

# Reputation, Entrenchment, and Dynamic Managerial Incentives

Noah Lyman<sup>1</sup>

**ABSTRACT:** I estimate a dynamic principal-agent model to study the impact of turnover frictions on managerial incentives. Firms hire CEOs of unknown quality and design compensation contracts. Quality is gradually learned, and firms make replacement decisions based on their beliefs. CEOs are entrenched, so replacement is costly. Entrenchment weakens termination-based incentives, increasing the provision of performance pay in equilibrium. The estimates imply a 9.1% reduction in average CEO pay upon the elimination of entrenchment, though effects are heterogeneous across the distribution of CEO quality. Finally, I find that the level of entrenchment is minimally affected by board independence. (*JEL* M50, D86, D83, G30)

## 1 Introduction

CEOs are rarely fired. In a given year, only 3% of CEOs are forcefully removed from their positions. This fact is often attributed to managerial entrenchment, a term encompassing a variety of inefficient mechanisms through which CEOs are protected from job loss.<sup>2</sup> As discussed by Taylor (2010), entrenchment leads boards to adopt inefficient firing policies at the expense of shareholder value. Furthermore, such protection from replacement can weaken the alignment of incentives between CEOs and shareholders, exacerbating moral hazard and increasing the incentive-aligning level of pay. While it is well understood that entrenchment increases the cost of CEO replacement, what impact this has on managerial incentives is an open question.

In this paper, I estimate a dynamic principal-agent model to quantify the impact of entrenchment on managerial compensation. Boards hire CEOs of uncertain quality which

---

<sup>1</sup>Warwick Business School, University of Warwick. Contact: [Noah.Lyman@wbs.ac.uk](mailto:Noah.Lyman@wbs.ac.uk). I thank my dissertation advisors Andrés Hincapié, Barton Hamilton, and Fei Li for their guidance and support. I also thank Luca Flabbi, Qing Gong, Todd Milbourne, Jeremy Bertomeu, and Iván Marinovic for their helpful comments. This work has also benefited from presentations at the Fuqua School of Business, Olin Business School, Kenan-Flagler Business School, the University of Mississippi, North Carolina State University, the College of Charleston, the University of South Carolina, the 2022 and 2023 Southeastern Economics Association meetings, the 2023 WashU Economics Graduate Student Conference, and the 2023 Summer Meeting of the Econometric Society.

<sup>2</sup>Entrenchment may arise through a number of anti-takeover mechanisms (ex: poison pills, golden parachutes, voting agreements) or through personal relationships between managers and board members.

is gradually learned as employment progresses. CEOs are privately informed of their actions and have limited liability, giving rise to moral hazard. Upon learning that their CEO is of sufficiently low quality, boards can fire their CEO and draw a replacement from a fixed population of executives. Firing a CEO subjects boards to both monetary and non-monetary costs, the latter reflecting entrenchment. Both incentive pay and the threat of termination help motivate the CEO to select efficient actions, though the incentive effects of termination are weakened by entrenchment.

I find that entrenchment has a considerable impact on the incentive-aligning level of CEO pay; the model predicts a 9.1% reduction in average CEO compensation upon the elimination of entrenchment. Entrenchment effectively induces a transfer of surplus from shareholders to the CEO, which I refer to as the *entrenchment premium*. Notably, this premium varies substantially across the distribution of managerial quality. The average marginal effect of entrenchment on compensation is \$907 thousand dollars per year for CEOs in the top quintile of the estimated quality distribution, compared to \$45 thousand per year for CEOs in the bottom quintile. In addition, I find that the level of entrenchment varies minimally with the level of board independence; CEOs monitored by boards with high and low levels of independence extract roughly the same benefits stemming from entrenchment. This is consistent with prior literature (Coles et al., 2014) and suggests that increasing board independence is not an exhaustive remedy for the issue of CEO entrenchment.

This paper makes several contributions to the literature. First, this paper makes a methodological contribution to the growing literature on CEO employment dynamics.<sup>3</sup> I leverage theoretical results from Demarzo and Sannikov (2017) to embed the equilibrium compensation contract within the canonical model of CEO turnover first employed by Taylor (2010), allowing compensation and employment dynamics to be studied in tandem. Second, I contribute to the literature on corporate governance by quantifying the impact of entrenchment on managerial compensation. The role of disciplinary termination in CEO compensation contracts is often overlooked; I endogenize termination and provide empirical evidence that the threat of termination has considerable effects on the provision of managerial incentives. Furthermore, I extend the model to allow for interactions between CEO entrenchment and board independence, permitting an assessment of

---

<sup>3</sup>The use of structural techniques is growing increasingly common in this literature. Taylor (2010) estimates a dynamic model of CEO turnover to measure the impact of entrenchment on firm value. Lippi and Schivardi (2014) use a similar approach to study the impact of concentrated ownership on executive selection. More recently, Ferraro (2021) and Hamilton et al. (2024) have extended the model of Taylor (2010) to respectively study female leadership and nepotism. Barry (2023) estimates a similar model to study the impact of shareholder voice on CEO pay.

whether the effects of entrenchment on managerial incentives are heterogeneous across board types. Finally, this paper contributes to the broader literature in labor economics on job matching under uncertainty about worker quality (Jovanovic, 1979; Miller, 1984; Moscarini, 2005). This literature typically views firms and workers as symmetrically informed about worker quality, which is gradually learned over time. This framework provides a natural explanation of the increasing wage-tenure relationship commonly observed empirically; conditional on a job match surviving, posterior beliefs increase on average, driving wages up with tenure. I present a second explanation for the wage-tenure relationship in environments with asymmetric information between the firm and worker. Namely, as tenure increases, termination incentives fade, placing upward pressure on the incentive-aligning level of compensation.

With a panel of publicly traded North American firms spanning from 1995-2019, I estimate the model using the simulated method of moments. Firm-level information is obtained from Compustat while CEO-level information is obtained from Execucomp. When estimating the extended model, I augment my sample with director data obtained from Institutional Shareholder Services (ISS). Using data provided by Peters and Wagner (2014) and Jenter and Kanaan (2015), I classify cases of CEO turnover as either forced (being fired) or voluntary (retiring). As mentioned previously, the rate of forced CEO turnover is quite low; on average, only 3% of CEOs are fired in a given year. Using a similar structural approach, Taylor (2010) and Hamilton et al. (2024) find strong evidence that this rate is far below its efficient level. I confirm these conclusions, and delve into their implications for managerial incentives, which are not considered in their papers.

Misalignment of incentives between shareholders and CEOs has been an utmost concern since the separation of corporate ownership and control in the early 20<sup>th</sup> century (Edmans et al., 2017). The objectives of shareholders and CEOs are unlikely to perfectly coincide, and perfect monitoring of the CEO is generally taken to be prohibitively costly (Hermalin and Weisbach, 1998). This gives rise to moral hazard, whose costs have received considerable empirical attention.<sup>4</sup> The literature generally agrees: conditional on employing a CEO, misalignment of incentives is harmful for firm value. However, securing a qualified CEO is no easy task. Boards face substantial uncertainty about CEO quality at time of hire,<sup>5</sup> and such information frictions at the hiring margin compound the moral hazard issue (Jovanovic and Prat, 2014; Demarzo and Sannikov, 2017). Given their limited information, boards cannot disentangle the effects of effort and quality when

---

<sup>4</sup>See for example: Margiotta and Miller (2000), Gayle and Miller (2009), Gayle et al. (2015), Page (2018), Ai et al. (2022).

<sup>5</sup>See for example: Hermalin (2005), Taylor (2010), Hamilton et al. (2024), Capron et al. (2024)

monitoring CEO performance, posing an identification problem for boards which CEOs can leverage for their private gain. Reminiscent of the ratchet effect in Laffont and Tirole (1988), CEOs who convey positive information to boards face more demanding incentive schemes later in their employment. CEOs thus have an incentive to convey negative information, which can be achieved by privately expropriating firm resources and increasing board pessimism about future performance. Such behavior is inefficient, and the problem is especially severe when the rate of board learning is slow. Indeed, the model estimates imply that boards learn about CEO quality quite slowly. After 15 years of CEO tenure, roughly 40% of the initial uncertainty remains. This suggests that boards' ability to precisely make inferences about the quality of their CEO is severely limited by the idiosyncratic noise present in the cash flow process. An implication is that some degree of "pay for luck" is to be expected in equilibrium (Bertrand and Mullainathan, 2001; Garvey and Milbourn, 2006; Daniel et al., 2019); even if a CEO is of relatively low quality, sequences of good luck may lead boards to perceive their CEO as high quality and award pay accordingly. The estimates suggest that the prevalence of pay for luck in equilibrium is substantial; among CEOs at the 90<sup>th</sup> percentile of the compensation distribution, roughly one in four have quality below that of the average replacement.

Low-quality CEOs can have a detrimental impact on shareholder value, as managerial quality has been found to be an important determinant of firm performance.<sup>6</sup> The existing literature has documented high variation in CEO quality, implying that the difference between a high and low-quality CEO is quite pronounced. I confirm this result; for the median-sized firm in my sample, the model estimates imply a \$155 million dollar difference in yearly cash flows between a CEO at the 95<sup>th</sup> and 5<sup>th</sup> percentiles of the quality distribution. Thus, firms stand to gain a substantial amount of value through the prompt replacement of low-quality CEOs. Despite this, the firing option is rarely exercised, as the estimated level of entrenchment is quite high. Entrenchment reflects a conflict of interest between shareholders and boards of directors, to whom CEO replacement duties are delegated. In particular, boards may consider non-pecuniary factors, independent of shareholder value, when making CEO termination decisions. As an example, these non-pecuniary factors might reflect personal relationships with the CEO which lead the board to view forceful replacement as undesirable. Additionally, firing a CEO may reflect poorly on the board, who hired the CEO to begin with. Thus, CEO dismissals may also subject directors to reputational costs. In light of these considerations, I follow Taylor (2010) and represent CEO entrenchment as a non-pecuniary cost incurred by the board

---

<sup>6</sup>See for example: Bertrand and Schoar (2003), Allgood and Farrell (2003), Pérez-González (2006), Bennedsen et al. (2007), Mackey (2008), Kaplan et al. (2012).

when firing their CEO.

A high degree of entrenchment is generally associated with weak corporate governance (Gompers et al., 2003; Bebchuk et al., 2009). In the corporate finance literature, managerial entrenchment has been argued to be an important determinant of a number of firm outcomes including investment decisions (Shleifer and Vishny, 1989), capital structure (Zwiebel, 1996; Berger et al., 1997), and dividend payout policy (Hu and Kumar, 2004). I contribute to this literature by quantifying both the effect of entrenchment on equilibrium CEO compensation and the sensitivity of this effect to board independence. Reduced-form analysis, which has proven to be challenging in this literature,<sup>7</sup> is poorly suited for my research question as compensation and termination decisions are endogenous. I tackle this challenge by instead taking a structural approach, modeling explicitly the relationship between turnover, entrenchment, and compensation. In the model, termination and compensation policies are jointly determined in equilibrium and serve as alternative levers through which the board can incentivize the CEO. Entrenchment increases the effective cost of forced termination, inducing boards to substitute and use more performance-pay in equilibrium. Given the estimates, this amounts to a roughly 9.1% increase in average yearly CEO pay. Further, I find that the level of CEO entrenchment decreases with board independence, but not drastically. I estimate the effective cost of replacing a CEO to be roughly 15% higher for boards with low levels of independence relative to those with high levels of independence.

The results of this paper highlight the importance of turnover frictions when studying moral hazard in executive labor markets. Considerations of turnover are often neglected in empirical studies on moral hazard, despite the consensus in the theoretical literature that turnover serves as a useful incentive device. The incentive effects of turnover have been studied as far back as Stiglitz and Weiss (1983), and more recently by Spear and Wang (2005) and DeMarzo and Fishman (2007) in discrete time settings. More recently, the incentive effects of turnover have received attention in a continuous time setting (Sannikov, 2008; Biais et al., 2010; Zhu, 2013; Demarzo and Sannikov, 2017; Grochulski and Zhang, 2023). Continuous-time methods improve the tractability of dynamic incentive problems, as the derivation of optimal incentives amounts to solving a partial differential equation, which can be done numerically with low computational burden. For improved tractability, I adopt a continuous-time approach in this paper. Importantly, I model CEOs as risk-neutral with limited liability, the latter constraint serving as the source of the moral hazard problem. Thus, termination serves as a punishment of last resort when boards' ability to punish CEOs through financial means is exhausted.

