firm chooses to retain their existing manager, the firm's value consists of this period's net profits after paying the manger's wages, plus the discounted expected future value of the firm. This future value equals the weighted average of the firm value conditional on the existing manager retiring or not in the next period. Alternatively, the firm can replace the existing manager by paying a switching cost, thereby allowing the firm to be matched with the optimal manager.

Figure 2 gives some insight into this optimization problem. The first panel compares the firm's expected value for both options available, retaining and replacing the current

in equilibrium there will be no other firm with a smaller level of productive assets also in the labor market who is matched with a more talented manager. Thus, while the labor market is matched in a positive assortative fashion, the economy as a whole is not.

CEO. The figure plots the ratio of the expected firm value conditional on retaining the existing CEO to the firm value conditional on replacing the CEO. Predictably, this ratio is at its greatest level when the firm is already matched with its optimal CEO. At this point, the firm's current level of productive assets would result in the firm being matched with a manager from the labor pool of equivalent talent to the incumbent CEO. However, doing so would require the firm to pay the switching cost. As the firm moves away from this point, the incumbent CEO becomes an increasingly inefficient match, offsetting the cost to replace the existing CEO. Furthermore, there are two points where the value of the firm becomes greater if the current CEO is replaced with a manager from the labor pool. This gives rise to an upper and lower level of productive assets, or "replacement thresholds" denoted by $\overline{A}(\theta)$ and $\underline{A}(\theta)$ respectively, for each CEO talent level. At these thresholds, it becomes optimal for the firm to replace the current manager and participate in the labor market.

Panel B plots these replacement thresholds as a percentage of the firm's productive assets when initially matched with the incumbent manager. While a firm with the smallest amount of productive assets requires roughly a 50% increase or 25% decrease in productive assets to replace managers, these switching thresholds increase with a firm's productive assets. A firm's choice to replace managers is influenced by the distribution of managers in the labor pool. If the dispersion in ability decreases among more talented managers in the labor pool, larger firms would have less of an incentive to replace their incumbent CEO for a replacement from the pool.

2.4 Managerial Talent Distribution

As previously noted, the decision to retain or replace an incumbent CEO depends not only on the cost that a firm incurs when changing CEOs, but also on the distributions of firms and managers participating in the labor market. In equilibrium, these distributions must be consistent with the optimal decision rule solving the dynamic programming problem the firm faces. Appendix II generates the stationary distribution of firms and managers in the labor market, endogenously determined as a function of the optimal decision rule.

While the firm's choice to participate in the labor market is contingent on the distribution of managers and competing firms in the labor market, these distributions are a function of the overall distribution of managers in the economy. Therefore, while the composition of the talent pool available for higher is endogenously determined in the model, the cross-section of all managers in the economy is exogenously determined.

To characterize this distribution, I rely heavily on the work of Gabaix and Landier (2008) who adapt the extreme value theorem to describe the talent levels of the economy's most talented managers. They propose that the change in talent levels for CEOs of the largest U.S. firms can be approximated by the following expression:

$$\theta'(x) = -Bx^{\alpha-1} \tag{8}$$

where x denotes the percentile of the talent distribution, with a decrease in ability associated with an increase in x. Therefore, I am able to characterize managerial talent with three parameters, θ_{max}, B and α. θ_{max} represents the ability of the most talented manager in the economy. B serves as a scaling factor that governs the average decrease in talent as you progress through the distribution. α is referred to as the tail-index, and determines the convexity of this decrease.[5]

2.5 Equilibrium Wages

While constructed in a similar fashion to the equilibrium wages in the static models of both Gabaix and Landier (2008) and Terviö (2008), their construction differs slightly in the dynamic framework presented here. The intuition behind the static model's wage structure

[5]See Gabaix, Laibson and Li (2005) for a list of common distributions and their tail indices

described by (2) is that wages must be set in such a way that any firm would be indifferent between being matched with their positive assortative counterpart or being matched with a manager that is marginally more talented but demands a higher wage. Similar intuition holds in the dynamic model considered here, with one difference.

While the wages outlined by (2) make a firm indifferent between two adjacent managers when only the current period's revenue is being considered, the choice of manager has an effect on the wages from multiple periods in a dynamic model. Therefore, wages must be set in such a way that the value of any firm seeking a new manager from the labor market will remain unchanged if they hire a manager marginally more talented than their equilibrium match. Suppose a firm with productive assets A is matched with a manger of talent level $\theta(A)$ in the labor market. Then, equilibrium wages must be set such that:

$$\frac{\partial W(T(p))}{\partial \theta(A)} = \frac{\partial V(\theta(A), A)}{\partial \theta(A)} \tag{9}$$

Note that the value of a firm, conditional on participating in the labor market, is a monotonic transformation of the firm's productive assets. Therefore, while firms are not positive assortatively matched with managers in the cross section of all firms, all firms and managers participating in the labor market are matched in such a fashion.

2.6 Formal Equilibrium Conditions

The optimal decision to replace a suboptimal manager, the distributions of CEOs and firms participating in the labor pool and competitive wages are all jointly determined in equilibrium. Therefore, the general equilibrium is one that satisfies the following conditions:

- Given the distributions $f_s^\star$ and $f_\theta^\star$ and a wage function $W^\star(\cdot)$, the optimal decision rule that solves equation (6) is characterized by $\underline{A}^\star(\theta_i)$, $\overline{A}^\star(\theta_i)$ for all θ_i.

- Given the optimal decision rule $\underline{A}^{\star}(\cdot)$, $\overline{A}^{\star}(\cdot)$ and the distributions $f_s^{\star}$ and $f_\theta^{\star}$, all wage offerings that satisfy equation (7) are characterized by the function $W^{\star}(\cdot)$.

- Given the optimal decision rule $\underline{A}^{\star}(\cdot)$, $\overline{A}^{\star}(\cdot)$, the stationary distribution of firm's participating in the labor market is $f_s^{\star}$.

- Given the distribution $f_s^{\star}$ and the optimal decision rule $\underline{A}^{\star}(\cdot)$, $\overline{A}^{\star}(\cdot)$, the stationary distribution of managers in the labor market is $f_\theta^{\star}$.

The general equilibrium solution is solved for numerically, as outlined in Appendix III.

3 Estimation Approach

3.1 Dataset

The validity of a structural estimation's results rest on two things, the underlying model's ability to represent reality and the moments used to identify the model. Therefore, the subset of the economy that the model is applied to must be carefully considered. The model presented herein makes two pivotal assumptions. The first assumption is that the value that a CEO adds is transferable and proportional across companies. Presumably, consistent with the findings of Murphy and Zabojnik (2004, 2007), CEOs of large firms rely more on the ability to manage people effectively and do not require as much firm-specific knowledge from the CEO. Therefore, it is plausible that their managerial ability is transferable across large firms. Additionally, the model also utilizes the extreme value theorem which is very robust in characterizing the tail of a distribution but loses its ability to describe the distribution as you move further towards the interior. For these reasons, the scope of this model should be restricted to large, non-specialized firms.

Therefore, I focus on the largest U.S. firms when estimating the model. I begin with the universe of managers reported in Execucomp from 1992 to 2011, which I merge with

Compustat. Given the specific nature of their industries, all financial companies (SIC codes 6000-6999) and utilities (SIC codes 4900-4999) are removed. All CEO successions are identified as a change in the annual CEO flag from one individual to another. Firms are then ranked based on their total market capitalization.[6] Because I seek to study only large firms, I retain only those CEOs whose firm ranking in the year prior to their appointment places them within the top 500 firms. However, such a methodology would under-sample CEOs with long lived tenures that are more likely to have entered their office prior to 1992. To counteract this issue, the terms of all CEOs in the top 500 firms in 1993 are also considered.[7]

Finally, the purpose of the model is to measure the distortionary effect that a switching cost has by preventing a firm from optimally replacing its CEO as the firm experiences shocks to productivity. Nevertheless, I must recognize that in reality other factors come into play regarding a firm's decision to change managers. For instance, the model presented in Taylor (2010) focuses on a firm's ability to learn about the quality of its manager, which is incorporated into the firing decision. Rather than adding more features to the model, reducing the model's tractability in the process, I instead exclude CEO spells where it is unlikely the turnover decision was made because of a change in the firm's productive assets. Therefore, I exclude any CEO spell whose tenure is two years or less. To avoid any bias this has when comparing empirical and simulated moments, I also apply this exclusion rule in all simulated spells.

Once the CEO terms are identified, financials of all corresponding firm-years are collected as well as the first fiscal year following the CEO's replacement. Profits are defined as a firm's earnings after depreciation divided by the average of lagged and contemporaneous total assets. CEO compensation is set equal to the reported total compensation.[8] All dollar

[6] Total market capitalization is set equal to Assets Total + Common Shares Outstanding * Price Close (Annual) - Common Equity - Deferred Taxes

[7] While some firms 1992 fiscal year filings are reported in Execucomp, some firms do not enter the sample until 1993. For this reason, 1993 was chosen instead of 1992.

[8] This corresponds to the Execucomp variable, TDC1, and is the sum of salary, bonus, restricted stock

values are converted to 2005 dollars using the GDP deflator. Summary statistics for the final sample are contrasted with the entire universe of Execucomp firms in Table I. Not surprising, firms in the final sample tend to be larger in both total assets and market capitalization, be more profitable, and have larger executive compensation packages.

Finally, I must identify the reason for each CEO's departure. Fortunately, within the context of the model, the dissolution of any match is mutually agreed upon by the firm and manager. The CEO would prefer wages equal to their market value while the firm would prefer a more optimal CEO match. The only classification that must be made is whether a CEO is *replaced* or *retired*. Therefore, any CEO who subsequently takes an officer position at another firm is classified as being *replaced*. For the remainder of the successions, all managers 62 years of age or older at the time of replacement are classified as being *retired* with the remainder being classified as *replaced*.

3.2 Model Parameters and Estimation

For a given parameter set $\Theta \equiv (\delta, \beta, \theta_{max}, B, \alpha, \gamma, \sigma_{\varepsilon}^2, c_{replace})$, the general equilibrium is computed numerically, as outlined in Appendix III.[9] To illustrate the resulting optimal policy function, Figure 3 plots the life cycle of one simulated firm and its policy function for a reasonable set of parameters. The dotted blue (black) line denotes the upper (lower) threshold of the firm's productive assets before the incumbent CEO is replaced. Notice that after the current CEO retires (hollow blue dots) the switching thresholds change. Following a CEO's retirement, the firm re-enters the labor market and is optimally matched. Therefore, because switching thresholds are a function of the incumbent CEO, they change following both CEO retirement and replacement.

Recall that the first parameters, δ, is the probability that the existing CEO retires. This

grants, Black-Scholes value of option grants, and LTIP payouts

[9]In the estimation, I set c_{retire} equal to zero, which is borne out in untabulated empirical results discussed later.

probability is taken as strictly exogenous in the context of the model, and is estimated using standard reduced-from techniques. In addition, I also abstain from estimating the second parameter, β, used to discount future cash flows because there are models far more suitable than the one presented here able to measure this factor. Therefore, I assume a discount factor of .90, but consider others in untabulated results for robustness sake.

While the first two parameters can be either estimated using a reduced-form approach, or we have reasonable priors regarding their value, the remaining six parameters must be estimated structurally. The first three of these, $(\theta_{max}, B, \alpha)$, jointly describe the distribution of managerial talent in the extreme tail of the population. The thickness of this tail is characterized by the tail index, α. Figure 4 illustrates the convexity of changes in managerial talent within the tail of the distribution.