---

<sup>7</sup>See for example: Lehn et al. (2007), Bebchuk et al. (2009), Chang and Zhang (2015)

On the theoretical front, the papers most similar to this one are Jovanovic and Prat (2014) (JP) and Demarzo and Sannikov (2017) (DS). Both papers develop dynamic models of moral hazard in which the principal and agent face symmetric uncertainty about agent quality. JP allows learning to be non-stationary while DS restricts attention to stationary learning. Uncertainty reduction over tenure is important to consider when studying the dynamic selection of workers, so I adopt the JP assumption of non-stationary learning. However, JP assumes agents are risk-averse with unlimited liability and does not consider turnover. I explicitly model turnover, both endogenous termination and exogenous retirement, and assume agents to be risk-neutral with limited liability. The assumption of limited liability is important in this paper, as it leaves a role for disciplinary termination in the equilibrium contract. Endogenous termination is included in DS, but agent replacement is not; rather, firms liquidate and cease operations when their agent is terminated. In my model, agents are replaced following an instance of turnover and firms continue operations. On the empirical front, the paper most similar to mine is Taylor (2010) who uses a structural approach to explain the low rate of forced CEO turnover observed empirically. Taylor finds that the key determinant of this empirical regularity is CEO entrenchment, and offers a compelling argument suggesting entrenchment is detrimental for firm value. I build upon his paper by embedding the optimal contract into his framework, allowing for a detailed analysis of the equilibrium response of CEO compensation to turnover frictions. I find that the response is considerable, and show that turnover frictions have major implications for the severity of moral hazard in the executive labor market.

The remainder of the paper proceeds as follows. Section 2 outlines the sample and key empirical patterns motivating the structural model. Section 3 presents the theoretical environment and derivation of firms' optimal compensation and turnover policies. Section 4 discusses model identification and the estimation procedure. Section 5 presents the model estimates. The impact of entrenchment on equilibrium compensation is analyzed through counterfactual experiments in Section 6. Section 7 presents the estimates and results of the extended model with board independence. Closing discussion and concluding remarks are contained in Section 8.

## 2 Data

The sample is a matched CEO-firm yearly panel of North American publicly-traded firms spanning 1994-2019. The panel is constructed by linking three sources of data. First, I obtain CEO-level information from Execucomp, which provides detailed data on com-

pensation packages and CEO tenure. I match this with Compustat which reports information on firm assets, industry classification, income, and other financial fundamentals. Firm performance is measured using their return on assets (ROA), defined as operating income per dollar in assets.<sup>8</sup> Firms with missing operating income or missing total assets are omitted from the final sample. Lastly, I match the sample with supplementary turnover data (Peters and Wagner, 2014; Jenter and Kanaan, 2015), which classifies instances of CEO turnover as forced or voluntary using the method outlined by Parrino (1997). The final sample consists of 42,513 firm-year observations, 3,627 distinct firms, and 8,191 distinct CEO employment spells. I observe 5,005 cases of CEO turnover, where 1,260 (25.2%) are forced and 3,745 (74.8%) are voluntary.<sup>9</sup> Throughout this section, I index CEOs by  $i$ , firms by  $j$ , and calendar years by  $t$ .

Summary statistics are reported in Table 1. CEOs in the sample are predominantly male; only 2.8% of CEOs are female. The average executive is 53 years in age. The majority of CEOs have some prior experience with their firms prior to becoming CEO. On average, an executive has roughly 9 years of firm-specific experience when appointed for the CEO position. The average length of employment as CEO is 8.29 years, though the spell length distribution exhibits substantial rightward skewness. Of the CEO employment spells which are not right-censored, roughly 26% end in forced termination.

The level of CEO pay is substantial and largely attributable to performance pay, with fixed salary making up only about 30% of total compensation for the average CEO (Figure 1). Performance pay is composed primarily of equity incentives, the use of which in executive compensation packages is well documented.<sup>10</sup> The use of performance pay is intended to align the interests of the CEO with those of the firm, mitigating the CEO's motivation to pursue private interests (Margiotta and Miller, 2000). In addition, the threat of forced termination provides incentives by serving as a punishment of last resort (Spear and Wang, 2005). Performance pay and the threat of termination thus complement each other in the incentive mix, rewarding the CEO in cases of positive performance and punishing the CEO in cases of persistent negative performance.

---

<sup>8</sup>ROA is defined as operating income ( $oibdp_{ijt}$  in Compustat) divided by total assets ( $at_{ijt}$  in Compustat):  

$$y_{ijt} = \frac{oibdp_{ijt}}{at_{ijt}} \times 100$$

<sup>9</sup>See Appendix B.1 for extended details on the turnover classification and construction of the sample.

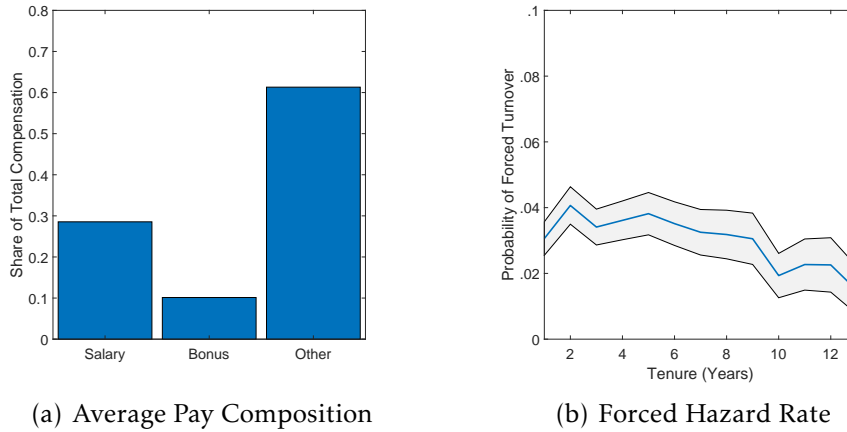
<sup>10</sup>Through the 1990s and beyond, stock and option packages surged to become the dominant component of CEO compensation. See Edmans et al. (2017) for a detailed discussion.

Table 1: Summary Statistics

	Mean	Std. Dev.	5 <sup>th</sup> Percentile	Median	95 <sup>th</sup> Percentile
<i>Panel (a): Firm Characteristics</i>					
Profitability (ROA)	11.9	11.8	-1.75	12.0	28.5
Total assets (\$ Billions)	17.1	103.5	.178	2.22	51.5
Total revenue (\$ Billions)	6.93	21.0	.136	1.66	27.4
<i>Panel (b): CEO Characteristics</i>					
Spell length	8.29	7.12	1	6	22
Eventually fired	.259	.438	0	0	1
Eventually retired	.743	.437	0	1	1
<i>Panel (c): CEO Compensation (\$ Millions)</i>					
Total compensation	6.32	12.2	.601	3.72	18.9
Salary	.866	.457	.327	.809	1.56
Bonus	.549	1.88	0	0	2.40
Bonus (Conditional on > 0)	1.23	2.66	.058	.643	3.88
Other compensation	4.91	11.9	.015	2.44	16.2

Notes: The unit of observation in Panels (a) and (b) is a firm-year. The unit of observation in Panel (c) is a CEO. All monetary values are expressed in 2015 dollars.

Figure 1: CEO Pay and Forced Turnover



Notes: Panel (a) plots the composition of CEO compensation packages on average. Roughly 30% of total CEO pay can be attributed to salary, while the rest is split between cash bonuses and other forms of incentive pay. “Other” is composed of equity incentives including restricted stock grants, option grants, and long-term incentive payouts. Panel (b) plots the rate of forced turnover over the first 15 years of CEO tenure. The likelihood of forced termination is low and gradually declines with tenure.

## 2.1 Stylized Facts

Next, I outline the key empirical facts motivating the structural model. The data suggest that variation in CEO quality is high; CEOs of relatively high quality generate substantially higher rates of profitability than their low-quality counterparts. Furthermore, firms are responsive to new information about the quality of their CEO, revealed by the firm’s financial performance. As information is generated, firms adjust their pay contracts, and



in the case of exceptionally poor performance, exercise their firing option. Low-quality CEOs are gradually forced out of employment, inducing positive selection of CEO quality over tenure. As tenure increases, firms substitute away from termination incentives towards monetary incentives. Thus, as CEO employment progresses, those who survive into later years of tenure are increasingly motivated by the use of incentive pay and decreasingly motivated by the risk of forced termination.

**CEOs are rarely fired.** Panel (b) of Figure 1 plots the rate of forced turnover over the first 13 years of CEO tenure. As previously documented in the literature, CEOs are unlikely to be fired. The likelihood of forced termination is highest in early years of tenure, peaking at roughly 4% and otherwise generally declining with tenure. There are many potential explanations underlying the low rate of forced termination. As discussed by Taylor (2010), replacing a CEO is quite costly, so boards may only exercise their firing option as a last resort. Alternatively, if CEOs are relatively homogeneous in the population, CEO replacement may have minimal impact on the trajectory of firm performance. I argue next, however, that this second possibility is unlikely.

**Variation in CEO quality is high.** CEO quality has been shown to be extremely consequential for firm performance (Allgood and Farrell, 2003; Bertrand and Schoar, 2003; Bennedsen et al., 2020). While CEO quality is not directly observed in the data, I create a reduced-form proxy for it using observed firm profitability. Specifically, I project firm profitability on a vector of firm and CEO characteristics, and obtain the shrinkage estimate of CEO quality given as a weighted average of the resulting profitability residuals. I begin by estimating the equation:

$$y_{ijt} = \lambda_0 + \lambda_1 \mathbf{C}_{ijt} + \lambda_2 \mathbf{F}_{ijt} + \tau_t + \gamma_j + \epsilon_{ijt} \quad (1)$$

where  $y_{ijt}$  denotes the return on assets for firm  $j$  employing CEO  $i$  at time  $t$ .<sup>11</sup>  $\mathbf{C}_{ijt}$  and  $\mathbf{F}_{ijt}$  are a vector of CEO and firm characteristics, respectively.<sup>12</sup>  $\tau_t$  and  $\gamma_j$  are year and firm fixed effects. Next, I use the residual component of Equation (1) to create a coarse proxy of CEO quality. For each CEO-firm match  $ij$ , denoting their length of employment

---

<sup>11</sup>The use of ROA as the main measure of firm performance is consistent with the related methodological literature (Taylor, 2010; Hamilton et al., 2024; Capron et al., 2024). I follow suit in the interest of maximizing comparability of results with those of previous studies.

<sup>12</sup>Specifically, I control for the CEO's age, gender, and tenure (as CEO), along with firm assets and total revenue.

by  $s_{ij}$ , the shrinkage estimate of CEO match quality  $\theta_{ij}$  is defined by:

$$\hat{\theta}_{ij} = \frac{\omega_j}{1 + \omega_j} \left( \frac{1}{s_{ij}} \sum_{ij'} \hat{\epsilon}_{ij't} \mathbf{1}[ij' = ij] \right) \quad (2)$$

$$\omega_j = \text{Var}(\hat{\epsilon}_{ij't} | j' = j) \quad (3)$$

$\hat{\theta}_{ij}$  is the James-Stein estimator of true match quality  $\theta_{ij}$  and  $\hat{\epsilon}_{ij't}$  are fitted residuals obtained from estimating Equation (1).<sup>13</sup> Table 2, summarizes the distribution of  $\hat{\theta}_{ij}$  across all CEOs in the sample. Its standard deviation is 5.4, which lies within the range found

Table 2: Distribution of CEO Quality  $\hat{\theta}_{ij}$

	Mean	Std. Dev.	5 <sup>th</sup> pct	50 <sup>th</sup> pct	95 <sup>th</sup> pct
Quality Proxy ( $\hat{\theta}_{ij}$ )	-.171	5.40	-7.15	0.00	6.51

in previous literature.<sup>14</sup> For the median-sized firm in the sample, the implied difference between cash flows generated by a CEO at the 5<sup>th</sup> and 95<sup>th</sup> percentile of the distribution of  $\hat{\theta}_{ij}$  is roughly \$303 million per year.<sup>15</sup> This staggering difference suggests that the difference between a high and low-quality CEO is quite pronounced.

**CEO quality is positively selected over tenure** Notably, the distribution of  $\hat{\theta}_{ij}$  evolves with CEO tenure. Figure 2 shows that the across-CEO average of  $\hat{\theta}_{ij}$  increases over tenure (Panel (a)) while its variance decreases (Panel (b)). As  $\hat{\theta}_{ij}$  is fixed within a given CEO, this pattern is suggestive of positive selection on quality; firms exercise their firing option in response to poor signals of CEO quality, inducing gradual attrition of low-quality CEOs.<sup>16</sup> As low quality CEOs gradually exit the sample, the average quality of those who survive is pushed upwards. Furthermore, the termination of low quality CEOs compresses the variance of the quality proxy among those employed, as the distribution becomes more concentrated around relatively high levels of quality.

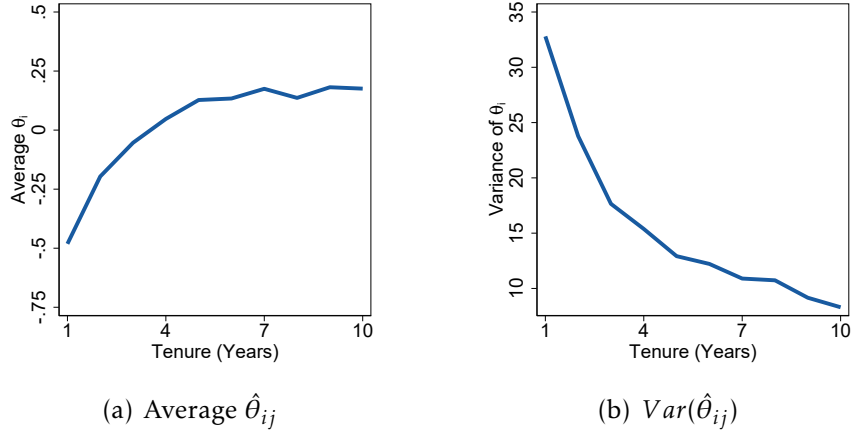
<sup>13</sup>Match quality  $\theta_{ij}$  is clearly measured with error in the data. Given this measurement error, the James-Stein estimator, while biased, minimizes mean-squared error among the set of admissible  $M$ -estimators.

<sup>14</sup>For example: Taylor (2010) reports a SD of 2.42 while Bertrand and Schoar (2003) report a SD of 7.

<sup>15</sup>The median-sized firm in the sample has \$2.219 billion in assets. The approximate \$303 value is obtained by:  $2219 \times \frac{6.51+7.15}{100} \approx 303$ .

<sup>16</sup>I show in Appendix B.1.2 that forced turnover decisions are sensitive to cumulative performance, while voluntary turnover decisions are statistically independent of cumulative performance.

Figure 2: Distribution of  $\hat{\theta}_{ij}$  over Tenure



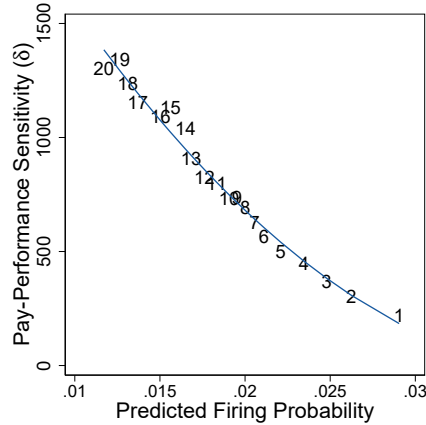
Notes: Panel (a) plots the average value of the CEO quality proxy  $\hat{\theta}_{ij}$  over the first 10 years of tenure while Panel (b) plots the variance. Together, the figures suggest that the distribution of CEO quality becomes increasingly concentrated among high values as tenure increases.

**Termination incentives and monetary incentives are substitutes.** Given the evolution of the CEO quality distribution, I next consider the co-evolution of the incentive mix. Let  $\delta_{ijt}$  denote contract  $ij$ 's pay-performance sensitivity in year  $t$  as calculated in Core and Guay (2002) and Coles et al. (2006).<sup>17</sup> Specifically,  $\delta_{ijt}$  is defined as the year  $t$  monetary return (in \$1000s) that CEO  $i$  would receive in response to a 1% increase in firm  $j$ 's stock price.

Figure 3 plots the evolution of pay sensitivity and the predicted firing probability over CEO tenure. Each point corresponds to a level of tenure. Long-tenured CEOs, who have exhibited relatively high levels of performance on average (Figure 2), face a very low risk of forced termination and are motivated primarily by monetary incentives. The opposite is true for newly-tenured CEOs; monetary incentives are weaker relative to long-tenured CEOs while termination-based incentives are stronger. Figure 3 gives a sharp depiction of this gradual substitution away from termination-based incentives towards monetary incentives. This substitution suggests a change in the firm's relative cost of providing termination versus monetary incentives. In particular, the observed pattern is consistent with termination incentives becoming increasingly costly relative to monetary incentives as CEO tenure increases. In other words, it is more costly to threaten to fire a proven, long-tenured CEO than a brand new CEO of uncertain quality.

<sup>17</sup> $\delta_{ijt}$  is calculated as the change in option portfolio value in response to a 1% increase in the firm's stock price, where options are valued using the standard model of Black and Scholes (1973) as modified by Merton (1973) to accommodate for dividend payouts. See Core and Guay (2002) for more details.

Figure 3: Incentive Pay, Termination Risk, and CEO Tenure



Notes: For each of the first 20 years of CEO tenure, Figure 3 scatters the average pay-performance sensitivity over the predicted firing probability conditional on firm characteristics, tenure, and the CEO's history of performance. Pay-performance sensitivity is measured following Core and Guay (2002) and Coles et al. (2006) and gives the number of dollars (in thousands) a CEO would receive in response to a 1% increase in their firm's stock price. The figure reveals that CEO's risk of termination decreases with tenure, while their pay sensitivity increases. The risk of termination has a weaker impact on incentives late in CEOs' careers; these weakened termination incentives are gradually replaced with financial incentives.

## 2.2 Discussion

The stylized facts presented in this section provide a descriptive glimpse into the role of reputation in CEO employment dynamics. It is worth emphasizing that the patterns observed are the product of complex endogenous relationships between compensation, turnover, and perceptions of CEO quality. Consequently, the reduced-form results reported here are almost certain to be contaminated by bias. For instance, Figure 2 suggests that CEO quality and tenure are correlated, in which case estimating CEO quality using residual performance will yield biased estimates.<sup>18</sup> To develop a more credible understanding of the influence of reputation and turnover on the provision of CEO incentives, it is essential to address the limitations of this descriptive analysis and provide a more rigorous account of the underlying decision problem.

To this end, I construct in the next section a dynamic agency model to capture the joint dynamics of CEO reputation, turnover, and incentives. I explicitly model boards' decision over both CEO dismissals and contract design, addressing both selection bias and the endogeneity of CEO reputation and incentive provision (Milbourn, 2003). This approach lays the groundwork for a more nuanced interpretation of the stylized facts, linking them to the strategic decision-making processes that govern CEO employment and compensation.

<sup>18</sup>See Chetty et al. (2014) for an analogous discussion in the context of estimating teacher value-added.

### 3 Model

The model features two types of decision makers: boards and CEOs. Each board acts on behalf of their firm and employs one CEO at a time. Firms' rate of profitability is determined by their level of productivity, turnover costs, idiosyncratic shocks, and their CEO's quality<sup>19</sup> and private actions.<sup>20</sup> Neither quality nor actions are directly observed by boards. Rather, quality is gradually learned by observing profitability and efficient actions are implemented through the board's design of a full-commitment contract. The contract specifies optimal compensation and termination policies which serve as alternative mechanisms through which incentives can be delivered to the CEO. Termination occurs when the board believes their CEO to be of sufficiently low quality, at which point a replacement CEO is hired and operations continue. The board designs contracts to maximize firm value plus non-pecuniary factors associated with CEO turnover. The presence of such non-pecuniary factors in boards' objective induces a wedge separating firm value and board utility. As such, the equilibrium contract may not be shareholder-optimal in the sense that it does not purely maximize firm value.

Both boards and CEOs are risk-neutral. Each firm<sup>21</sup>  $j$  employs a CEO  $i$  whose current level of tenure is denoted by  $t \geq 0$ . CEO quality  $\theta_i$  is drawn from population distribution  $N(\theta_0, \delta_0^2)$  and remains fixed over time.<sup>22</sup>  $\theta_i$  is not known by either party, but is gradually learned as cash flows are realized.  $\tilde{\theta}_{ijt}$  and  $\tilde{\theta}_{ijt}^a$  respectively denote the board and CEO's estimate of  $\theta_i$  at tenure  $t$ , which do not coincide in general. If the board believes quality to be sufficiently low, they will fire their CEO and hire a replacement; firing occurs at stopping time  $T$ . Otherwise, CEOs retire stochastically at Poisson rate  $\lambda$ ; retirement occurs at random time  $R$ .<sup>23</sup> I let  $\tau_{ij} = \min\{T, R\}$  denote the time at which CEO  $i$ 's employment within firm  $j$  ends, whether by retiring or being fired. Firms incur cost  $c$  in the case of firing or retirement, representing the monetary costs associated with replacing a CEO. Additionally, boards incur non-pecuniary cost  $\pi$  when firing their CEO, reflecting

---

<sup>19</sup>Throughout this paper, "quality" is taken to be synonymous with skill, ability, or fitness for the CEO position. It is beyond the scope of this study to engage in a detailed philosophical inquiry of what constitutes managerial quality—whether it reflects the ability to select profitable projects, interpersonal skills, strategic vision, or other factors. Instead, I follow related literature and treat quality as a latent, CEO-specific attribute with an additive effect on firm performance. For an in depth discussion of what comprises managerial quality, see Kaplan et al. (2012).

<sup>20</sup>All firms operate independently of one another, so broader market equilibrium considerations are not addressed in this paper.

<sup>21</sup>Because each firm has a single board, the  $j$  subscript can be seen as indexing both firms and boards.

<sup>22</sup>I abstract from learning by doing and show in Appendix B.1.1 that the assumption of time-invariant CEO quality is supported in the data.

<sup>23</sup>The assumption of exogenous CEO retirement is empirically motivated. I show in Appendix B.1.2 that voluntary CEO turnover is statistically independent of CEOs' history of performance.

CEO entrenchment.<sup>24</sup> Both  $c$  and  $\pi$  are measured as a fraction of firm assets, allowing turnover costs to vary with firm size. The tenure index  $t$  resets to zero upon the replacement of a CEO. Boards operate over an infinite horizon, hiring successive CEOs whenever the previous one departs.

**Firm Profitability** Each firm  $j$  has two unique characteristics: total assets  $b_j$  and a productivity parameter  $\gamma_j \sim N(0, \sigma_\gamma^2)$ , both of which I assume to be known and constant over time.<sup>25</sup> Firm cash flows gross of CEO pay are denoted by  $dX_{ijt}$  and have the dynamics:

$$dX_{ijt} = b_j \overbrace{(\gamma_j dt + dY_{ijt})}^{\text{Profitability Rate}} \quad (4)$$

$$dY_{ijt} = (\theta_i - a_{ijt})dt + \sigma_W dW_{ijt} \quad (5)$$

$$X_{ij0} = -b_j c + X_{(i-1)j\tau_{(i-1)j}} \quad (6)$$

For firm  $j$  with CEO  $i$ , the rate of profitability at tenure  $t$  is obtained by dividing cash flows  $dX_{ijt}$  by assets  $b_j$ . Profitability has two components. The drift term  $\gamma_j dt$  represents the firm's known contribution to profitability, which is independent of the CEO.  $dY_{ijt}$  is the contribution of unobservables to profitability, which I henceforth refer to as *residual performance*.  $dY_{ijt}$  increases in the CEO's quality and decreases in their private action  $a_{ijt} \geq 0$ , representing the diversion of firm cash flows towards the CEO's private consumption.<sup>26</sup>  $W_{ijt} = \int_0^t dW_{ijs}$  is a standard Brownian Motion on probability space  $\{\Omega, \mathcal{F}, P\}$ ,<sup>27</sup> and the parameter  $\sigma_W$  measures the volatility of performance shocks. At the beginning of CEO  $i$ 's employment spell, cumulative cash flows initialize at  $X_{ij0}$ . This initial condition is given by the cumulative level of cash flows at the time of departure for the previous CEO, denoted by  $X_{(i-1)j\tau_{(i-1)j}}$ , minus the monetary turnover cost  $b_j c$ . Because  $c$  is represented as a fraction of assets, scaling by firm assets converts the units of turnover costs to

<sup>24</sup>Modeling entrenchment as a non-pecuniary cost follows the specification of previous structural analyses of CEO entrenchment (Taylor, 2010; Hamilton et al., 2024; Capron et al., 2024). This reflects boards' distaste for forced CEO dismissals stemming from, for example, personal relationships with the CEO, reputational costs, or the fact that director turnover often follows CEO turnover.