While Figure 4 illustrates how the distribution of talent changes with α, it is harder to visualize how managerial talent varies with B. Fortunately, given values for θ_{max} and α, there is a one-to-one relationship between B and the average talent level of managers over a given percentile range. Therefore, I instead estimate $\overline{\theta} \equiv \frac{1}{500}\sum_{i=1}^{500}\theta_i$, the average of the 500 most talented individuals. Furthermore, this parameter also encompasses the average productivity of assets in place and should be interpreted as the average talent level scaled by this productivity factor.[10] Prior literature has begun to identify these values (Terviö (2008), Gabaix and Landier (2008)) using CEO compensation within the context of frictionless models. In contrast, if there is a friction whose effect on a firm's participation in the labor market is a function of the firm's size, the distribution of managers becomes endogenous.

The fourth parameter, γ, captures the magnitude of shocks to a firm's productive assets. Additionally, because there is an idiosyncratic noise component embedded in firm profitabil-

[10]While the difference in talent among CEOs matters for the firm's decision to retain or release a manager, the average value does not. The firm's optimal policy function will remain unchanged if every CEO's talent level was increased by a constant. However, this parameter is necessary when considering certain model moments.

ity, ε, the productivity shocks cannot be estimated using only the variance of a firm's profits over time without jointly considering the variance of this noise term. Similarly, the final value of interest, $c_{replace}$, which measures the cost of changing managers depends on the value provided by a more efficient firm-CEO match, making reduced-form analysis problematic. For these reasons, I elect to use a structural model to measure these parameters of interest.

3.3 Simulated Method of Moments

Given model parameters, Appendix III outlines the process to solve for the general equilibrium. However, the task remains to estimate the appropriate model parameters that correspond to empirical evidence. To estimate the model, I rely on the simulated method of moments (SMM), a technique that will be discussed briefly.[11] Given a set of parameters, Θ, rational firm follow an optimal policy function. Therefore, using this decision rule, the actions of a panel of firms within an economy can be simulated through time. From these firm decisions, a set of sample moments can be simulated. The SMM procedure is used to find the parameter set $\Theta^{\star}$ resulting in a set of sample moments that most closely matches the same set of moments measured empirically.

More formally, given a parameter set Θ, N simulations are performed, each of which generates a data panel, $y(\Theta)_i$. Let $M(\mathbf{Z})$ be a vector of moments generated from data $\mathbf{Z}$. The difference between the empirical moments observed and those generated in the simulated economies is defined as follows:

$$\ddot{M}(\Theta) \equiv M(\mathbf{X}) - \frac{1}{N}\sum_{i=1}^{N} M\left(y(\Theta)_i\right) \tag{10}$$

[11] For a more thorough discussion of SMM see McFadden (1989), Pakes and Pollard (1989), Hennessy and Whited (2005, 2007), Taylor (2010)

where $\mathbf{X}$ is the panel of empirical data. Then the SMM estimator satisfies the following:

$$\Theta^{\star} = \arg\min_{\Theta} Q\left(\mathbf{X}, \Theta\right) \equiv \ddot{M}\left(\Theta\right)' \hat{W} \ddot{M}\left(\Theta\right) \tag{11}$$

where $\hat{W}$ is the efficient weighting matrix, which is set to the inverse of the variance-covariance matrix of the moments in the empirical data.

Following Pakes and Pollard (1989), the parameter standard errors are computed as:

$$\left(1 + \frac{1}{N}\right) \left[\left(\frac{\partial \ddot{M}\left(\Theta^{\star}\right)}{\partial \Theta^{\star}} \right)' \hat{W} \left(\frac{\partial \ddot{M}\left(\Theta^{\star}\right)}{\partial \Theta^{\star}} \right) \right]^{-1} \tag{12}$$

where N is equal to the number of simulations, to adjust for a simulation bias. Conceptually, a parameter's standard error is reduced when moments are more sensitive to a change in the parameter, or when the moments are measured more precisely in the empirical data.

While computationally intensive, the minimization process is quite similar to GMM estimation. However, the parameter identification is dependent on the moments specified, and thus great care should be taken when selecting from a group of possible candidates. I now turn my attention to the moments used in the model estimation.

3.4 Identifying Moments

An ideal moment candidate has three distinct characteristics. A moment should be strongly correlated with one of the structural parameters. A second feature of the ideal moment is being either uncorrelated or correlated in the opposite direction with other structural parameters. Finally, the ideal moment is one that can be precisely measured in the empirical data. The motivation for this lies in the optimal weighting matrix, which depends on the relative precision with which each moment is estimated.

Therefore, an array of moments must be selected that is able to identify the distribution of CEO talent, the magnitude of shocks to productivity, and the switching cost associated with replacing the incumbent CEO. Ultimately, I use a broad spectrum of 15 moments related to CEO tenure length, firm performance around CEO changes, mean, persistence and variability of firm profitability, and changes in relative market values to identify the model parameters.

The first two identifying moments used are based on the length of an average CEO spell. The moments are set equal to the coefficients of α and β_1 in the following model:

$$d_{i,t} = \alpha + \beta_1 \lambda^{(7+)} + \nu_t + \varepsilon_{i,t} \tag{13}$$

where $d_{i,t}$ is an binary variable set to unity if CEO i is replaced, and zero otherwise, and $\lambda^{(7+)}$ is a dummy variable that takes on a value of one if the CEO's tenure at time t is greater than or equal to seven years of service, and zero otherwise. Year fixed effects are also included to control for any time trends in CEO turnover, which would only add noise to the estimation process. The coefficients on α and β_1 help to estimate the size of productivity shocks relative to the switching cost. For instance, if the cost to change managers is relatively high and productivity shocks are relatively low, a firm would be less likely to cross the switching thresholds early on in a CEO's tenure, leading to a larger value of β_1. However, if shocks are large relative to the switching cost, it would be more likely that a firm crosses the switching threshold within the first few years of a CEO's spell, leading to a larger value on the coefficient of α. When estimating (13) only those spells in which the CEO is replaced are kept. Neither switching costs nor productivity shocks play a role in the decision to retire in my model, thus all spells where the manager retires are removed from the sample. The frequency of managerial retirement will be estimated in a more traditional way, which will be discussed below.

While these two moments help identify the magnitude of shocks relative to the switching cost, they do not identify productivity shocks independent of other parameters. Therefore, the variance of within-spell firm profitability is also considered. Specifically, the third moment equals the expected variance of the residual estimated from the following pooled linear regression:

$$ROA_{i,t} = \eta_i + \nu_t + \varepsilon_{i,t} \tag{14}$$

Because the moment is intended to identify the magnitude of a firm's productivity shocks, firm-CEO fixed effects are included to absorb any variance in the profitability measure attributable to differences in CEO ability. Additionally, time fixed effects are included to control for the effects of economy wide shocks. This moment will also help to identify $\theta_{max},$. If assets are generally more productive, θ_{max}, will be larger leading to a larger variance of ROA for a given series of shocks to firm productivity. To further identify $\theta_{max},$, the average return on assets is estimated from the following pooled regression and serves as fourth final moment:

$$ROA_{i,t} = \alpha + \nu_t + \varepsilon_{i,t} \tag{15}$$

The next five moments chosen help identify shocks to firm productivity and are based on changes in a firm's relative value in the economy. Firms are first ranked according to market value each year, and the annual change in firm ranking is calculated for each firm. Five moments are then generated from the conditional variance of the residual in the following specification:

$$\Delta Rank_{i,t} = \boldsymbol{\lambda} + \boldsymbol{\lambda} \cdot \varepsilon_{i,t} \tag{16}$$

where $\boldsymbol{\lambda}$ is a vector of five dummy variables segmenting firms into quintiles based on lagged firm rank and $\Delta Rank_{i,t}$ is the change in a firm's ranking with respect to market value. Therefore, the first moment is set to the expected value of ε^2 conditional on having a lagged firm value in the first quintile, and so forth.

While the previous moments help to identify the switching cost relative to productivity shocks, the level of the cost is yet to be identified. For this, I turn to firm performance in the two years around a CEO turnover event. Specifically, four moments are estimated from the following regression:

$$ROA_{i,t} = \beta_1 \phi_0^U + \beta_2 \phi_{-1}^U + \beta_3 \phi_0^D + \beta_4 \phi_{-1}^D + \eta_i + \nu_t + \varepsilon_{i,t} \tag{17}$$

$ROA_{i,t}$ is set equal to the earnings after depreciation divided by the average of lagged and contemporaneous total assets of the firm at time t. Fixed effects are also included at the year and firm-CEO pair levels. The effects of a manager's turnover on profitability in the years surrounding the event are captured with the dummy variables taking on the generic form ϕ_t^s with the two possible superscripts U, and D which stand for up, and down respectively. Up represents the scenario in which the firm's relative value has increased over the term of the exiting CEO, represented by an increase in the firm's lagged size rank from the CEO's first year to their final year. Similarly, the subscript t, represents the number of years from the turnover event. Thus, in the year immediately preceding a CEO being terminated following a decrease in firm size, ϕ_{-1}^D would take on a value of one. By segregating the turnover event into two mutually exclusive groups, I allow for differing effects on firm profits around the turnover event. For instance, the cost of changing CEOs may be offset to a greater extent by a firm who has grown in size, relative to a firm that has shrunk. This would lead to a more negative coefficient of ϕ_{-1}^D when compared to that of ϕ_{-1}^U.

To identify the model parameters related to the distribution of CEO talent, the next moment is set to the elasticity of total CEO compensation to firm value:

$$ln\left(Pay_{i,t}\right) = \alpha + \beta_1 \cdot ln\left(Size_{i,t}\right) + \nu_t + \varepsilon_{i,t} \tag{18}$$

where $ln\left(Size_{i,t}\right)$ equals the market value of firm i at time t. Gabaix and Landier (2008) use this pay-size elasticity to estimate the tail index of CEO ability in a static model. Its inclusion allows for their estimation to be tested in a dynamic setting when firms and CEOs may not be matched in a positive assortative manner.

While other parameters have been discussed, the idiosyncratic component of profitability has yet to be tied to specific moments. For this, one final moment is considered. To identify what portion of a firm's profitability in a given period is due to noise and not to productivity shocks, I include the coefficient of lagged profitability from an AR(1) model:

$$ROA_{i,t} = \alpha + \beta_1 \cdot ROA_{i,t-1} + \varepsilon_{i,t} \tag{19}$$

3.5 Empirical Values of Moments

Each of the moments described above is defined within a reduced-form linear framework. Thus, ordinary least squares regressions provide consistent estimates for their values. The efficient weighting matrix used in the objective function is estimated by applying the seemingly unrelated regressions technique to the moment equations detailed above with error terms clustered on the firm level, and taking the inverse of the corresponding covariance matrix. This weighting matrix puts more weight on matching moments that are precisely measured in the data.

Finally, the probability of a CEO retiring is estimated using a parametric hazard model. Because the model assumes the CEO has an equal likelihood of retiring each year, the underlying hazard function in the estimation process is defined to have an exponential form. The entire history of each CEO included in the sample is first collected. All CEOs who are 62 years or older when they step down and do not take another executive position are assumed to have retired. Furthermore, if a CEO was forced out at a younger age and has since aged to the point where they have effectively left the labor pool, we must consider them to be retired. Therefore, in these cases where the CEO was replaced at a younger age but turns 62 before 2011, the last year observable, they are assumed to have retired when they turned

62. Finally, for all managers who are either still with their firm or have not turned 62 by 2011 the observations are considered to be right censored.