<sup>25</sup>The analytic results would be unchanged if  $\gamma_j$  or  $b_j$  varied over time, as long as they are observed separately from  $dY_{ijt}$ . The assumption of constant firm assets is equivalent to assuming that all profits are immediately paid as dividends to shareholders. This allows me to abstract from dividend payout decisions, which are outside of the scope of this paper. Assets are denoted in millions of dollars; I take as given that  $b_j > 1$  for all  $j$ , so all firms have at least \$1 million in assets. For reference, the minimum value of the empirical distribution of assets is roughly \$57 million.

<sup>26</sup>It is equivalent to interpret  $a_{ijt}$  as managerial effort, where the case of  $a_{ijt} = 0$  corresponds to full effort exertion.

<sup>27</sup> $\Omega$  is the set of all sample paths of  $\{W_{ijt}\}_{t \geq 0}$ , with  $\omega \in \Omega$  denoting an arbitrary sample path.  $P$  is a probability measure over  $\mathcal{F}$ , a  $\sigma$ -algebra over  $\Omega$ .

dollars.

**Preferences and outside options** CEOs' flow utility is given by:<sup>28</sup>

$$u(w_{ijt}, a_{ijt}) = w_{ijt} + \phi_j a_{ijt} \quad (7)$$

$$\phi_j = b_j^\alpha \quad \alpha \in (0, 1) \quad (8)$$

$w_{ijt}$  is the tenure- $t$  realization of compensation specified by the  $\mathcal{F}_{ijt}$ -adapted process  $w : [0, T] \times \Omega \rightarrow \mathbb{R}_+$  and  $\phi_j$  measures the size-dependent rate of cash flow diversion. Concretely, diverting  $a_{ijt}$ % of cash flows from the firm yields  $\phi_j a_{ijt}$  dollars directly to the CEO, where  $\phi_j$  increases with firm size. Hence, cash flow diversion is more profitable for CEOs in larger firms, implying that CEOs' incentive for misbehavior grows with firm size (Gayle and Miller, 2009).<sup>29</sup> Imposing  $\alpha < 1$  renders cash flow diversion inefficient and allows us to restrict attention to contracts which implement no diversion ( $a_{ijt} = 0$ ) for all  $t$ .<sup>30</sup>

Upon exiting from the firm, the CEO receives outside option  $C(\tilde{\theta}_{ijt}^a)$  which depends explicitly on their reputation at time of departure, capturing the influence of career concerns (Fama, 1980; Gibbons and Murphy, 1992; Holmström, 1999; Fee and Hadlock, 2015).<sup>31</sup> Specifically, I assume CEO outside options are an increasing, linear function of their perceived quality  $\tilde{\theta}_{ijt}^a$ :

$$C(\tilde{\theta}_{ijt}^a) = \mu + \frac{\phi_j}{r} \tilde{\theta}_{ijt}^a \quad (9)$$

Hence, high-quality CEOs will (on average) have a high outside payoff, and thus will be more expensive to retain than their low-quality counterparts. The assumption that  $\frac{dC}{d\theta^a} = \frac{\phi_j}{r}$  is not innocuous. I illustrate in Appendix A.1 that this slope assumption implies

<sup>28</sup>As is conventional, I assume compensation and effort are additively separable in the interest of tractability. For an example of optimal incentive provision with multiplicative utility functions, see Edmans et al. (2008).

<sup>29</sup>Further, this will imply that CEOs employed in large firms are compensated more than CEOs in small firms, all else equal. This is consistent with the well-known fact that CEO pay increases with firm size (Gabaix and Landier, 2008).

<sup>30</sup>See Appendix B.2 for a proof of this statement.

<sup>31</sup>Tying outside options to beliefs implies that managers transmit information about their fitness as CEO to the market while employed in a CEO position. At first glance, this may appear inconsistent with the earlier assumption that boards have no information about the quality of incoming CEOs, aside from knowledge of the distribution  $N(\theta_0, \delta_0^2)$ . However, it is important to note that the vast majority of CEOs in the sample have never held a prior CEO position (Cziraki and Jenter, 2024). In fact, over 95% of managers in my sample have only held one CEO position in their career. Consequently, at the time of hire, managers generally have had no prior opportunity to publicly demonstrate their quality as a CEO, leaving firms reliant on the prior distribution of quality to form beliefs.

that the marginal benefit associated with an increase in private beliefs is the same within and outside of the firm, vastly improving both the analytic and numerical tractability of the model. Furthermore, this slope assumption has little effect on the key qualitative results of the model.

Both parties discount the future at rate  $\rho$ . However, given the possibility of exogenous separation, the arrival rate of CEO retirement  $\lambda$  will be absorbed into the discount rate. For notational brevity, I therefore define  $r \equiv \rho + \lambda$  as the effective discount rate. Given a strategy  $a : [0, T] \times \Omega \rightarrow \mathbb{R}_+$ , the CEO's net present value of employment at tenure  $t$  is given by:<sup>32</sup>

$$U_{ijt}^a = \mathbb{E}_t^a \left[ \int_t^T e^{-r(s-t)} u(w_{ijs}, a_{ijs}) ds + \lambda \int_t^T e^{-r(s-t)} C(\tilde{\theta}_{ijs}^a) ds + e^{-r(T-t)} C(\tilde{\theta}_{ijT}^a) \right] \quad (10)$$

The first term reflects the discounted flow payoffs accumulated during employment. The second term reflects the possibility of future retirement, in which case the CEO departs and collects their outside option. If the CEO is fired prior to the arrival of a retirement shock, they will also depart and collect their outside option, as reflected by the third term. Note that  $U_{ijt}^a$  is the continuation payoff given an arbitrary strategy  $a$ , which in general does not coincide with the board's recommended strategy, denoted by  $a^*$ .

I assume boards are risk-neutral and maximize the expected net present value of cash flows net of CEO pay and the non-pecuniary cost of termination. At the outset of the contractual relationship, the firm's optimal payoff is given by:

$$V_{ij0} = \max_{\mathcal{C} \in \mathbb{C}} \mathbb{E}_0 \left[ \int_0^T e^{-rt} dX_{ijt} - \int_0^T e^{-rt} w_{ijt} dt + \lambda \int_0^T e^{-rt} V_T dt + e^{-rT} (V_T - b_j \pi) \right] \quad (11)$$

The board maximizes over the space of admissible contracts  $\mathbb{C}$ , where a contract  $\mathcal{C} = (w, a, T)$  is a triple specifying a compensation process  $w$ , action process  $a$ , and stopping time  $T$ , all of which are  $\mathcal{F}_{ijt}$ -adapted. Upon turnover, the board receives value  $V_T$  which simply denotes the value of the next CEO employment spell:

$$V_T \equiv V_{(i+1)j0} \quad (12)$$

Following any instance of turnover, the board immediately draws a successor CEO ( $i+1$ ) from distribution  $N(\theta_0, \delta_0^2)$  and continues operations. Note that given the initial condition for profitability (6), the monetary turnover cost  $c$  is reflected in the term  $V_{(i+1)j0}$ .

---

<sup>32</sup>Throughout the model,  $\mathbb{E}_t^a[x] = \int_{\Omega} x dP_t^a$ , where  $P_t^a$  is the tenure- $t$  probability measure arising from having observed the action process  $a$ . On the other hand,  $\mathbb{E}_t[x] = \int_{\Omega} x dP_t$ , where  $P_t$  is the tenure- $t$  probability measure having not observed the action process. Given that the CEO observes  $a$  and the firm does not, the two parties will condition their expectations on different information.



Additionally, in the case of firing the CEO, reflected by the last term of (11), the board incurs non-pecuniary cost  $b_j\pi$ . This represents CEO entrenchment, which raises the effective cost of forceful CEO replacement. Hence at stopping time  $T$ , the board incurs cost  $b_j(\pi+c)$  and continues operations with new CEO  $i+1$ . Note here that the board's problem is stationary and the turnover value  $V_T$ , the value of the subsequent CEO's employment, is independent of the current employment spell. When deriving the optimal contract,  $V_T$  can thus be treated as a constant.

**Learning** Though CEO quality is unknown, information about  $\theta_i$  is continuously generated as firm profitability is realized. The distributions  $N(\tilde{\theta}_{ijt}, \tilde{\delta}_{ijt})$  and  $N(\tilde{\theta}_{ijt}^a, \tilde{\delta}_{ijt}^a)$  represent the respective beliefs of the board and CEO given information up to  $t$ , where  $\tilde{\theta}_{ijt} = \tilde{\theta}_{ijt}^a$  in equilibrium. Off the equilibrium path however, given their private knowledge of the process  $a$ , the CEO may form beliefs which differ from the board's. Considering deviations in the board and CEO's beliefs if necessary for establishing incentive compatibility, the details of which I relegate to Appendix A.1. I assume rational expectations, so the initial beliefs of the board and CEO coincide with the population distribution:

$$\tilde{\theta}_{ij0} = \tilde{\theta}_{ij0}^a = \theta_0 \quad (13)$$

$$\tilde{\delta}_{ij0}^2 = \delta_0^2 \quad (14)$$

As CEO tenure increases, beliefs adjust in response to realized performance. In particular, given the normality of both  $\theta_i$  and  $W_{ijt}$ , the posterior mean will be an increasing, linear function of cumulative performance. Define cumulative performance  $Y_{ijt}$  and cumulative action  $A_{ijt}$  by:

$$Y_{ijt} = Y_{ij0} + \int_0^t (\theta_i - a_{ijs}) ds + \sigma_W \int_0^t dW_{ijs} \quad (15)$$

$$A_{ijt} = \int_0^t a_{ijs} ds \quad (16)$$

The parameters of the equilibrium belief distribution  $N(\tilde{\theta}_{ijt}, \tilde{\delta}_{ijt}^2)$  are then given by:

$$\tilde{\theta}_{ijt} = \frac{\delta_0^{-2}\theta_0 + \sigma^{-2}(Y_{ijt} - Y_{ij0} + A_{ijt})}{\tilde{\delta}_{ijt}^{-2}} \quad (17)$$

$$\tilde{\delta}_{ijt}^2 = (\delta_0^{-2} + \sigma^{-2}t)^{-1} \quad (18)$$

Beliefs depend only on the CEO's cumulative contribution to profitability  $Y_{ijt}$ , as the firm productivity component of profitability  $\gamma_j$  is independent of the CEO. Additionally, beliefs are conditioned on the CEO's cumulative action  $A_{ijt}$ . In equilibrium, the CEO will always pick the board's recommended action, in which case the board correctly infers  $A_{ijt}$  and shares the same estimate of  $\theta_i$  as the CEO. The equilibrium law of motion for  $\tilde{\theta}_{ijt}$  follows from Ito's lemma:

$$d\tilde{\theta}_{ijt} = \frac{\tilde{\delta}_{ijt}^2}{\sigma_W^2} (dY_{ijt} - (\tilde{\theta}_{ijt} - a_{ijt}^*)dt) \quad (19)$$

$$= v_{ijt}\sigma_W dZ_{ijt} \quad (20)$$

where  $dZ_{ijt} = \sigma_W^{-1}(dY_{ijt} - (\tilde{\theta}_{ijt} - a_{ijt}^*)dt)$  is the *innovation process*, tracking the realization of performance  $dY_{ijt}$  net of expectations  $(\tilde{\theta}_{ijt} - a_{ijt}^*)dt$ . Beliefs adjust in response to these signals with sensitivity  $v_{ijt} = \tilde{\delta}_{ijt}^2/\sigma_W^2$ , which I define as the rate of learning. The posterior variance is deterministic and decreases monotonically with  $t$ :

$$d\tilde{\delta}_{ijt}^2 = -v_{ijt}\tilde{\delta}_{ijt}^2 dt \quad (21)$$

Note that the variance of beliefs depends only on tenure, so is unaffected by CEO action choices and hence will be the same on or off the equilibrium path for a given  $t$ .

The posterior mean for the board and CEO on the other hand will in general not coincide off the equilibrium path. In particular, deviations from the efficient action  $a^*$  will lead boards to misinterpret the realized signal of quality. Relative to the board's expectations, cash flow diversion induces a low realization of contemporaneous performance.<sup>33</sup> This leads the board's beliefs to drift downwards relative to the CEO's, inducing a gap in expectations about future performance.<sup>34</sup> This is reminiscent of the ratchet effect as discussed by Laffont and Tirole (1988). The CEO benefits from conveying negative information to the board, as it eases their future incentive load. To prevent this in equilibrium, the board compensates the CEO via an information rent. Additionally, the CEO's incentive to convey negative information is limited by the risk of termination. If the boards' beliefs fall too low, the CEO will be fired.

<sup>33</sup>Given that the action process  $a^*$  attains its lower bound in equilibrium, it suffices to restrict attention to positive deviations. See Appendix B.2 for more discussion.

<sup>34</sup>CEOs' capacity to manipulate boards' beliefs about their quality has received empirical attention in the accounting literature. For example, Ali and Zhang (2015) note that CEOs' incentive to manipulate earnings reports is greatest in early years of tenure when there is relatively more uncertainty about quality, as is the case in the model presented here. More recently, evidence has been presented suggesting that the pervasiveness of such strategic earnings management is minimal (Bertomeu et al., 2021). This supports the notion that the "no manipulation" equilibrium considered here is a reasonable approximation of reality.

**CEO Turnover** CEOs employment can end either through endogenous termination or exogenous retirement, where retirement shocks arrive at rate  $\lambda$ . Conditional on a retirement shock, the board immediately draws a replacement CEO at cost  $c$  and continues operations. Firing occurs when the board's beliefs  $\tilde{\theta}_{ijt}$  drop below the endogenous threshold  $\theta_f(t)$ . Let  $V(\tilde{\theta}_{ijt}, t, U_{ijt})$  denote the board's optimal payoff given state  $(\tilde{\theta}_{ijt}, t, U_{ijt})$ , where  $U_{ijt}$  denotes the CEO's promised equilibrium payoff.  $\theta_f(t)$  is defined as the level of  $\tilde{\theta}_{ijt}$  such that the board is indifferent between continuing with CEO  $i$  and drawing a new CEO ( $i + 1$ ):

$$V(\theta_f(t), t, U_{ijt}) = V_T - b_j\pi \quad (22)$$

The stopping time  $T$  denotes the first time that  $\tilde{\theta}_{ijt}$  reaches  $\theta_f(t)$ . Concretely:

$$T = \inf\{t < \infty \mid \tilde{\theta}_{ijt} = \theta_f(t)\} \quad (23)$$

**Board's Problem** The optimal contract  $\mathcal{C} = (w, a, t)$  delivers the CEO a payoff of  $U_{ij0}$  and maximizes the board's objective:

$$\begin{aligned} V(\tilde{\theta}_{ijt}, t, U_{ijt}) = \mathbb{E}_t \left[ \int_t^T e^{-r(s-t)} dX_{ijs} - \int_t^T e^{-r(s-t)} w_{ijs} ds \right. \\ \left. + \lambda \int_t^T e^{-r(s-t)} V_T ds + e^{-r(T-t)} (V_T - b_j\pi) \right] \quad (24) \end{aligned}$$

subject to:

$$U_{ij0} = \mathbb{E}_0 \left[ \int_0^T e^{-rt} u(w_{ijt}, a_{ijt}) dt + \lambda \int_0^T e^{-rt} C(\tilde{\theta}_{ijt}) dt + e^{-rT} C(\tilde{\theta}_{ijT}) \right] \quad (\text{PK})$$

$$U_{ijt} \geq C(\tilde{\theta}_{ijt}) \quad \forall t \leq T \quad (\text{IR})$$

$$U_{ij0} \geq U_{ij0}^{\hat{a}} \quad \text{for any other } \hat{a} \quad (\text{IC})$$

(PK) simply defines the CEO's promised value. (IR) and (IC) are the participation and incentive-compatibility constraints, respectively. Rather than directly using (IC), I use a first-order approach following Williams (2011) and derive the CEO's first-order-incentive-compatibility condition (FOIC).

### 3.1 First-Best Case

Before deriving the optimal contract, it is useful to analyze the first-best case. Actions  $a_{ijt}$  are observable and the board ensures the first-best action  $a_{ijt}^* = 0$  is selected for all  $t$ . This

lack of information asymmetry implies that the board and CEO's beliefs are identical (i.e.  $\tilde{\theta}_{ijt} = \tilde{\theta}_{ijt}^a$ ). Here, the board solves a pure optimal stopping problem, monitoring performance and determining when to fire and replace their current CEO. Define  $\theta_f^{FB}(t)$  as the first-best firing threshold, which is characterized in the following proposition.

**Proposition 1** (First-Best Firing Threshold). *When actions are observable, the firing threshold which maximizes the board's payoff is given by:*

$$\theta_f^{FB}(t) = -r\pi + b_j^{-1} \left( \rho V_T - \frac{v_{ijt}^2 \sigma_W^2}{2} V_{\theta\theta}(\theta_f^{FB}(t), t) \right) \quad (25)$$

*Proof.* See Appendix A.1. □

Condition (25) implicitly defines the optimal firing boundary.<sup>35</sup> In the case of no entrenchment,  $\pi = 0$  and (25) is shareholder-optimal in the sense that it maximizes firm value. When  $\pi > 0$ , a wedge is induced which separates firm value from the board's optimal payoff  $V$ . Thus, in the presence of entrenchment, the board's enacted firing rule does not coincide with the shareholder-optimal firing rule. In particular, when CEOs are entrenched, the board retains some CEOs which shareholders would have preferred to see terminated ex-post. The same is true in the second-best case with unobservable CEO actions, which I consider next.

### 3.2 Second-Best Case

To characterize incentives in the second-best case, I apply the stochastic maximum principle to derive necessary conditions for the incentive-compatibility of the efficient action path (Bismut, 1973; Williams, 2011; Demarzo and Sannikov, 2017). The necessary conditions are summarized in the following theorem.

**Proposition 2** (Necessary Conditions). *Under the efficient strategy  $a^*$ , the CEO's promised value has equilibrium law of motion:*

$$dU_{ijt} = \left( rU_{ijt} - w_{ijt} - \lambda C(\tilde{\theta}_{ijt}) \right) dt + \beta_{ijt} \sigma_W dZ_{ijt} \quad (26)$$

where  $\beta_{ijt}$  is a sensitivity process representing the incentives provided by the contract. The efficient strategy  $a^*$  is incentive compatible if:

$$\beta_{ijt} \geq v_{ijt} \Gamma_{ijt} + \phi_j \quad (27)$$

---

<sup>35</sup>Obtaining a closed-form representation of  $\theta_f^{FB}(t)$  is not feasible, but it can be computed numerically. See Appendix B.3 for a detailed exposition of the numerical solution.

where  $\Gamma_{ijt} = \frac{\partial U_t^a}{\partial \theta^a}$  is the CEO's information rent, the benefit of having marginally more optimistic beliefs relative to the board.  $\Gamma_{ijt}$  has equilibrium lower bound:

$$\Gamma_{ijt} \geq \frac{\phi_j}{r} \equiv \Gamma_{ijt}^* \quad (28)$$

which holds with equality when (27) binds.

*Proof.* See Appendix A.1. □

Equation (26) is the standard representation of the CEO's promised value in continuous time, showing that  $U_{ijt}$  is an Ito process with respect to the standard Brownian Motion  $Z_{ijt}$ .  $U_{ijt}$  has drift  $(rU_{ijt} - w_{ijt} - \lambda C(\tilde{\theta}_{ijt}))$ , stating that the CEO's promised value accumulates at rate  $r$  net of the CEO's expected payoff  $(w_{ijt} + \lambda C(\tilde{\theta}_{ijt}))$ .<sup>36</sup> Promised value has volatility  $\beta_{ijt}\sigma_W$ , measuring the sensitivity of the CEO's payoff to performance innovations  $dZ_{ijt}$ . The sensitivity process  $\beta_{ijt}$  is chosen implicitly by the board, and is the key instrument through which incentives are delivered to the CEO.

Condition (27) is the first-order counterpart of the incentive compatibility constraints (IC). The above proposition, however, says little about the participation constraint (IR). If the CEO's promised value falls below their outside option  $C(\tilde{\theta}_{ijt})$ , it is optimal for the CEO to leave the firm. The optimal compensation process ensures that  $U_{ijT} = C(\tilde{\theta}_{ijT})$  while  $U_{ijt} > C(\tilde{\theta}_{ijt})$  for all  $t < T$ , so CEO departure only occurs through firing when it is optimal for the board. Condition (27) implies that  $U_{ijt}$  changes with  $\tilde{\theta}_{ijt}$  according to:

$$\frac{dU_{ijt}}{d\tilde{\theta}_{ijt}} = \frac{dU_{ijt}}{dY_{ijt}} \left( \frac{d\tilde{\theta}_{ijt}}{dY_{ijt}} \right)^{-1} = \frac{\beta_{ijt}}{v_{ijt}} \geq \frac{\phi_j}{v_{ijt}} + \Gamma_{ijt} \quad (29)$$

The condition (29) paired with the bound on the CEO's information rent (28) implies a lower bound on the CEO's continuation payoff  $U_{ijt}$  in any optimal contract. For a given

---

<sup>36</sup>Speaking heuristically, prior to the realization of the retirement shock, the CEO's expected payoff over interval  $\Delta t$  is:

$$e^{\lambda\Delta t} w_{ijt} + \lambda e^{\lambda\Delta t} C(\tilde{\theta}_{ijt}) + o(\Delta t)$$

The first two terms respectively represent the probability of zero and one retirement shocks arriving over interval  $\Delta t$ . The probability of  $> 1$  shocks arriving is negligible, represented by the third term. Taking the limit as  $\Delta t \rightarrow 0$  yields  $w_{ijt} + \lambda C(\tilde{\theta}_{ijt})$ .

level of tenure  $t$  and termination boundary  $\theta_f(t)$ ,  $U_{ijt} \equiv U(\tilde{\theta}_{ijt}, t, \theta_f(t))$  must exceed:

$$U_{ijt} = C(\theta_f(t)) + \int_{\theta_f(t)}^{\tilde{\theta}_{ijt}} \frac{dU_t}{d\tilde{\theta}_{ijt}} d\theta \quad (30)$$

$$\geq C(\theta_f(t)) + (\tilde{\theta}_{ijt} - \theta_f(t)) \frac{\phi_j}{v_{ijt}} + \int_{\theta_f(t)}^{\tilde{\theta}_{ijt}} \Gamma_{ijt} d\theta \quad (31)$$

$$\geq C(\theta_f(t)) + (\tilde{\theta}_{ijt} - \theta_f(t)) \left( \frac{\phi_j}{v_{ijt}} + \frac{\phi_j}{r} \right) \equiv U_{ijt}^* \quad (32)$$

The first inequality comes from (29), while the second comes from (28).  $U_{ijt}^*$  is thus the lower bound of the CEO's promised utility in any incentive-compatible contract implementing  $a_{ijt}^* = 0$  for all  $t$ . The board seeks a compensation process  $w$  which minimizes  $U_{ijt}$  subject to the constraint  $U_{ijt} \geq U_{ijt}^*$ . Given this constraint, the following proposition summarizes the cost-minimizing compensation process.

**Proposition 3** (Wage Determination). *The firm's relaxed problem can be stated as follows. The board offers value  $U_{ijt}$  with volatility  $\beta_{ijt}\sigma_W$  such that  $U_{ijt} \geq C(\tilde{\theta}_{ijt})$  for all  $t \leq T$ , holding with equality only when  $t = T$ . The incentive-compatibility constraint (27) implies a lower bound on  $U_{ijt}$  in equilibrium:*

$$U_{ijt} \geq C(\theta_f(t)) + (\tilde{\theta}_{ijt} - \theta_f(t)) \left( \frac{\phi_j}{v_{ijt}} + \frac{\phi_j}{r} \right) \equiv U_{ijt}^* \quad (33)$$

Thus, the board maximizes (24) subject to  $U_{ijt} \geq U_{ijt}^*$  and participation constraint (IR) for all  $t$ . Given termination boundary  $\theta_f(t)$ , the cost-minimizing compensation process is given by:

$$w_{ijt} = \rho\mu + \kappa_{1t}\tilde{\theta}_{ijt} + \kappa_{2t}\theta_f(t) \quad (34)$$

$$\kappa_{1t} = \phi_j \left( \frac{\rho}{r} + \frac{r}{v_{ijt}} + \frac{v_{ijt}}{r} \right) \quad (35)$$

$$\kappa_{2t} = \phi_j \left( 1 - \frac{r}{v_{ijt}} \right) \quad (36)$$

*Proof.* See Appendix A.1. □

The compensation process above ensures that the constraint  $U_{ijt} \geq U_{ijt}^*$  binds for all  $t$ . This implies that it is in the CEO's interest to commit themselves to the efficient action process for the duration of their employment and rules out premature endogenous CEO departures. Given the dependence of compensation on the termination boundary  $\theta_f(t)$ , the replacement cost parameters will have direct effects on the provision of CEO pay, the

magnitude of which will be determined empirically. To close the model, I summarize the second-best termination boundary in the proposition below.