The estimation results for each of the sample moments along with the likelihood of retiring are reported in Table II. The first column estimates the percent of managerial turnover annually. Over the first six years of a manager's tenure, the probability of turnover occurring in a given year conditional on surviving up to that point is 4.2%. This probability increases by 2.3% conditional on the spell being at least 7 years in length. The second column reports the moments based on the cross-section of firm profitability. A firm's ROA has a cross-sectional mean of 10.31% and exhibits a persistent behavior through time. The third column indicates pay-size elasticity of CEO compensation is 0.42 for the sample considered.

In the fourth column, there is evidence for a change in firm profitability around a CEO's departure. Interestingly, while firm profits are lower in the year preceding a turnover event whenever the firm has decreased in size, although not significant at any traditional confidence level, the same is not true if the firm has increased in size. In addition, while profits are lower in both of these groups when managerial turnover occurs, the point estimate is more negative for firms whose market capitalization has decreased over the CEO's tenure. In untabulated results, CEO retirement has no statistically significant effects on firm profits in the years surrounding the retirement date. For this reason, I set the switching cost following a CEO's retirement to zero to reduce the number of free parameters that must be estimated.

The fifth column reports the variance of the change in firm rank with respect to market value conditional on lagged firm rank. Firms with a larger lagged market value exhibit a lower variance in their ranking within the economy. Finally, the final column reports the average probability of a CEO retiring in a given year. Thus, when solving the model, the probability of a manager retiring (δ) is set equal to 0.0574.

4 Estimation Results

The model is estimated using a simulated annealing algorithm, with the methodology used to compute the simulated moments detailed in Appendix IV. The resulting parameter estimations are presented in Panel A of Table III, along with their respective standard errors. The first parameter of particular interest is the cost that a firm incurs when replacing its CEO. The estimate indicates that if a firm were to change managers, they would effectively be paying a penalty equal to 2.18% of their assets in place. Relative to the median firm's annual return on assets, this represents a decrease of 20.0% which corresponds to a dollar cost of $81.8 million in the sample.[12] Using detailed information on CEOs in Denmark between 1992 and 2003, Bennedsen, Perez-Gonzales and Wolfenzon (2012) find a CEO's death is associated with an 18% decline in operating return on assets. However, this finding is slightly larger than the estimate of Taylor (2010) who evaluates the cost to be 1.33% of a firm's assets. When considering the decreased productivity a firm possibly faces when changing management, as well as the severance packages paid to executives among other costs, this cost is reasonable. While this estimate gives a sense of the contemporaneous cost a firm faces, it does not represent the overall impact on the economy. To examine the distortionary effect of this cost, in the next section I will generate a counterfactual economy free of this friction where talent can be optimally allocated.

Beyond this switching cost, the distribution of talent among the top managers is also estimated. The parameters indicate there is considerable dispersion in managerial ability. When benchmarked against the average level of talent among all managers, 0.17, the most talented manager with ability, 0.30, can generate gross revenues that are larger by a factor of approximately 1.77. While these parameters give a sense of the disparity in talent at the extreme, it does not characterize how quickly talent diminishes at the tail of the distribution.

[12]With respect to the mean return on assets, this represents 23.2% of annual ROA, or in dollar terms, a $94.2 million loss.

This rate of decrease is captured by the parameter α. The point estimate of 0.72 is very similar to the value of 0.66 that Gabaix and Landier (2008) estimate when using only executive compensation. However, it is hard to assign an economic value to the surplus managers provide based strictly on these estimates. In the next section, I will address this question more thoroughly using counterfactuals.

The estimate of the parameter γ indicates that the annual shock to a firm's productive assets is bounded by a 30.8% increase or decrease. Combining this with the uniform nature of the shock distribution, this parameter represents a standard deviation in a firm's productive assets of 17.78%.[13] In addition to this persistent shock to asset productivity, firm profitability is affected by a non-persistent, idiosyncratic component with a standard deviation of 0.0263.

While every parameter estimated is very statistically significant, this is not a sufficient indicator of their ability to replicate all the moments observed in the data. This hypothesis can be formally assessed using a χ^2 test, similar to over-identifying test presented in Hansen (1982) used in GMM estimations:

$$\frac{S}{1+S} N \cdot Q\left(\mathbf{X}, \Theta^{\star}\right) \tag{20}$$

where Q is the objective function being minimized in (9), N is the number of observations in the sample, and S is the number of simulations performed. Under the null that all J moments generated from a model with K parameters are equal to their empirically observed analogs, (18) follows a χ^2 distribution with $J-K$ degrees of freedom. As Panel B of Table III reports, the null that the model can replicate all 15 moments is rejected at the .01% confidence level. This model is asked to explain many different features of the data including the time-series variance and persistence of firm profitability within CEO spells, the average amount of CEO turnover, changes in relative firm value, and cross-sectional patterns in CEO compensation;

[13] Using the formula for standard deviation of a uniform distribution: $\frac{b-a}{\sqrt{12}} = \frac{2\cdot 30.8\%}{\sqrt{12}} = 17.78\%$

therefore the result is not wholly surprising. However, before discussing the model's economic implications, I examine where the model breaks down in its ability to replicate particular moments. Panel C reports the 15 moments estimated empirically and in the simulated data and a t-test of their differences.

CEOs in the simulated economy have an increased probability of experiencing turnover after their sixth year, similar to the pattern observed empirically. However, the model tends to over-estimate the turnover in the first six years, and slightly under-estimate the likelihood of turnover in later years. There is a statistically significant difference in the turnover rate in the first six years of a CEO's tenure but not in terms greater than six years.

When attempting to match moments related to firm profitability, the model fares considerably better. Both the average and persistence of a firm's return on assets is similar in the model relative to their empirical counterparts. However, the model tends to underestimate the with-spell variance of firm profitability.

Interestingly, the equilibrium wages the model yields results in an elasticity of CEO pay to firm value within 0.05 of the estimated elasticity from Execucomp. The difference between the two values cannot be rejected at the 5% confidence level.

The results are mixed when the model attempts to mimic the array of moments related to profitability around CEO turnover events. While only one moment out of six is rejected at the .1% confidence level, the model does fail to replicate other patterns among these moments. Empirically, the decrease in profitability tends to be larger for years when turnover occurs compared to the year prior. While a similar pattern is generated in the model, the simulations tend to under-estimate the cost following an increase in firm value and over-estimate the cost following a decrease in firm value.

Finally, when evaluated on its performance in matching changes to a firm's relative ranking within the economy, the model tends to under-state the variability of these changes in value. The one exception is the fourth quintile corresponding to relatively small firms in the economy. The model tends to over-state the change in firm ranking for firm's whose lagged ranking qualifies them for this quintile.

5 Distortionary Effects of a Switching Cost

While the estimated parameters values presented in the previous section give some insight into the cost of changing managers and the degree to which a better manager adds value, their economic magnitudes are still unclear. To better understand how this friction affects the matching of firms and managers, I perform a counterfactual that measures the distortionary effects that a switching cost has on the optimal allocation of talent across firms.

As discussed in the previous section, the estimation results imply that whenever a firm replaces its manager it incurs a one-time cost of 2.18% of total firm assets. While interesting, this does not represent the complete cost this inefficiency imposes on the economy. By multiplying this estimated switching cost by the likelihood of replacing a manager in a given year and evaluating it as an annuity of annual costs one can obtain a back-of-the-envelope estimate, which serves as a lower bound on the decrease in firm value. However, this understates the extent of the distortionary effect as it does not incorporate the implicit cost of being inefficiently matched with a suboptimal manager following a shock to productivity. For this reason, I compare the expected value of firms in my model to those in a counterfactual economy absent any switching costs.

The optimal policy function associated with the parameter estimates from the previous section serves as a base case, denoted as *Matching Friction*. Alternatively, a second economy is considered that is free of any switching costs, denoted as *Optimal Matching*. In this case, the firm's optimal decision is to always enter the labor market and be re-matched with a CEO.[14] As an estimate of total firm value in each scenario, I start the economy and allow

[14]If a firm has not experienced a change in rankings since the previous period, they will be re-matched with the CEO that had in the previous period.

150 periods to pass to establish a steady-state. At this point I assign a size rank to each firm. 75 periods are then simulated, wherein firms' assets experience periodic shocks, managerial replacement occurs according to the corresponding policy function, and profits are realized in each period. For each firm, this series of 75 cash flow realizations is then discounted back to the period when firms were initially ranked. Therefore, the discounted value of the cash flow series represents an estimate of total firm value. This simulation is repeated 100,000 times and the present value calculations are averaged across all simulations to form an expected firm value for each size ranking.

Panel A of Figure 5 plots the ratio of a firm's value under the *Optimal Matching* scenario relative to the *Matching Friction* case. There is little differentiation in expected firm value for larger firms, while the difference increases as firm size becomes smaller. The distortion to firm value caused by the switching cost reaches its greatest point slightly after the median firm. Recall that the moment estimates from the previous section indicate that there is less volatility in the size rankings of large firms relative to small firms, both in the model and empirically. Thus, productivity shocks are less likely to affect the relative ranking of a large firm in the economy. This reduces the difference in ability between their incumbent CEO and a replacement they could obtain in the labor market. Ultimately, firms in the frictionless economy who are able to replace their CEOs free of a switching cost have a median (mean) value 5.06% (4.21%) larger than those who pay a switching cost and are suboptimally matched with managers.

While the previous two scenarios capture the overall difference in firm value caused by the matching friction, they do not quantify what portion of this distortion is the inefficiency of the manager-firm match and what portion is the cost experienced when replacing managers. To address this question, a third scenario is considered in which each firm experiences a series of switching costs identical to the *Matching Friction* scenario, but is also efficiently matched with managers in every period. Thus, the difference in a firm's value under this scenario, denoted *Switching Cost*, and *Matching Friction* represents the economic magnitude of the inefficient match. The reduction in value solely attributable to the inefficient matching of firms and managers exhibits a similar pattern to the overall decrease in firm value. At its greatest point, optimal matching results in an increase in firm value of 2.73%. For the overall economy, the median (mean) firm value increases by 1.84% (1.72%) when managers and firms are optimally matched but still experience the same switching cost.

Panel B of Figure 5 contrasts the cost of suboptimal matching to the overall decrease in firm value. For each firm, the ratio of the decrease in firm value associated with inefficient matching to the decrease in value associated with both inefficient matching and the explicit switching cost is reported. While the economy's largest firms experience productivity shocks that lead to inefficient matches, these shocks seldom result in the optimal decision to replace managers. Thus, the implicit cost associated with an inefficient match dominates the explicit cost of switching managers for these firms. While the contribution of this implicit cost to the overall destruction in firm value decreases when considering smaller firms, at its minimum point it still represents two-thirds of the overall effect. Overall, suboptimal matching represents 76.2% of the overall reduction in firm value for the median firm.

While the existence of a switching cost does lead to the inefficient matching between firms and managers, it is still unclear how equilibrium wages change in a dynamic setting relative to prior static model. Contrasting the wages from the model presented here with a static framework will give insight into how sensitive the predictions of Gabaix and Landier (2008) and Terviö (2008) are to a multi-period setting. Therefore, I take the estimated model parameters and construct the one-period equilibrium wages from (3).

Figure 6 plots the ratio of equilibrium wages from the dynamic model to these static wages. Overall, equilibrium wages are larger in the dynamic framework relative to a one-period model. This disparity is largest for the lowest ranked firms in the economy.[15] However,

[15] This could be attributed to the technical requirement a firm's productive assets be bounded for a

the average increase in wages in the dynamic model is only 3.10%. Therefore, it is unlikely that the predictions from a static model would be materially influenced by the construction of wages from a dynamic framework.

6 Conclusion

Recent work by both Gabaix and Landier (2008) and Terviö (2008) evaluate the importance of CEOs based on the compensation received in a static equilibrium. However, firms and industries continually receive shocks to the productivity of their assets. In a frictionless economy, as a firm grows in prominence they should be optimally matched with a more talented manager. However, while firms grow and shrink in size, empirically they do not replace existing managers in a manner consistent with a frictionless labor market. If there is a friction in the matching process, how do predictions from a static model change? Furthermore, how much firm value is destroyed due to this friction?

To answer these questions, this paper proposes a dynamic competitive equilibrium model with time-varying firm productivity and a switching cost that must be paid when replacing an existing CEO. The model is estimated using SMM, yielding an interesting set of findings. Firms experience a switching cost equal to 2.18% of their assets in place. Ultimately, this friction prevents the optimal re-matching of CEOs and firms in a dynamic setting. Using counterfactual analysis, it is estimated that the median (mean) firm's value would increase by 5.06% (4.21%) when able to replace managers free of any such switching cost. Furthermore, the lion's share of this value destruction is due to the inefficiency of firm-CEO matches in the cross-section and not the explicit cost when switching managers.

While the model estimated here is relatively straightforward as firms are able to observe the ability of managers, the overall cost they experience due to factors such as lost produc-

stationary distribution to be reached. Therefore, small firms would be more likely to grow, increasing the premium for managers in this region of the distribution.

tivity, severance pay, and search costs negatively impacts firm value in a substantial way. Furthermore, reduced-form estimates of this cost will understate the overall impact to firm value, as they do not account for the cost of a suboptimal match between firms and CEOs.

Appendix I

Solution to the Dynamic Programming Problem

This appendix solves the dynamic programming problem faced by the firm. I begin by introducing some notation. Let f_s be the probability distribution of firms participating in the labor market with respect to their level of productive assets. Let f_θ be the probability distribution of talent for managers participating in the labor market. Define A_t, θ_t, and $W(\theta_t)$ to be the productive assets of a firm, the talent level of its incumbent manager, and its wages paid at time *t*. Finally, let $\theta(A_t) \equiv F_\theta^{-1}(F_s(A_t))$, indicating the talent level of the manager a firm with productive asset A_t would be matched with in the labor market.

Now, I re-examine the value maximization problem faced by a firm. Each firm maximizes expected value through its firing decision:

$$V(\theta_{t-1}, A_t) \equiv \max_{\{d_s\}_{s=t}^{\infty}} E_t\left[\sum_{s=t}^{\infty} \beta^{s-t}\pi_s\right] \tag{AI.1}$$

where θ_{t-1} is the talent level of the incumbent manager at the close of the previous period. Furthermore, because all agents are risk-neutral and the idiosyncratic component of profitability in equation (4) has mean zero, the solution to (AI.1) can be found by excluding this term altogether. Therefore, by substituting (4) into (AI.1) and separating out the current period's dividend, I get the following:

$$V(\theta_{t-1}, A_t) \equiv \max_{\{d_s\}_{s=t}^{\infty}} A_t \cdot \left(\theta_t - d_t \cdot c_{replace}\right) - W(\theta_t) + E_t\left[\sum_{s=t+1}^{\infty} \beta^{s-t}\pi_s\right] \tag{AI.2}$$

$$\equiv \max_{d_t}\left\{\begin{array}{c}\left(A_t\theta_t - W(\theta_t) + \max\limits_{\{d_s\}_{s=t+1}^{\infty}} E_t\left[\sum_{s=t+1}^{\infty}\beta^{s-t}\pi_s\right]\Big|\theta_t = \theta_{t-1}\right), \\ \left(\left(A_t \cdot \left(\theta_t - c_{replace}\right) - W(\theta_t) + \max\limits_{\{d_s\}_{s=t+1}^{\infty}} E_t\left[\sum_{s=t+1}^{\infty}\beta^{s-t}\pi_s\right]\Big|\theta_t = \theta(A_t)\right)\right)\end{array}\right\} \tag{AI.3}$$

I now focus on the first value in (AI.3), which corresponds to the case when the firm retains their current CEO. While the firm has chosen to retain the incumbent manager, she will retire with probability δ at the end of the period after the profits have been realized. Therefore, the value of the firm conditional on retaining the existing manager is:

$$V(\theta_{t-1}, A_t)^{retain} = A_t\theta_{t-1} - W(\theta_{t-1})$$
$$+ (1-\delta) \cdot \left(\max_{\{d_s\}_{s=t+1}^{\infty}} E_t[\textstyle\sum_{s=t+1}^{\infty} \beta^{s-t}\pi_s] \middle| \theta_t = \theta_{t-1} \right)$$
$$+ \delta \cdot \left(\max_{\{d_s\}_{s=t+1}^{\infty}} E_t[\textstyle\sum_{s=t+1}^{\infty} [\beta^{s-t}\pi_s] - \beta \cdot c_{retire} \cdot A_{t+1}] \middle| \theta_{t+1} = \theta(A_{t+1}) \right) \quad \text{(AI.4)}$$

By factoring out a β term from both summations, and given the distribution of size shocks, f_x, the switching cost can be separated out by conditioning on the firm's current level of assets, yielding:

$$V(\theta_{t-1}, A_t)^{retain} = A_t\theta_{t-1} - W(\theta_{t-1})$$
$$+ \beta(1-\delta) \cdot \int \left(\max_{\{d_s\}_{s=t+1}^{\infty}} E_t[\textstyle\sum_{s=t+1}^{\infty} \beta^{s-t-1}\pi_s] \middle| \theta_t = \theta_{t-1},\, A_{t+1} = A_t x \right) f_x \, dx$$
$$+ \beta\delta \cdot \int \left(\max_{\{d_s\}_{s=t+1}^{\infty}} E_t[\textstyle\sum_{s=t+1}^{\infty} \beta^{s-t-1}\pi_s] \middle| \theta_{t+1} = \theta(A_{t+1}),\, A_{t+1} = A_t x \right) f_x \, dx \quad \text{(AI.5)}$$
$$- \beta\delta \cdot \int c_{retire} \cdot (A_t x) \cdot f_x \, dx$$

Note that first integrand is represents (AI.1) one period into the future. Furthermore, the second integrand is of the same form with one exception. The firm no longer has any choice over retaining or replacing the incumbent CEO in period *t+1* following their retirement at the end of time *t*. However, a manager with talent level $\theta(A_{t+1})$ at the end of time *t* is, by definition, the optimal match at time *t+1*, thereby satisfying equation (AI.1). Therefore, (AI.5) can be simplified to the following:

$$V(\theta_{t-1}, A_t)^{retain} = A_t\theta_{t-1} - W(\theta_{t-1}) + \beta(1-\delta) \cdot \int V(\theta_{t-1},\, A_t x) f_x \, dx$$
$$+ \beta\delta \cdot \left[\int V(\theta(A_t x),\, A_t x) f_x \, dx - \int c_{retire} \cdot (A_t x) f_x \, dx \right] \quad \text{(AI.6)}$$

$$= A_t\theta_{t-1} - W(\theta_{t-1}) + \beta \cdot \left[(1-\delta) \cdot E_t[V(\theta_{t-1}, A_{t+1})] + \delta \cdot E_t[V(\theta(A_{t+1}), A_{t+1}) - c_{retire} \cdot A_{t+1}] \right] \quad \text{(AI.7)}$$

$$= A_t\theta_{t-1} - W(\theta_{t-1}) + \beta \cdot E_t\left[(1-\delta) \cdot V(\theta_{t-1}, A_{t+1}) + \delta \cdot [V(\theta(A_{t+1}), A_{t+1}) - c_{retire} \cdot A_{t+1}] \right] \quad \text{(AI.8)}$$

Intuitively, (AI.8) can be broken down into three parts. The first term represents the dividends earned in the current period if the manager is retained. The second component represents the discounted expected value of the firm next period conditional on the manager not retiring. The last element is the discounted expected value of the firm if the current manager does retire.

Alternatively, the firm's other option is to replace the current manager. Fortunately, this case can be represented much more concisely, while still remaining intuitive. Recall that if the firm replaces its current manager, its value can be represented by the following expression:

$$V(\theta_{t-1}, A_t)^{replace} = A_t\theta(A_t) - W\big(\theta(A_t)\big) + \left(\max_{\{d_s\}_{s=t+1}^{\infty}} E_t\left[\textstyle\sum_{s=t+1}^{\infty} \beta^{s-t}\pi_s\right]\middle|\theta_t = \theta(A_t)\right) - A_t \cdot c_{replace} \qquad \text{(AI.9)}$$

Note that this value is equivalent to equation (AI.1) when $\theta_{t-1} = \theta(A_t)$ with the exception of the switching cost suffered by the firm. Furthermore, $\theta_{t-1} = \theta(A_t)$ implies that the CEO of the firm at the close of *t-1* is the optimal manager of the firm at time *t* given the assets of the firm at time *t*. Thus, the value of the firm when replacing the current CEO is:

$$\begin{aligned} V(\theta_{t-1}, A_t)^{replace} &= A_t\theta(A_t) - W\big(\theta(A_t)\big) + \beta \cdot E_t\big[(1-\delta) \cdot V(\theta(A_t), A_{t+1}) \\ &+ \delta \cdot [V(\theta(A_{t+1}), A_{t+1}) - c_{retire} \cdot A_{t+1}]\big] - A_t \cdot c_{replace} \end{aligned} \qquad \text{(AI.10)}$$

$$V(\theta_{t-1}, A_t)^{replace} = V(\theta(A_t), A_t) - A_t \cdot c_{replace} \qquad \text{(AI.11)}$$

Thus, by substituting equations (AI.8) and (AI.11) into the original maximization problem that the firm faces the following Bellman equation is achieved:

$$V(\theta', A_t) \equiv \max_{d_t} \begin{cases} A_t\theta' - W(\theta') + \beta \cdot E_t\big[(1-\delta) \cdot V(\theta', A_{t+1}) + \delta \cdot [V(\theta(A_{t+1}), A_{t+1}) - c_{retire} \cdot A_{t+1}]\big] \\ V(\theta(A_t), A_t) - A_t \cdot c_{replace} \end{cases} \qquad \text{(AI.12)}$$

Appendix II

Equilibrium Distributions of Firms and Managers in the Labor Market

iv. Distribution of Competitive Firm Sizes

In equilibrium, the stationary distribution of firms competing in the labor market is a function of the optimal policy function of the firm. While asset productivity of a firm has been treated as a continuous random variable up to this point, the steady state distribution is more easily explained in a discrete context. Therefore, let a firm's productive assets, A, take on values from the discrete set $\{A_0, A_1, \dots, A_N\}$. Then, given the optimal policy function that solves the dynamic programming problem, $(\underline{A}, \overline{A})$, a firm will not replace its manager so long as its productive assets satisfy the following condition: $A \in [\underline{A}, \overline{A}]$.