**Proposition 4** (Second-Best Firing Threshold). *When CEO actions are unobservable, the firing threshold which maximizes the firm's payoff is given by:*

$$\theta_f(t) = -r\pi + b_j^{-1} \left( \rho V_T - \frac{v_{ijt}^2 \sigma_W^2}{2} V_{\theta\theta}(\theta_f(t), t, C(\theta_f(t))) - \frac{\beta_{ijt}^2 \sigma_W^2}{2} V_{UU}(\theta_f(t), t, C(\theta_f(t))) \right) \quad (37)$$

*Proof.* See Appendix A.1. □

The optimal firing threshold is derived by applying standard smooth-pasting and value-matching conditions to the board's HJB equation. The second-best threshold retains many of the same features as the first-best counterpart. The key difference arises from the inclusion of the CEO's continuation payoff in the state. Given the CEO's limited liability, firms have limited ability to punish their CEO through financial means. Termination serves as an alternative method of punishment. In particular, if CEOs' continuation payoff falls to their outside option  $C(\tilde{\theta}_{ijt})$ , the board optimally fires and replaces the CEO. Thus, poor performance gradually drives down CEOs' continuation payoff, and in extreme cases results in the termination of their employment.

## 4 Identification and Estimation

I fix the value of the discount rate  $\rho = .05$  and estimate the remaining 9 parameters using the Simulated Method of Moments. I first discuss the identification of the parameters  $\sigma_W$ ,  $\sigma_\gamma$ ,  $\theta_0$ , and  $c$ , whose key identifying information comes from firm profitability data. Under the assumption that both CEO quality  $\theta_i$  and firm productivity  $\gamma_j$  are time-invariant, variation in ROA within a given employment spell is generated entirely by idiosyncratic shocks. Thus, within-spell variation in ROA helps to identify  $\sigma_W$ . Across-firm variation in ROA on the other hand is informative of  $\sigma_\gamma$ . Let  $y_{ijt}$  denote observed ROA, and let  $\mathbb{E}_j$  and  $Var_j$  denote the mean and variance operators conditioned on firm  $j$ . To separately identify  $\sigma_\gamma$  and  $\sigma_W$ , I target the following two moments:

$$Var(\mathbb{E}_j[y_{ijt}]) \quad (38)$$

$$\mathbb{E}[Var_j(y_{ijt})] \quad (39)$$

The moment (38) is the across-firm variance of the within-firm average of ROA, and carries information about the standard deviation of firm productivity  $\sigma_\gamma$ . The moment (39) is the within-firm variance of  $y_{ijt}$  averaged across all firms, which informs the estimate

of the idiosyncratic volatility  $\sigma_W$ . The mean of match quality  $\theta_0$  affects the mean of profitability, so is pinned down by the empirical average of ROA. The monetary cost of turnover, measured by  $c$ , is identified by variation in ROA around episodes of CEO turnover (forced or voluntary).

The parameters  $\pi$  and  $\lambda$  are identified off of CEO turnover data. The probability of forced termination strictly decreases in the non-pecuniary cost of forced turnover  $\pi$ . The empirical forced termination rate thus carries information about the level of entrenchment. With this in mind, I target the coefficients of the auxiliary model:

$$d_{ijt} = \lambda_0 + \lambda_1 \textit{tenure}_{ijt} + \lambda_2 \textit{tenure}_{ijt}^2 + \xi_{ijt} \quad (40)$$

where  $d_{ijt} \in \{0, 1\}$  is an indicator for forced turnover. The arrival rate of retirement shocks  $\lambda$  is recovered by simply matching the simulated and empirical rates of CEO retirement.

The parameters  $\mu$ ,  $\alpha$ , and  $\delta_0$  are identified off of the compensation data. Equation (34) shows that the unconditional expectation of model-implied compensation increases with  $\mu$ . This estimate of this parameter therefore informed by the empirical average of compensation. Additionally, compensation varies with firm size through the specification of the cash flow diversion rate (8). Under this parameterization, targeting the empirical correlation between compensation and firm size is sufficient for recovering the parameter  $\alpha$ . Specifically, I match the coefficients of the following auxiliary model:

$$\log(w_{ijt}) = \beta_0 + \beta_1 \log(\textit{assets}_{ijt}) + \beta_2 t_{ij} + \epsilon_{ijt} \quad (41)$$

where the auxiliary parameters  $\beta_0$  and  $\beta_1$  are informative of  $\mu$  and  $\alpha$ , respectively. Finally, the compensation process (34) implies that conditional on recovering the parameters  $\alpha$ ,  $\sigma_W$ , and  $\lambda$ , the sensitivity of compensation to CEO performance is pinned down by the parameter  $\delta_0$ . I thus match the simulated and empirical pay-performance sensitivities summarized in the simple auxiliary model:

$$\Delta \log(w_{ijt}) = \beta_0 + \beta_1 y_{ijt} + \eta_{ijt} \quad (42)$$

$\Delta \log(w_{ijt})$  is year over year growth and CEO pay and the coefficient  $\beta_1$  measures the sensitivity of pay to performance. In total, I estimate the model using a  $13 \times 1$  vector of moments denoted by  $\hat{M}$ . I obtain the optimal weighting matrix as the inverse of the covariance matrix of  $\hat{M}$ .<sup>37</sup>

---

<sup>37</sup>See Appendix B.4 for details on the computation of the optimal weighting matrix, estimation algorithm, and standard error computation.



The estimation algorithm proceeds as follows. Let  $\Theta \in \mathbb{R}^9$  denote an arbitrary vector of structural parameters. Given  $\Theta$ , I obtain the value function  $V$  by numerically solving the board’s HJB equation, from which the optimal termination boundary and compensation process can be computed. Given the optimal policies, I simulate 5000 firms 20 times each. Firms draw an initial CEO, and thereafter performance, beliefs, turnover, and compensation evolve as specified in the previous section. The simulation proceeds for 50 periods, where a period corresponds to a calendar year. Using the simulated data, I compute the same 13 moments as were computed in the empirical sample. If the simulated moments are sufficiently close to their empirical counterparts, the algorithm halts and returns the estimate  $\hat{\Theta}$ . Otherwise, a new candidate parameter vector is chosen and the procedure repeats. Standard errors for the estimates are computed based upon the asymptotic distribution of the SMM estimator as presented by Duffie and Singleton (1993). Details on model fit can be found in Appendix A.2.

## 5 Results

Parameter estimates and standard errors are reported in Table 3. The estimated standard deviation of the CEO quality distribution ( $\sigma_\theta$ ) is .021, or 2.1% of total assets per year, which is comparable to what has previously been found in the literature. Taylor (2010) for example reports an estimate of 2.4%. Interpreting this estimate in terms of dollars: a one standard deviation increase in CEO quality implies a \$46.6 million dollar increase in average yearly cash flows for the median-sized firm in the sample.<sup>38</sup> The estimate of  $\sigma_\gamma$ , the standard deviation of firm productivity, is significantly higher; a one standard deviation increase in productivity yields a \$228.6 million increase in yearly cash flows for a median-sized firm. Using the volatility estimates  $\sigma_\theta$ ,  $\sigma_\gamma$ , and  $\sigma_W$ , I decompose the variance of firm profitability: 2.9% and 68.0% of the variation of profitability can be respectively attributed to variation in CEO quality and firm productivity. The remaining 29.1% is attributed to idiosyncratic variation orthogonal to CEO and firm characteristics.<sup>39</sup>

The above calculation may, at first glance, seem to suggest that the impact of CEO quality on firm performance is negligible. This is not the case; from shareholders’ perspective, whether an incumbent CEO is of high or low quality is paramount. To see this,

<sup>38</sup>The median firm in the sample has approximately \$2.2 billion in assets.

<sup>39</sup>Interpreting the increment  $dy_{ijt}$  as firm ROA, under the assumption that CEO quality  $\theta_i$  and firm productivity  $\gamma_j$  are independent we have that  $Var(dy_{ijt}) = (\sigma_\theta^2 + \sigma_\gamma^2 + \sigma_W^2)dt$ . The shares of the variation in ROA attributed to each component  $k \in \{\theta, \gamma, W\}$  are then computed as  $\sigma_k^2 / (\sigma_\theta^2 + \sigma_\gamma^2 + \sigma_W^2)$ .

Table 3: Structural Estimates

$\theta_0$	$\delta_0$	$\sigma_W$	$\sigma_\gamma$	$\lambda$	$\alpha$	$\mu$	$c$	$\pi$
.117	.021	.067	.103	.088	.467	12.0	.007	.037
( $6.9e^{-4}$ )	( $9.1e^{-4}$ )	( $2.1e^{-4}$ )	( $1.4e^{-4}$ )	(.001)	(.001)	(.278)	(.002)	(.004)

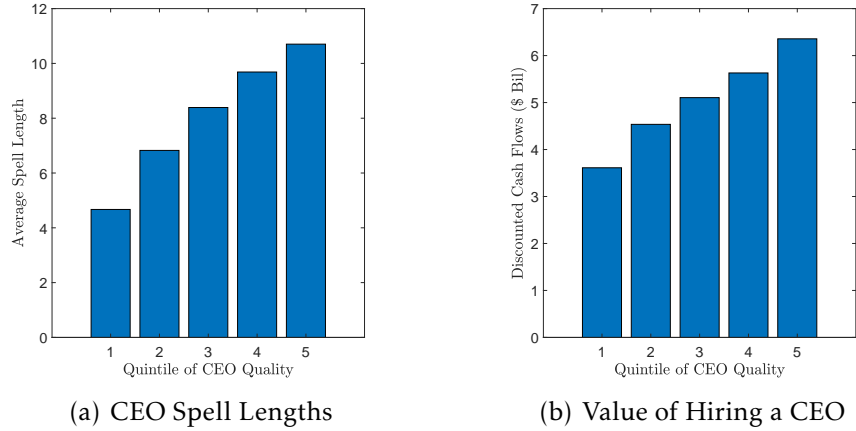
Notes: Standard errors are included in parenthesis.

first note that the implied difference in profitability between a CEO at the 5th and 95th percentile of the quality distribution is 7% per year, or \$155 million in cash flows per year for the median-sized firm in the sample.<sup>40</sup> Furthermore, this gap in performance accumulates as CEO spells progress. To get a sense of how this gap is amplified over CEO careers, note first that the average CEO spell length in the sample is roughly 8.3 years (Table 1). With model estimates in hand, I can compute this average separately for CEOs across the distribution of quality. In Figure 5(a), I plot the average spell length for CEOs within 5 quintiles of the estimated quality distribution. As expected, CEOs of relatively low levels of quality see shorter average spell lengths than their high-quality counterparts, highlighting some degree of proficiency in boards' screening faculties. Given these spell lengths, Figure 5(b) plots the implied level of expected discounted cash flows over a CEO spell for a median-sized firm, conditional on hiring the median CEO in each bin of quality. For example, the median CEO generates roughly \$5.1 billion in cash flows over their career (if employed by a median-sized firm). For a CEO in the top quality quintile, this increases by roughly \$1.2 billion. Thus, from the perspective of forward-looking shareholders, the difference between a high and low-quality CEO is in the order of billions of dollars.