The stationary distribution of firms entering the labor market requires the probability distribution of a firm's productive assets upon entering the labor pool, conditional on having assets A_i when the incumbent manager was hired. To do this, I must separate a firm's dynamics into two parts, 1) the changing of its assets due to productivity shocks and 2) the need to enter the labor market, either by choice or due to managerial retirement. Therefore, two Markov chains will be constructed, each with dimensions (N+1) *x* (N+1). The first of these, Π, represents the probability of moving from A' to A in a single period. This Markov chain also features an additional state, indexed by *N+1*, that serves as an absorbing state for firms who have already entered the labor market. Let $p_{i,j} \equiv P(A = A_j | A' = A_i)$, then:

$$\Pi = \begin{bmatrix} p_{0,0} & \cdots & p_{0,N} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ p_{N,0} & \cdots & p_{N,N} & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \tag{AII.1}$$

The second Markov chain, Γ, represents the likelihood of a firm seeking out a new manager given their current level of assets, A_j. This probability is composed of two factors. First, every firm is at risk of their manager retiring at the end of the previous period. Secondly, if A_j is outside the switching threshold, $(\underline{A}, \overline{A})$, the firm will enter the labor market with certainty. The last column of this matrix represents the firm's likelihood of changing managers, while the diagonal values represent the firm's likelihood of retaining their current manager. All other elements have a value of zero, as a firm will either retain their current manager and stay in their current state or replace their manager thereby entering the absorbing state:

$$\Gamma = \begin{array}{c} 0 \\ \vdots \\ \underline{A}-1 \\ \underline{A} \\ \vdots \\ \overline{A} \\ \overline{A}+1 \\ \vdots \\ \mathrm{n}+1 \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & \dots & \dots & 1-\delta & \dots & \dots & \vdots & \vdots & \delta \\ 0 & \dots & \dots & 0 & 1-\delta & 0 & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & 1-\delta & \vdots & \vdots & \delta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{AII.2}$$

Therefore, the probability of a firm transitioning from $\mathrm{A_i}$ to $\mathrm{A_j}$ after *t* periods can then be found in the *ith* row and *jth* column of the following matrix:

$$(\Pi\Gamma)^{(t)} \equiv \prod_{i=1}^{t} (\Pi \cdot \Gamma) \tag{AII.3}$$

The probability that a firm enters the labor market with assets $\mathrm{A_j}$ conditional on having initial assets $\mathrm{A_i}$ is found in (AII.4), below. Intuitively, the ith row of $(\Pi\Gamma)^{(t)}$ represents the distribution of a firm's assets after *t* periods without entering the labor market, conditional on starting with assets $\mathrm{A_i}$. Multiplying this value by Π and taking the element in the *i*th row and *j*th column yields the probability of transitioning to $\mathrm{A_j}$ in one additional period. This value is then multiplied by the probability of entering the labor market given assets $\mathrm{A_j}$, which is equal to 1 if the $\mathrm{A_j}$ is beyond the replacement threshold or δ if $\mathrm{A_j}$ is within the optimal switching threshold but the current CEO retires. Summing over all the possible number of periods before entering the labor market gives the final probability:

$$P_{i,j} \equiv \begin{cases} \sum_{t=1}^{\infty} \left[(\Pi\Gamma)^{(t)} \cdot \Pi\right]_{i,j} \cdot \delta \ \ \text{if } \mathrm{A_j} \in [\underline{A}, \overline{A}] \\ \sum_{t=1}^{\infty} \left[(\Pi\Gamma)^{(t)} \cdot \Pi\right]_{i,j} \cdot 1 \ \ \text{if } \mathrm{A_j} \notin [\underline{A}, \overline{A}] \end{cases} \tag{AII.4}$$

Finally, given $P_{i,j}$ for all possible values of *i* and *j*, the stationary distribution of firms upon entering the labor market, represented by the column vector Π of dimension 1 *x* N, satisfies the following condition:

$$\Pi = \Pi \cdot \begin{bmatrix} P_{0,0} & \cdots & P_{N,0} \\ \vdots & \ddots & \vdots \\ P_{0,N} & \cdots & P_{N,N} \end{bmatrix} \tag{AII.5}$$

v. Distribution of Labor Pool Talent

Given the distribution of firms entering the labor market, it is relatively straightforward to find the distribution of talent levels for managers in the labor market. This pool has two contributing sources, newly born CEOs in the economy and CEOs who have been released from their current firm. These two groups must be characterized separately and then aggregated into one distribution.

In the framework of the model, the distribution of newly born CEOs is assumed to be exogenous. Furthermore, in equilibrium the measure of these CEOs being born is equal to the percentage that retired in the same period, similar to an overlapping generation framework. Using the extreme value theorem the ability of a manager in the xth percentile is defined as:

$$\theta(x) = \theta_{max} - \frac{B}{\propto} \cdot x^{\propto} \tag{AII.6}$$

Alternatively, the distribution of managers participating in the labor pool after previously running a firm is closely related to the distribution of firm's entering the labor pool. I first discretize the distribution of managers. For each of the firms in the discrete set $\{A_0, A_1, \ldots, A_N\}$, assign a manager according to (AII.7):

$$\theta_i = \theta(A_i) = F_\theta^{-1}(F_s(A_i)) \tag{AII.7}$$

Next, consider a firm with an initial assets A_i. The firm's choice to enter the labor market will be made due to managerial replacement or retirement. Therefore, I first factor out the probability of retirement. If the firm's assets upon entering the labor market are still within the switching threshold of the firm, the CEO retired with certainty. Alternatively, if the firm is outside the switching threshold, the probability that their CEO retired is equal to δ. Therefore, by leveraging (AII.4), the probability of a firm with initial assets A_i releasing their current manager of talent level θ_i back into the labor market is:

$$Prob(\theta_i \text{ re-entering} | A_i) = \left(1 - \sum_{j | A_j \in [\underline{A}, \overline{A}]} P_{i,j}\right) \cdot (1 - \delta) \tag{AII.8}$$

Finally, the distribution of CEOs that re-enter the labor market is proportional to (AII.8), which is conditional on having initial assets A_i, multiplied by the probability of having initial assets A_i. By referencing the stationary distribution of firm sizes in the labor market found in equation (AII.5), the distribution of talent re-entering the pool is proportional to:

$$f_\theta(\theta_i) = \Pi_i \cdot \left(1 - \sum_{j|A_j \in [\underline{A}, \overline{A}]} P_{i,j}\right) \cdot (1 - \delta) \qquad \text{(AII.9)}$$

Therefore, aggregating (AII.6) and (AII.9) yields the overall talent distribution of managers found in the labor market.

Appendix III

Solution to the General Equilibrium

i. General Equilibrium Solution

The general equilibrium solution is found by first assuming a distribution of firms and managers participating in the labor market as well as a wage schedule, and solving the dynamic programming problem faced by a firm given these distributions. Wages are then recalculated given the value function, and the dynamic programming problem is solved for again. After finding the fixed point of this problem, the optimal policy function generated by the solution is used to generate a new pair of distributions for firms and managers participating in the labor market. Given these new distributions, the solution to the maximization problem is calculated again. This process is repeated until the current policy function is identical to that of the previous iteration, at which point the general equilibrium has been found.

I rely on numerical techniques to solve the dynamic programming problem and the associated firm size and talent distributions. Firms are assumed to take on discrete levels of productive asset whose logged values are evenly spaced over a grid of 1,200 points. The assets at the upper and lower endpoints are chosen to match the assets of the largest and 500th firm, respectively, in our sample as of 1993. While these assets remain fixed from iteration to iteration, the associated probability mass function does not, and must be solved for. However, in order to solve the dynamic programming problem, a size distribution must first be assumed. Therefore, consistent with the findings of Gabaix and Landier (2008), the initial distribution is assumed to be governed by Zipf's law.[16] All atomistic firms that lie on a discrete grid point are assumed to have managers of equal talent levels. Therefore, given parameter values governing the distribution of manager ability, talent levels can be computed (AII.6) using the probability mass associated with each point. Finally, given these firm sizes and talent levels, an initial guess for the wages earned by each manager is generated according to (3).

Given parameter values, manager and firm distributions, and wages, (6) is solved for by value function iteration. The iterative process continues until the difference in values from one iteration to the next

[16] While the distribution of size is assumed to follow Zipf's law, the general equilibrium solution is not dependent on these starting values. However, while any distribution could be assumed, the speed of convergence does depend on the initial probability mass function. As I do not have a prior on the general form of the final distribution, Zipf's law seems as suitable as any for a starting guess.

converges to a sufficiently small amount.[17] However, the resulting values are implicitly a function of the initial series of wages specified, which do not necessarily satisfy (7). Competitive wages are such that a firm's value upon entering the labor market would be unchanged if it was matched with a manager one ranking better than its optimal match, thus making the firm indifferent between the two managers. To find this equilibrium, the firm values from the previous step are used to update the wage levels and firm values are re-computed. This process is repeated until the difference in wages from one iteration to the next is sufficiently small. At this point, the equilibrium wages have been found given a distribution of firms and managers participating in the labor market. This also yields a decision rule that each firm will follow when deicing to retain or replace its manager.

This optimally policy function can then be used in equations (AII.5) to update the probability mass function of the firms participating in the labor market. Finally, the distribution of managers in the pool can be updated by aggregating the resulting probability mass function of CEOs re-entering the pool from equation (AII.9) with the newly born individuals entering the pool for the first time.

After repeating this cycle until every firm's decision remains unchanged from one solution of the optimization problem to the next, the optimal policy function in a competitive equilibrium has been generated, given an initial parameter set.

[17] The value function is iterated over until the sum of the absolute changes from one iteration to the next is less than 0.01.

Appendix IV

Simulated Moment Generation

Given a set of model parameters, the general equilibrium is first computed from which the optimal policy function is derived. Next 500 firms are created, each one given an initial size according to Zipf's law and optimally matched with a CEO drawn from the distribution governed by the model parameters.[18] After creation, 150 years are allowed to pass so a steady state can be established. In each year, every firm's level of productive assets experiences a shock drawn from a uniform distribution with mean one and support dictated by the model parameters. Additionally, each firm's manager faces a chance of retirement with probability δ. If either a firm's CEO retires or the productivity shock it received moves it beyond the switching threshold, it enters the labor pool to seek a replacement. The pool of CEOs available to choose from is comprised from two sources. First, all the managers released by their respective firms re-enter the labor market. The second source consists of a group of newly born managers, whose total number equals the number of CEOs that retired in the previous period, is drawn from the distribution governed by the model's parameters. Each firm competing in the pool is assigned a new manager according to their relative ranking amongst all other competing firms, and wages are set accordingly. Each firm then generates revenues based on amount of assets at the start of the period and the current manager in place, and pays out all revenue net wages to the shareholders as a dividend.

After 150 years have passed, the entire time-series of profits for the current CEO spells are collected. The economy is then permitted to continue for another 18 years. Thus the final panel consists of 500 firms over a 19 year span, consistent with the data collected from Compustat and Execucomp.

Once the full panel has been simulated, moments are calculated from the generated data in exactly the same way as they were computed on the actual data. The economy is then reset, and is simulated again until 64 sets of moments have been generated. At this point, the objective function from (8) is computed. I seek to find the parameter set corresponding to the global minimum of the objective function, which may

[18] In addition to Zipf's law, which dictates the spacing between firms, the size of one firm must be specified to pin down the distribution. Therefore, the size of the 500th firm is set to roughly correspond to the size of the 500th firm in the sample in 1993.

also have many local minima. Therefore, standard convex optimization routines should be avoided. Ideally, the global minimum would be found using a grid search over the state space of feasible parameter values; however the computationally intensive nature of the problem makes this impractical. For these reasons, a simulated annealing routine is used to estimate the model, where the initial temperature is estimated from a set of 50 random parameter values. To avoid any unnecessary instability in the optimization process, for each set of parameters considered the same random seed is used when initializing the economy for the first time.