Clearly, it is in shareholders' interest to have low-quality CEOs promptly replaced. The speed at which this happens is dependent on the effective cost of CEO turnover and boards' ability to make inferences about CEO quality. Recall that the effective cost of forced CEO replacement is given by  $c + \pi$ . The estimated monetary cost of turnover ( $c$ ) is 0.7% of total firm assets which amounts to roughly \$15.1 million for the median-sized firm. This cost reflects expenses associated with finding a replacement CEO, severance payouts, and general disruptions to profitability resulting from onboarding new personnel. Compared to the monetary cost  $c$ , boards' estimated utility cost of forced turnover ( $\pi$ ) is significantly higher. While turnover entails a monetary cost of 0.7% of total assets, boards behave as if turnover costs  $c + \pi \times 100 = 4.4\%$  of total assets. This induces a wedge separating the shareholder-optimal firing policy from boards' enacted policy. Namely, when CEOs are entrenched, the rate of CEO termination is lower than share-

<sup>40</sup>The implied difference in profitability is obtained using the formula  $2 \times 1.65 \times \delta_0 = 9.24\%$

Figure 4: CEO Quality, Spell Lengths, and Discounted Cash Flows



Notes: Figure 5(a) plots the average spell lengths for CEOs across the five quintiles of the CEO quality distribution. Given these spell lengths, Figure 5(b) approximates the discounted cash flows expected over a CEO spell given that quality  $\theta_i$  lies on the midpoint of the respective quintile (i.e. the 10<sup>th</sup>, 30<sup>th</sup>, 50<sup>th</sup>, 70<sup>th</sup>, and 90<sup>th</sup> percentiles). The net present value of cash flows is computed as:

$$(1 - e^{-\hat{T}_q p_q})$$

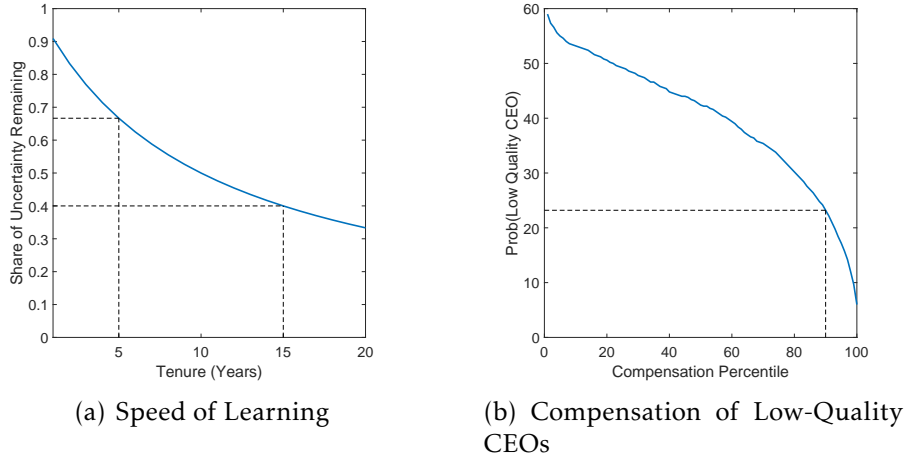
where  $\hat{T}_q$  is the average spell length of a CEO in quintile  $q \in \{1, 2, 3, 4, 5\}$  and  $p_q$  is the value of  $\theta_i$  which lies in the midpoint of each quintile.

holders would prefer.

In addition, board firing decisions depend on their quality of information at any given point in time. Given the high degree of uncertainty over CEO quality, it may be difficult for boards to correctly infer that their CEO is of low quality. As not to erroneously fire a high-quality CEO, boards may prefer to defer their option to replace until they have acquired a sufficient amount of information pertaining to their CEO. The rate of information acquisition is governed here by boards' speed of learning, given analytically by the signal to noise ratio  $v_{ijt} = \tilde{\delta}_{ijt}^2 / \sigma_W^2$ . The high estimate of  $\sigma_W$  relative to  $\delta_0$  implies that boards learn quite slowly. To illustrate this, I plot in Figure 6(a) the share of quality uncertainty remaining over the first 20 years of CEO tenure. Roughly 66% and 40% of the quality uncertainty remains after 5 and 15 years of tenure, respectively. The low rate of information acquisition decreases boards' willingness to commit to costly turnover, further increasing the employment lengths of low-quality CEOs.

As boards have limited ability to disentangle the effects of their CEO and exogenous shocks on firm performance, some degree of pay for luck may be expected in equilibrium. As beliefs adjust, either upwards or downwards, compensation adjusts as a result. Given the normality of firm beliefs and performance, the magnitude of compensation adjustment is the same in response to positive or negative performance shocks of the same magnitude. Thus, pay for luck is symmetric, as documented empirically by Daniel et al. (2019). To quantify the degree of pay for luck in equilibrium, Figure 6(b) plots, across all percentiles of the compensation distribution, the probability that a CEO has quality

Figure 5: Information Acquisition and Pay for Luck



Notes: Panel (a) plots the share of uncertainty remaining over the first 15 years of CEO tenure. I define the share of quality uncertainty remaining at tenure  $t$  as  $(\delta_{ijt}/\delta_0)^2$ . Panel (b) plots the simulated probability that within a given percentile of the cross-sectional compensation distribution, a CEO is of low quality. A CEO is referred to as low quality if and only if  $\theta_i < \theta_0$ .

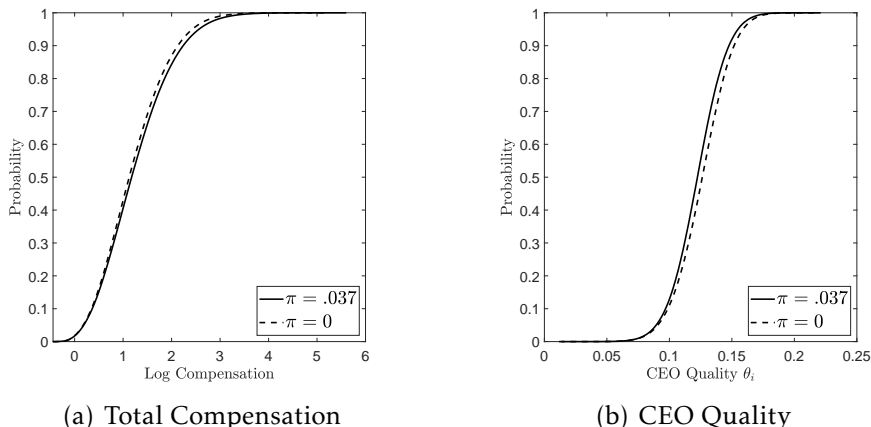
below that of the population mean  $\theta_0$ . The Figure suggests that the degree of pay for luck is substantial. As an example, among CEOs with compensation in the 90<sup>th</sup> percentile, roughly one in four of them has quality below that of the average replacement. The high degree of pay for luck can be attributed to the substantial amount of noise present in the board’s learning process.

## 6 Entrenchment and the CEO Pay Premium

A primary purpose of this paper is to analyze the impact of entrenchment on CEO compensation. As is typical in models of moral hazard, by virtue of the asymmetry in information between board and CEO, boards set equilibrium compensation above CEOs’ outside option in order to achieve incentive alignment. I illustrate in this section that entrenchment weakens CEO incentives, increasing the board’s cost of incentive provision. This result is explained by straightforward economic intuition. When it comes to incentivizing the CEO, boards have two levers at their disposal: the “carrot” of performance-based pay and the “stick” of forced termination. Entrenchment increases the cost of forced termination thereby increasing boards’ utilization of the carrot relative to the stick in equilibrium. Thus, we should expect the CEOs’ rents to be larger in the presence of entrenchment relative to a counterfactual environment with no entrenchment. The model confirms this intuition; counterfactual simulations reveal a 9.1% decrease in the average

rate of CEO compensation upon the elimination of entrenchment.<sup>41</sup> Figure 7(a) compares the distributions of total CEO compensation with and without entrenchment.

Figure 6: Simulated CDFs of Total Compensation and CEO Quality



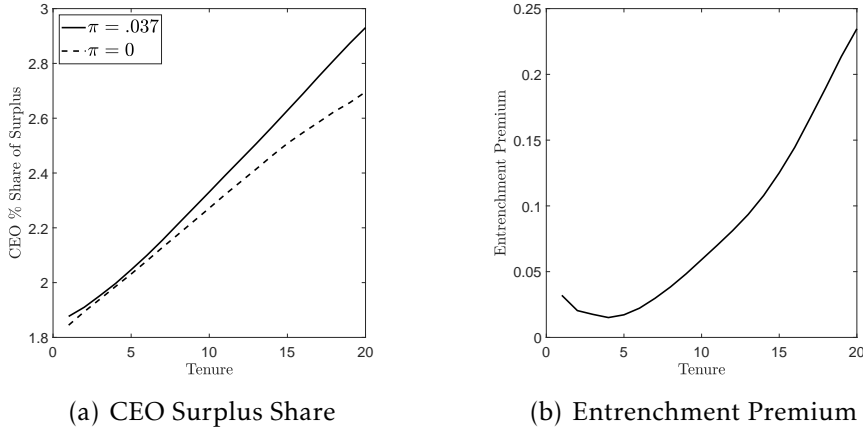
Notes: Figure 7(a) plots the cumulative distribution of CEO compensation with and without entrenchment. Eliminating entrenchment (by setting  $\pi = 0$ ) induces a leftward shift in the compensation distribution. This shift is the product of both the direct impact of entrenchment on the level of CEO pay, and a rightward shift in distribution of  $\theta_i$  among retained CEOs (Figure 7(b)).

Eliminating entrenchment induces a leftward shift in the compensation distribution. When faced with a higher risk of termination following poor performance, the level of compensation necessary to induce efficient CEO behavior decreases. This shift in the aggregate pay distribution, however, does not accurately reflect the direct effect of entrenchment on compensation. As illustrated in Figure 7(b), which plots the distribution of CEO quality conditional on retention, eliminating entrenchment induces a rightward shift in the conditional distribution of  $\theta_i$  following from more aggressive screening on CEO quality. All else equal, this has a positive effect on the average level of compensation. The change in compensation shown in Figure 7(a) is thus dampened by this shift in the distribution of quality among retained CEOs. To isolate the direct effect of entrenchment, I plot in Figure 8(a) average CEO compensation as a fraction of the CEO's perceived contribution to output (i.e. their share of total surplus) with and without CEO entrenchment. CEO surplus is strictly larger in the presence of entrenchment; boards pay an additional premium to offset CEOs' adverse incentive effects stemming from their decreased risk of performance-induced termination.

I refer to the gap between the two lines as the *entrenchment premium*, which is the excess surplus extracted by the CEO when they are entrenched relative to the counterfactual case with no entrenchment. Notably, the entrenchment premium rises with tenure. Given the positive correlation between CEO tenure and quality (Figure 5(a)), this implies

<sup>41</sup>Entrenchment is eliminated by setting the value of the parameter  $\pi = 0$ .

Figure 7: Entrenchment, Tenure, and CEO Rent Extraction



Notes: Figure 8(a) plots the average share of surplus extracted by the CEO over the first 20 years of tenure, separately in the estimated model (in which  $\pi = .037$ ) and a counterfactual version of the model in which  $\pi = 0$ . The CEO's share of surplus is defined as the ratio of their compensation to their average output:  $w_{ijt}/\theta_{ijt}$ . Entrenchment, as a result of its adverse effects on CEO incentives, increases the cost of incentivizing the CEO, therefore increasing the CEO's share of surplus. The difference between these two lines, which I for brevity refer to as the *entrenchment premium*, is plotted in Figure 8(b).

that per unit of output, high-quality CEOs on average extract more surplus than their low-quality counterparts. Thus, when CEOs are entrenched, the benefits associated with hiring a high-quality CEO are not fully passed through to shareholders and are instead partially retained by the CEO. The heterogeneous effects of entrenchment on CEO pay across the quality distribution are further illustrated in Table 4, which decomposes total compensation for CEOs across the five quintiles of the estimated quality distribution. For each quintile, I calculate both the level and share to total compensation associated with four of its key determinants: outside options, asymmetric information, entrenchment, and the monetary cost of turnover. This decomposition is carried out in four steps. First, I calculate average pay in the first-best environment with no informational asymmetry and costless turnover, giving the level of pay purely attributable to CEOs' outside options. Next, I add asymmetric information (i.e. moral hazard) to the contractual environment but keep turnover costs fixed to zero; this gives the marginal change in average compensation attributable to asymmetric information between board and CEO. Similarly, I add in (one at a time) entrenchment and monetary turnover costs to compute their marginal effects on average compensation.