References

Adams, R., H. Almeida and D. Ferreira. 2005. Powerful CEOs and Their Impact on Corporate Performance. *Review of Financial Studies* 18, 1403–32.

Almazan, A. and J. Suarez. 2003. Entrenchment and Severance Pay in Optimal Governance Structures. *Journal of Finance* 58, 519–47.

Bennedsen, M., F. Perez-Gonzales and D. Wolfenzon. 2012. Estimating the Value of the Boss: Evidence from CEO Hospitalization Events. Working Paper, Stanford University.

Bertrand, M. and A. Schoar. 2003. Managing with style: The effects of managers on firm policies. *Quarterly Journal of Economics* 118, 1169–1208

Bond, S. and C. Meghir. 1994. Dynamic Investment Models and the Firm's Financial Policy. *Review of Economic Studies* 61, 197–222

Carter, M.E., F. Franco and I. Tuna. 2010. Premium Pay for Executive Talent: An Empirical Analysis. Working Paper, Boston College.

Coles, J., M. Lemmon and J. F. Meschke. 2010. Structural Models and Endogeneity in Corporate Finance: The Link Between Managerial Ownership and Corporate Performance. *Journal of Financial Economics* 103, 149–68.

Eisfeldt, A. and C. Kuhnen. 2013. CEO Turnover in a Competitive Assignment Framework. *Journal of Financial Economics* 109, 351–72.

Falito, A., D. Li and T.T. Milbourn. 2013. Which Skills Matter in the Market for CEOs? Evidence from Pay for CEO Credentials. Working Paper, University of Washington in St. Louis.

Gabaix, X. and A. Landier. 2008. Why Has CEO Pay Increased so Much? *Quarterly Journal of Economics* 123, 49–100.

Graham, J.R., S. Li and J. Qiu. 2012. Managerial Attributes and Executive Compensation. *Review of Financial Studies* 25, 144–86.

Hennessy, C. and T. Whited. 2005. Debt Dynamics. *Journal of Finance* 60, 1129–65.

———. 2007. How Costly is External Financing? Evidence from a Structural Estimation. *Journal of Finance* 62, 1705–45.

Huson, M., Parrino, R. and L. Starks. 2001. How Internal Monitoring Mechanisms and CEO Turnover: A Long-Term Perspective. *Journal of Finance* 56, 2265–97.

Kang, Q., Q. Liu and R. Qi. 2010. The Sarbanes-Oxley Act and Corporate Investment: A Structural Assessment. *Journal of Financial Economics* 96, 291–305.

Kaplan, S.N., M.M. Klebanov and M. Sorensen. 2012. Which CEO Characteristics and Abilities Matter? *Journal of Finance* 67, 973–1007.

McFadden, D. 1989. A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica* 57, 995–1026.

Murphy, K. and J. Zabojnik. 2004. CEO Pay and Appointments: A Market-Based Explanation for Recent Trends. *The American Economic Review* 94, 192–6.

———. 2007. Managerial Capital and the Market for CEOs. University of Southern California, Working Paper.

Pakes, A. and D. Pollard. 1989. Simulation and the Asymptotics of Optimization Estimators. *Econometrica* 57, 1027–57.

Parrino, R. 1997. CEO Turnover and Outside Succession: A Cross-Sectional Analysis. *Journal of Financial Economics* 46, 165–97.

Roberts, M. and J. Tybout. 1997. The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs. *The American Economic Review* 87, 545–64.

Rosen, S. 1981. The Economics of Superstars. *The American Economic Review* 71, 845–58.

Sattinger, M. 1993. Assignment Models of the Distribution of Earnings. *Journal of Economic Literature* 31, 831–80.

Terviö, M. 2008. The Difference that CEOs Make: An Assignment Model Approach. *American Economic Review* 98, 642–68.

Taylor, L. 2010. Why are CEOs Rarely Fired? Evidence from Structural Estimation. *Journal of Finance* 65, 2051–87.

———. 2013. CEO Wage Dynamics: Estimates from a Learning Model. *Journal of Financial Economics* 108, 79–98.

Yermack, D. 2006. Golden Handshakes: Separation Pay for Retired and Dismissed CEOs. *Journal of Accounting and Economics* 41, 237–56.

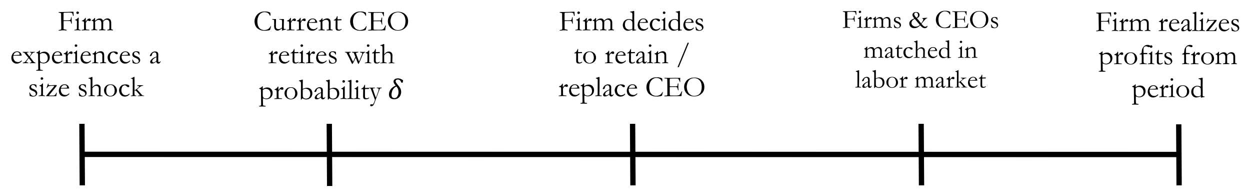


Figure 1
Timing of Model
This figure depicts the timeline of events within each period of the dynamic model.

Panel A

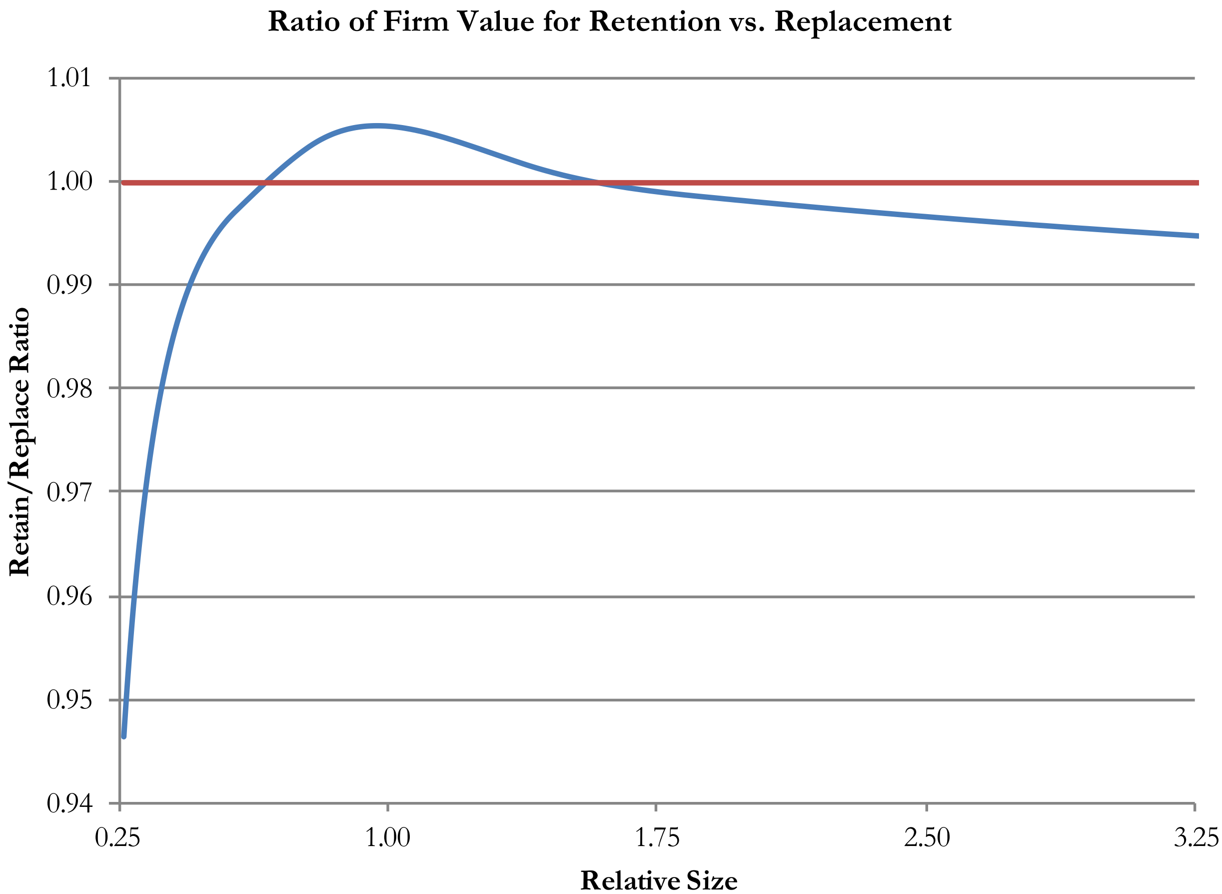
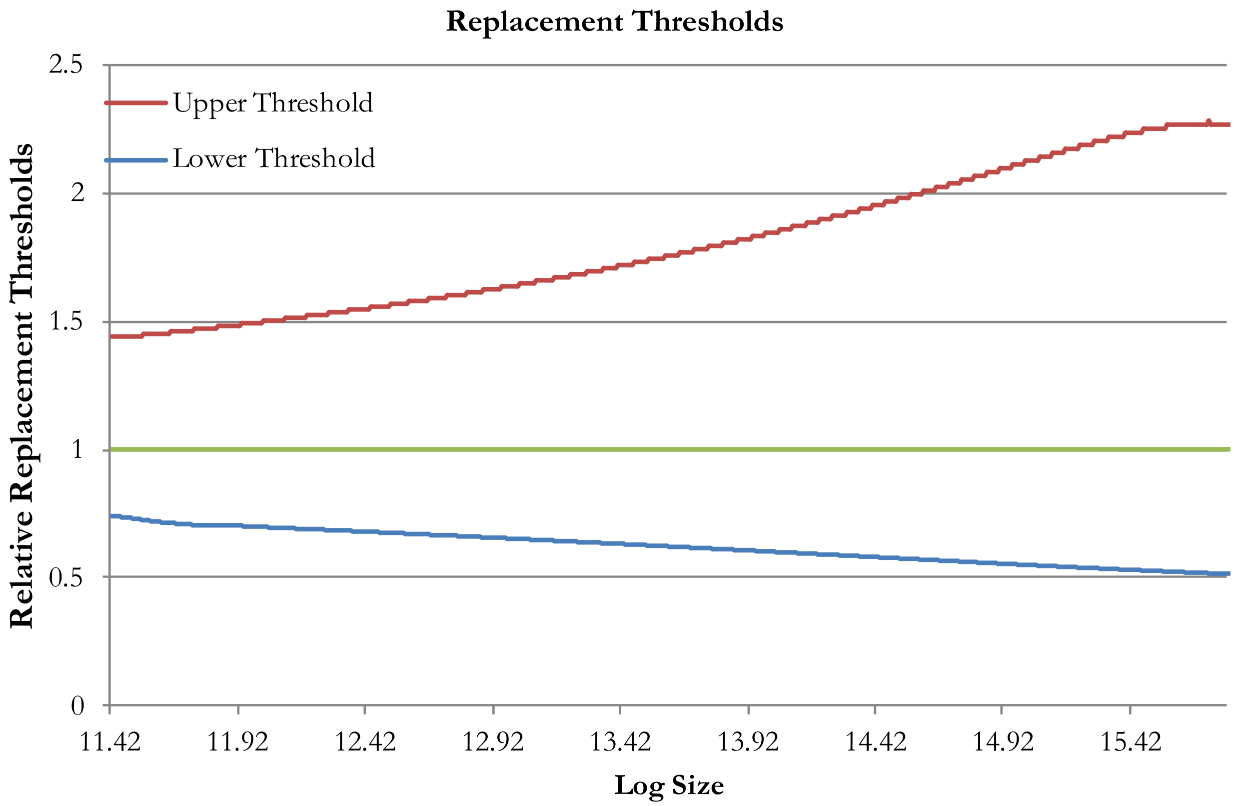


Figure 2

Value of Retaining versus Replacing Managers

This figure illustrates a firm's expected value given the retention or replacement of the incumbent CEO. The y-axis represents the expected firm value when retaining the existing manager to the expected firm value when replacing the incumbent manager. The x-axis represents the ratio of the firm's current level of productive assets to level of productive assets when the incumbent manager was hired. The two points at which the blue line crosses 1.00 represent the switching thresholds where the firm finds it optimal to replace the existing manager. Panel B illustrates how these optimal switching thresholds vary with the level of productive assets at which point the incumbent CEO was hired.

Panel B


Replacement Thresholds
Upper Threshold
Lower Threshold
Relative Replacement Thresholds
Log Size
2.5
2
1.5
1
0.5
0
11.42
11.92
12.42
12.92
13.42
13.92
14.42
14.92
15.42

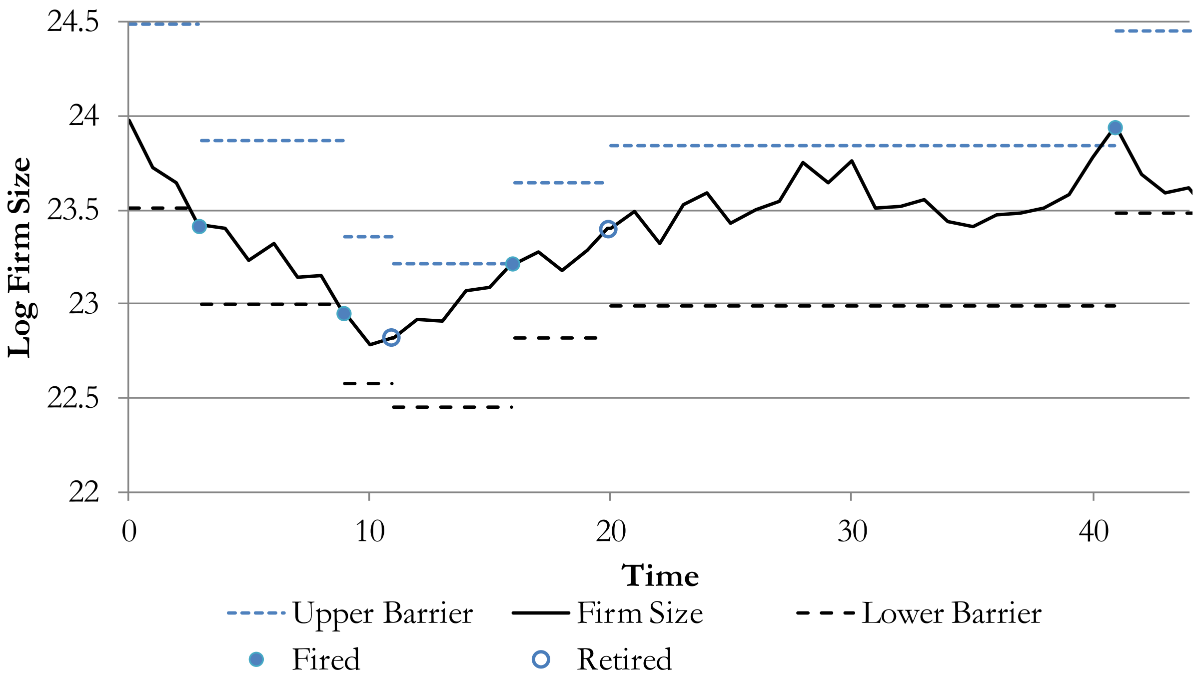


Figure 3
Simulated Path of One Firm

This figure depicts the actions of one simulated firm through time. The solid black line represents the level of the firm's productive assets for each point in time. The dashed black (blue) line illustrates the optimal lower (upper) replacement threshold for the firm. Solid blue markers represent instances where the current CEO was replaced, while hollow blue markers represent periods where the incumbent CEO retired.

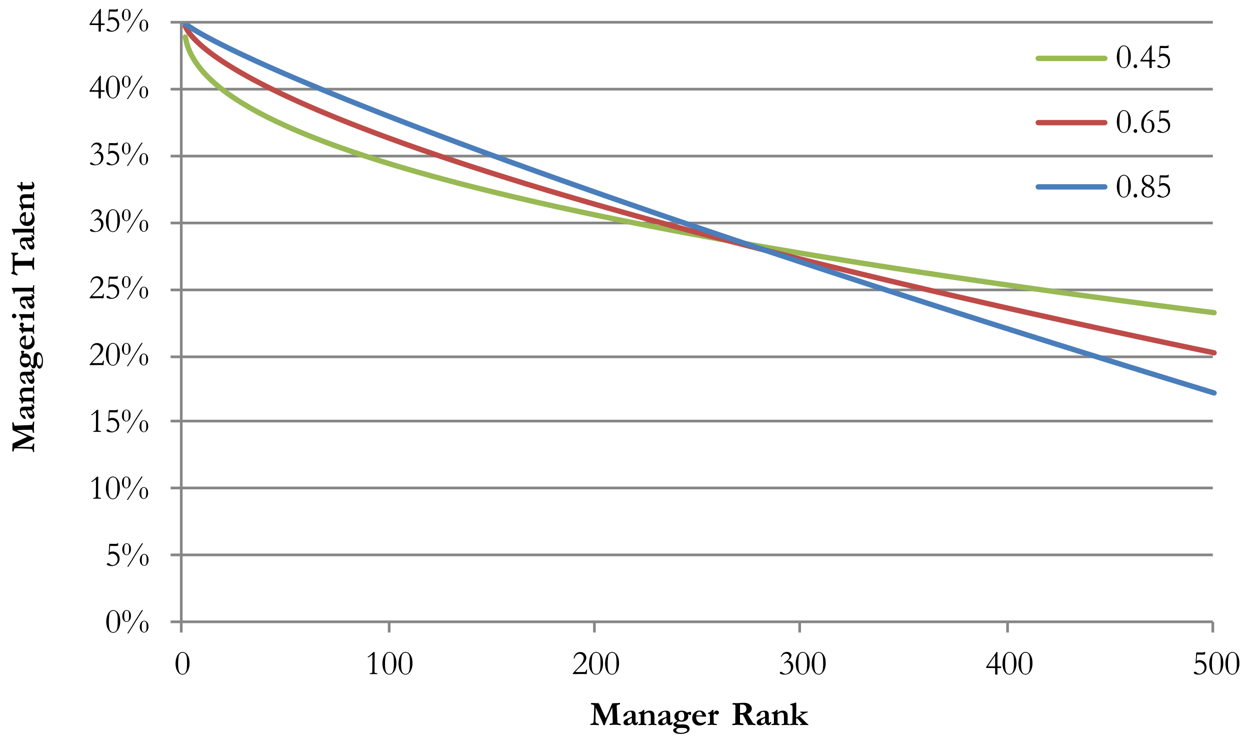


Figure 4
Talent Distribution

This figure illustrates how the convexity of the talent distribution varies with the tail index. The green, red and blue lines plot the talent level of the top 500 most talented mangers for tail indices of 0.45, 0.65, and 0.85 respectively.

Panel A

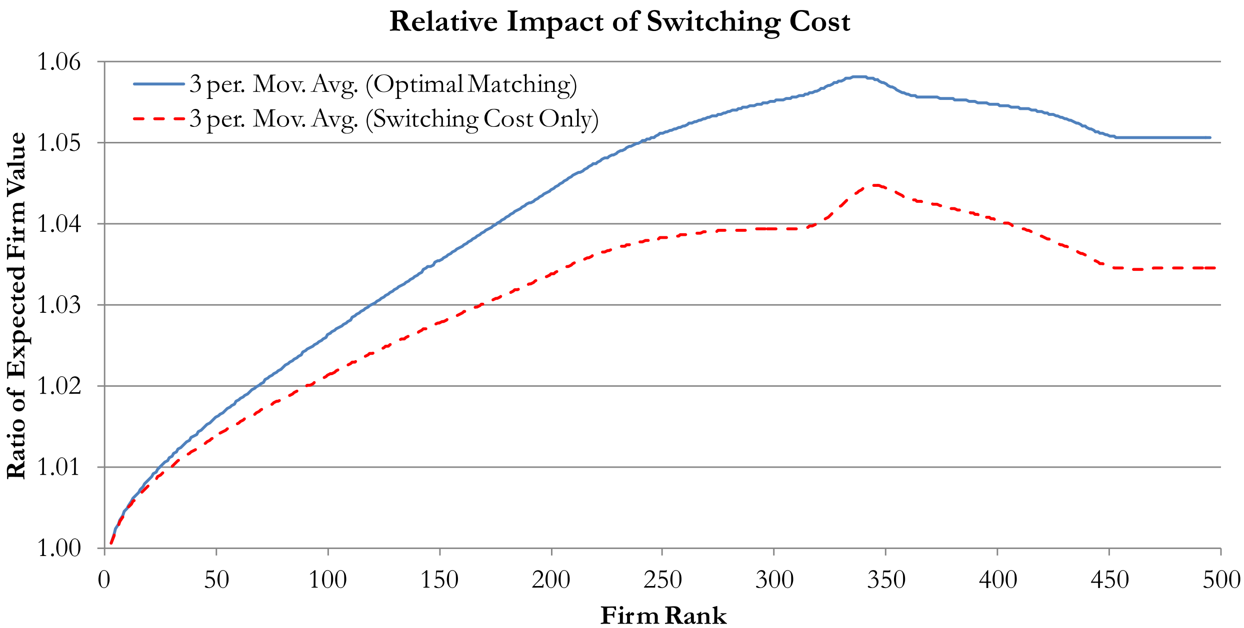
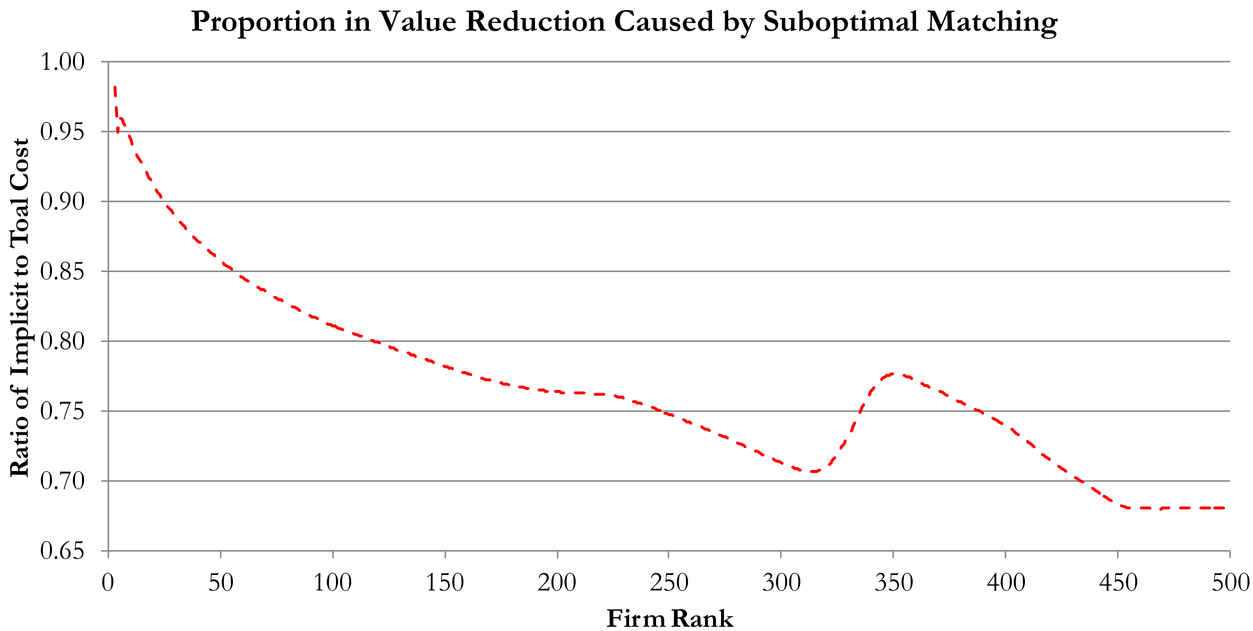


Figure 5

Distortionary Effects of a Switching Cost

This figure illustrates relative expected firm values for a given size ranking under three scenarios (Panel A) and the proportion of value reduction due to suboptimal matching (Panel B). *Optimal Matching* represents an economy with a frictionless labor market. *Matching Friction* represents an economy with a switching cost which leads to the sub-optimal matching of firms and managers. *Switching Cost* represents an economy where firms experience an identical series of explicit switching costs as *Matching Friction*, but are also optimally matched with managers in each period. Panel A reports the three firm moving average of the ratio of expected firm value for the *Optimal Matching* and *Switching Cost* scenarios relative to the *Matching Friction* scenario. Panel B reports the ratio of the decrease in firm value when firms are sub-optimally matched with managers but do not pay the switching cost to the *Switching Cost* scenario.

Panel B


Proportion in Value Reduction Caused by Suboptimal Matching
Ratio of Implicit to Toal Cost
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0
50
100
150
200
250
300
350
400
450
500
Firm Rank

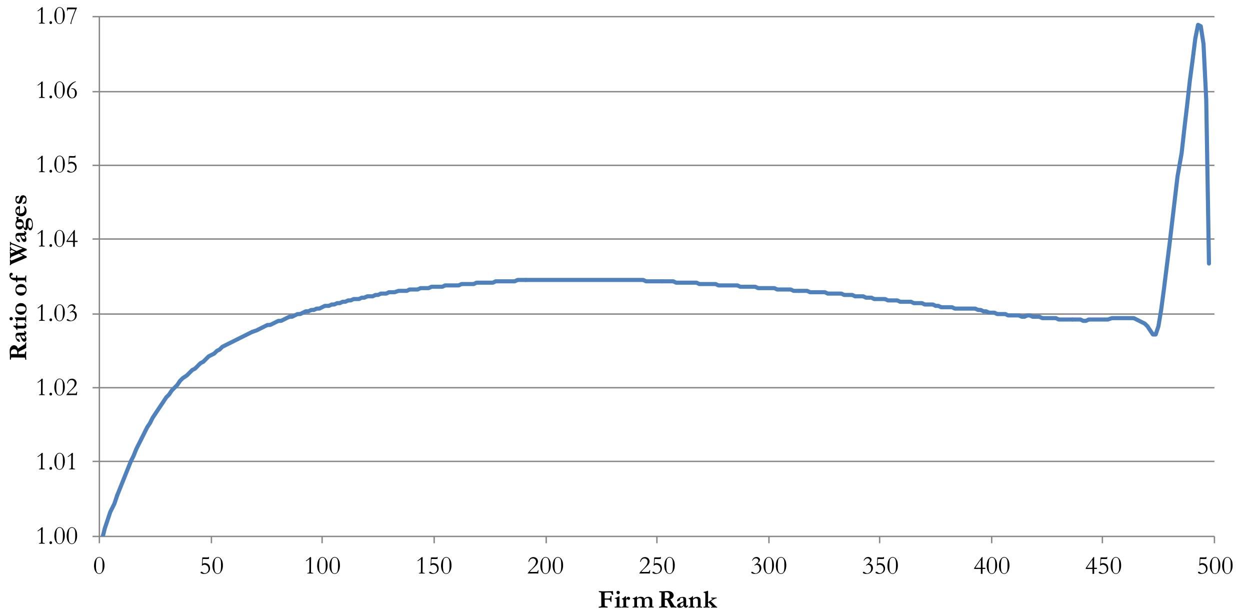


Figure 6
Equilibrium Wages Within a Dynamic and Static Framework

This figure illustrates ratio of equilibrium wages in the dynamic model to the wages within a static framework, using the final estimates of the model parameters. The y-axis reports the ratio of wages under the two models. The x-axis represents the size ranking of the firm being considered, ranked in descending order of total firm value.

Table I
Summary Statistics

	Execucomp		Non-Financial		Final Sample	
	Mean	N	Mean	N	Mean	N
Operating Income ($M)	683.95	33,339	529.00	26,454	1,439.03	8,214
Assets – Total ($M)	9,986.4	45,124	4,370.3	38,232	15,161.2	8,215
Return on Assets	8.60%	31,054	9.33%	24,957	11.53%	8,180
Total Market Capitalization ($M)	12,950.1	44,943	7,361.4	38,089	26,422.5	8,177
Total Executive Compensation ($M)	4.754	32,831	4.727	25,991	8.068	7,956

This table reports summary statistics for the entire universe of Execucomp observations, all non-financial firms, and the final sample used when computing empirical moments.

Table II
Moment Coefficients

	Turnover	ROA	Ln(Pay)	ROA	Var(Δ Rank)	Retire
Intercept	0.042					
	(13.36)					
7+ Years Tenure	0.023					
	(3.33)					
Mean		0.1031				
		(41.13)				
Lagged ROA		0.8527				
		(62.97)				
Variance		0.0027				
		(13.05)				
Ln(Market Value)			0.4217			
			(14.23)			
Turnover (Up, -1)				0.0006		
				(0.16)		
Turnover (Up, 0)				-0.0084		
				(-2.05)		
Turnover (Down, -1)				-0.0065		
				(-1.26)		
Turnover (Down, 0)				-0.0116		
				(-2.08)		
Lagged Rank Q1					266.21	
					(6.20)	
Lagged Rank Q2					1349.92	
					(8.37)	
Lagged Rank Q3					2407.13	
					(15.16)	
Lagged Rank Q4					3452.38	
					(19.57)	
Lagged Rank Q5					3509.87	
					(11.40)	
Prob. Hazard						0.057
						(73.95)

This table reports sample moment estimates using OLS regressions. *Turnover (Up, t)* is a dummy variable taking a value of one when a firm experiences CEO turnover at time *t* conditional of the firm's market value ranking increasing over the CEO's tenure. *Lagged Rank Q1* is a dummy variable taking on a value of one when the firm's lagged size ranking places it in the quintile of firms with the largest market value in the sample. Reported are *t*-statistics with standard errors clustered at the firm level.

Table III
Estimation Results

	Panel A: Parameter Estimates					
	Switching Cost (c)	Maximum Talent (θ_{max})	Average Talent ($\bar{\theta}$)	Tail Index ($\propto$)	Shock Bounds (γ)	Idiosyncratic Noise (σ_{ε}^{2})
Point Estimate	0.0218	0.3014	0.1698	0.7204	0.3084	0.0007
Standard Error	(0.0005)	(0.0005)	(0.0011)	(0.0030)	(0.0010)	(0.0035)

	Panel B: Over-Identification Test
	$\chi_{(9)}$
Test Statistic	95.60
p-value	<.0001

	Panel C: Individual Moment Fits		
	Empirical Moment	Simulated Moment	p-value
Turnover Intercept	0.0423	0.0485	0.0445
7+ Years Tenure	0.0230	0.0180	0.4757
Average ROA	0.1031	0.1012	0.3517
ROA Persistence	0.8527	0.8487	0.7554
ROA Var., within spell	0.0027	0.0018	< 0.0001
Pay-Size Elasticity	0.4217	0.3804	0.1292
Turnover (Up, -1)	0.00063	0.0129	0.0177
Turnover (Up, 0)	-0.0084	0.0056	0.0598
Turnover (Down, -1)	-0.0065	-0.0118	0.3228
Turnover (Down, 0)	-0.0116	-0.0336	< 0.0001
Var(Δ Rank), Q1	266.21	90.79	< 0.0001
Var(Δ Rank), Q2	1349.92	660.79	< 0.0001
Var(Δ Rank), Q3	2407.13	1885.57	0.0006
Var(Δ Rank), Q4	3452.38	3947.49	0.0040
Var(Δ Rank), Q5	3509.87	3145.56	0.2258

This table reports parameter estimates (Panel A), a test of over-identification (Panel B), and a test of equality between empirical and simulated moments (Panel C). Parameter estimate standard errors are reported in parentheses.

Table IA.I
Estimation Results without Pay-Size Elasticity Moment

Panel A: Parameter Estimates						
	Switching Cost (c)	Maximum Talent (θ_{max})	Average Talent ($\bar{\theta}$)	Tail Index ($\propto$)	Shock Bounds (γ)	Idiosyncratic Noise (σ_{ε}^{2})
Point Estimate	0.0214	0.2820	0.1638	0.6758	0.3025	0.0005
Standard Error	(0.0017)	(0.0007)	(0.0020)	(0.0017)	(0.0016)	(0.0041)

Panel B: Over-Identification Test	
	$\chi_{(8)}$
Test Statistic	91.76
p-value	<.0001

Panel C: Individual Moment Fits			
	Empirical Moment	Simulated Moment	p-value
Turnover Intercept	0.0423	0.0463	0.1995
7+ Years Tenure	0.0230	0. 0183	0.5094
Average ROA	0.1031	0.1028	0.8714
ROA Persistence	0.8527	0.8534	0.9522
ROA Var., within spell	0.0027	0.0019	< 0.0001
Pay-Size Elasticity	0.4217	0.3757	0.0910
Turnover (Up, -1)	0.00063	0.0133	0.0148
Turnover (Up, 0)	-0.0084	0.0049	0.0713
Turnover (Down, -1)	-0.0065	-0.0127	0.2435
Turnover (Down, 0)	-0.0116	-0.0356	<0.0001
Var(Δ Rank), Q1	266.21	84.42	< 0.0001
Var(Δ Rank), Q2	1349.92	613.45	< 0.0001
Var(Δ Rank), Q3	2407.13	1754.90	< 0.0001
Var(Δ Rank), Q4	3452.38	3716.70	0.1246
Var(Δ Rank), Q5	3509.87	2991.18	0.0847

This table reports parameter estimates (Panel A), a test of over-identification (Panel B), and a test of equality between empirical and simulated moments (Panel C) analogous to Table III, while excluding the elasticity of CEO pay to firm value in the estimation procedure. Parameter estimate standard errors are reported in parentheses.