Irrespective of the level of quality, outside options comprise the largest share of CEO pay. This constitutes the lower bound of compensation subject to retaining the CEO, which steadily increases with respect to CEO quality.<sup>42</sup> When adding asymmetric infor-

<sup>42</sup>Note that unlike standard models of moral hazard, equilibrium pay is not constant in the first-best case. This is due to the functional form of CEOs' outside options 9, under which outside options increase

Table 4: Composition of CEO Pay

	(1)	(2)	(3)	(4)	(5)
	Outside Option	Asymmetric Information	Entrenchment	Monetary Turnover Cost	Total Pay
<i>Quintile 1</i>					
Marginal Effect	\$2.89 M	\$388 K	\$45 K	\$122 K	<b>\$3.44 M</b>
Share of Total Pay	83.9%	11.3%	1.3%	3.5%	
<i>Quintile 2</i>					
Marginal Effect	\$2.95 M	\$602 K	\$254 K	\$157 K	<b>\$3.96 M</b>
Share of Total Pay	74.4%	15.2%	6.4%	4.0%	
<i>Quintile 3</i>					
Marginal Effect	\$3.00 M	\$813 K	\$471 K	\$195 K	<b>\$4.48 M</b>
Share of Total Pay	70.0%	18.1%	10.5%	4.4%	
<i>Quintile 4</i>					
Marginal Effect	\$3.07 M	\$1.16 M	\$683 K	\$217 K	<b>\$5.13 M</b>
Share of Total Pay	59.8%	22.6%	13.3%	4.2%	
<i>Quintile 5</i>					
Marginal Effect	\$3.20 M	\$2.02 M	\$907 K	\$249 K	<b>\$6.37 M</b>
Share of Total Pay	50.2%	31.6%	14.2%	3.9%	

Notes: Table 4 decomposes CEO compensation in four steps. To compute the marginal effect of outside options on total pay, I simulate the first-best case and additionally set both the monetary and non-monetary turnover costs ( $c$ ,  $\pi$ ) to zero. Here, equilibrium pay is set equilibrium pay to exactly compensate CEOs for their outside options. In step two, I introduce moral hazard (keeping  $c = \pi = 0$ ), in which case both the participation and incentive compatibility constraints must be satisfied. In step three, I set the entrenchment parameter  $\pi$  to its estimated value. Finally, I set the monetary turnover cost  $c$  to its estimate value in step four. In steps 2-4, the marginal effect on total pay is given by the difference in average compensation from the previous step. Summing the first four columns with each quintile gives total average compensation in the estimated model (Column 5). Shares of total pay are calculated as the marginal effect in each step divided by total pay.

mation, the incentive compatibility constraint requires satisfaction, in which case equilibrium compensation lies above CEOs' outside options. Observe that CEOs' gains from information asymmetry are heterogeneous; high-quality CEOs see by far the largest pay increases in response to moral hazard. This result is explained by differences in performance across the CEO quality distribution. Note first that in the presence of asymmetric information, compensation is tied to CEOs' performance histories. Under a performance-based pay scheme, the potential for rent extraction increases with respect to CEO performance, which on average is higher if a CEO is of high quality. In other words, consistent with Milbourn (2003), high-quality CEOs are generally awarded more performance-based pay than their low-quality counterparts, generating disparate effects of moral hazard on total compensation.

Entrenchment increases the incentive-aligning level of pay across all levels of CEO quality. Again, however, the magnitude of its effect is heterogeneous. As discussed above, entrenchment weakens termination-based incentives, increasing boards' utiliza-

with respect to public beliefs about CEO quality.

tion of performance pay in equilibrium. This disproportionately benefits high-quality CEOs given their relatively high levels of average performance; CEOs in the top quality quintile see nearly a roughly \$907 thousand dollar increase in average yearly compensation attributable to entrenchment. The monetary costs associated with turnover have a similarly adverse impact on CEO incentives, though its effects on compensation are relatively more stable across the quality distribution. Though they have similar effects on CEO compensation, entrenchment and monetary turnover costs have distinct effects on shareholder welfare. From shareholders' perspective, it is efficient for the boards acting on their behalf to consider monetary turnover costs when monitoring their CEOs and designing their incentive contracts, as monetary turnover costs have a direct impact on firm cash flows. Entrenchment, on the other hand, effectively induces a transfer of surplus from shareholders to the CEO. When  $\pi > 0$ , the rate of CEO termination is inefficiently low while the rate of CEO pay is inefficiently high, increasing CEO welfare at the expense of shareholders. In principle, these inefficiencies can be mitigated by a well-functioning board who neglects any private interests which may conflict with those of shareholders. In the next section, I examine empirically whether the degree of CEO entrenchment is sensitive to board characteristics, focusing specifically on the role of board independence.

## 7 Entrenchment and Board Independence

To further dissect the determinants of CEO entrenchment, I re-estimate the model allowing  $\pi$  to vary with board independence.<sup>43</sup> Director information comes from Institutional Shareholder Services (ISS), which covers a subset of firms in my full sample. For each firm in the sample, I compute the share of directors classified as independent. Note that this variable is missing for 65% of my full sample, which I accommodate for in the estimation procedure. As with the full sample, I numerically compute optimal termination and compensation policies and subsequently simulate 5000 firms 20 times each. For a given firm in the simulation, let the share of independent directors on the board be denoted by  $s_j$ . To mimic the data, I set this variable to "missing" for 65% of firms. Conditional on not being classified as missing, I draw a value of  $s_j$  from its empirical distribution and classify firms into one of two groups: high or low. "High" firms are those whose value of  $s_j$  lies above the empirical median of  $s_j$ , while "low" firms are those with values below the

---

<sup>43</sup>See Appendix B.1 for additional details on the sample construction and Appendix B.4 for additional details on the re-estimation procedure.



median.<sup>44</sup> Each group of firms is subject to their own respective entrenchment parameter  $\pi_{miss}$ ,  $\pi_{low}$ , and  $\pi_{high}$ , which are identified off of variation in termination rates across groups. As before, the remaining structural parameters are fixed across firms. I present the estimates of this extended model in Table 5.

Table 5: Structural Estimates - Entrenchment and Board Independence

$\theta_0$	$\delta_0$	$\sigma_W$	$\sigma_\gamma$	$\lambda$	$\alpha$	$\mu$	$c$	$\pi_{miss}$	$\pi_{low}$	$\pi_{high}$
.118	.020	.067	.103	.088	.467	12.0	.011	.026	.036	.029
( $7.4e^{-4}$ )	( $6.3e^{-4}$ )	( $2.1e^{-4}$ )	( $1.4e^{-4}$ )	(.001)	(.001)	(.279)	(.002)	(.003)	(.016)	(.010)

Notes: Standard errors are included in parenthesis.

The estimate of  $\pi_{high}$  is lower than that of  $\pi_{low}$ , suggesting that independent boards are more aggressive monitors of CEO performance (Hermalin and Weisbach, 1998). In particular, the estimates imply a roughly 15% decrease in the effective cost of CEO replacement upon switching from a low to high level of board independence. This difference, however, is not significant, as the confidence intervals of the estimates of  $\pi_{high}$  and  $\pi_{low}$  overlap. Though minor, this differential has implications for the effects of entrenchment on CEO pay across the board types, which I assess by conducting the same decomposition exercise as in Table 4. By assumption, different types of boards are subject to the same parameter values with the exception of  $\pi_k$  for  $k \in \{\text{missing, high, low}\}$ . Thus, I focus specifically on the differential impact of entrenchment on CEO pay across the two types of boards. I plot the results of this exercise in Table 6, which shows the implied increase in compensation stemming from entrenchment.

As before, high-quality CEOs benefit the most from entrenchment irrespective of board type. Additionally, as  $\pi_{high} < \pi_{low}$ , the effect of entrenchment on CEO pay is mitigated to some extent as board independence increases. The effect however does not disappear entirely, as the estimate of  $\pi_{high}$  is still significantly larger than zero. This is consistent with prior findings in the literature, namely that of Coles et al. (2014), who argue that board independence alone is not an exhaustive remedy for the issue of inefficient CEO monitoring practices. They make the important observation that even if a given board consists of a high proportion of independent directors, it may still be possible that the CEO exercises a great deal of influence over them. In particular, CEOs may strategically appoint sympathetic independent directors who share some alignment over private objectives. Such a “co-opted” director, while classified as independent, may still induce some degree of CEO entrenchment. Considerations of strategic board appointments made by the CEO, and the subsequent impact on CEO incentives, is important but

<sup>44</sup>The median of  $s_j$  in the sample is .83.

Table 6: Marginal Effect of Entrenchment on CEO Pay

	High Independence ( $\pi_{high}$ )	Low Independence ( $\pi_{low}$ )
<i>Quintile 1</i>	\$41.9 K	\$135 K
<i>Quintile 2</i>	\$233 K	\$369 K
<i>Quintile 3</i>	\$407 K	\$566 K
<i>Quintile 4</i>	\$613 K	\$782 K
<i>Quintile 5</i>	\$849 K	\$1.04 M

Notes: Given the estimates of  $\pi_{high}$  and  $\pi_{low}$ , Table 6 computes the marginal effect of CEO entrenchment on compensation under boards with high and low independence in exactly the same fashion as in Table 4. As before, I compute the effect separately across the five quintiles of the quality distribution. Because the entrenchment parameter  $\pi_k$  is the only parameter which varies with respect to board performance, I suppress the marginal effects of the other model channels as they are mechanically the same. The effect of entrenchment on compensation is higher under boards with low levels of independence given the higher degree of entrenchment under these boards.

beyond the scope of the analytic framework of this paper.

## 8 Conclusion

This paper shines light on the impact of turnover frictions on dynamic managerial incentives. The theoretical model, motivated by a set of empirical facts, provides a tractable framework in which moral hazard, entrenchment, and reputation can be studied comprehensively. The results highlight the substitutability of financial incentives and termination-based incentives when motivating managers. Entrenchment increases the cost of CEO replacement, weakening termination-based incentives and increasing the level of compensation needed to induce efficient CEO behavior. The model predicts a considerable decrease in managerial pay upon the elimination of entrenchment.

Through counterfactual experiments, I compute the entrenchment-induced transfer in surplus from shareholders to the CEO, which I refer to as the “entrenchment premium.” I decompose equilibrium CEO compensation and show that this premium increases with CEO quality. Thus, entrenchment partially mitigates shareholders’ benefits associated with the employment of a high-quality CEO. I show further that this premium has minimal sensitivity to the level of board independence. This is consistent with the findings of Coles et al. (2014), who argue that strategic appointments of sympathetic independent directors by the CEO weakens the relationship between board independence and monitoring efficiency. An important extension of this paper would take this addi-

tional layer of strategic CEO behavior into consideration. While in this paper, as in Taylor (2010), I assume that the level of entrenchment is fixed over time, it may be more realistic to allow CEOs to “entrench themselves” through the strategic manipulation of board characteristics. Determining to what extent such behavior can be prevented through the design of a proper incentive contract is an important question for future research.

# A Appendix

## A.1 Proofs of Key Theorems

Throughout the model appendix, I suppress the  $ij$  subscripts to ease notational burden.

*Proposition 1.* Given state  $(\tilde{\theta}_t, t)$ , the value of the firm continuing with their current CEO is:

$$rV(\tilde{\theta}, t) = b\tilde{\theta} + V_t(\tilde{\theta}, t) + \frac{v_t^2 \sigma_W^2}{2} V_{\theta\theta}(\tilde{\theta}, t) + \lambda V_T \quad (43)$$

(43) is the firm's Hamilton-Jacobi-Bellman (HJB) equation, the standard recursive representation of the value function in continuous time. To derive the firing threshold, I apply the value matching and smooth pasting conditions along the stopping boundary:

$$V(\theta_f^{FB}(t), t) = V_T - b\pi \quad (44)$$

$$V_t(\theta_f^{FB}(t), t) = \frac{\partial V_T}{\partial t} = 0 \quad (45)$$

Inserting these conditions into the HJB equation (43) evaluated at  $\theta_f^{FB}(t)$  yields:

$$\theta_f^{FB}(t) = -r\pi + b^{-1} \left( \rho V_T - \frac{v_{ijt}^2 \sigma_W^2}{2} V_{\theta\theta}(\theta_f^{FB}(t), t) \right) \quad (46)$$

□

*Proposition 2.* To derive necessary conditions in the second-best case, I first reformulate the CEO's problem. Recall that off of the equilibrium path, the board and CEO will not share the same beliefs about quality. It is thus convenient to define  $\alpha_t = \tilde{\theta}_t^a - \tilde{\theta}_t$  as the gap in beliefs at tenure  $t$ . Rewriting the innovation process in (20), we see that the board believes performance follows:

$$dY_t = \tilde{\theta}_t dt + \sigma_W dZ_t \quad (47)$$

whereas the CEO knows that profitability truly follows:

$$\begin{aligned} dY_t &= \tilde{\theta}_t dt + \sigma_W dZ_t^a \\ &= (\alpha_t - \hat{a}_t + \tilde{\theta}_t) dt + \sigma_W dZ_t \end{aligned} \quad (48)$$

Hence, when the CEO deviates from the board's assumed action path  $\{a_t = 0\}$ , the board and CEO assign different probability measures to realizations of performance. Let  $P_t^0$

and  $P_t^a$  denote the probability measures arising from the board and CEO's information ( $\mathcal{F}_t$  and  $\mathcal{F}_t^a$ ), respectively. By Girsanov's Theorem:

$$dP_t^a = \Lambda_t dP_t^0 \quad (49)$$

$$\Lambda_t = \exp\left(\frac{1}{\sigma_W} \int_0^t (\alpha_s - \hat{a}_s) dZ_s - \frac{1}{2\sigma_W^2} \int_0^t (\alpha_s - \hat{a}_s)^2 ds\right) \quad (50)$$

$$d\Lambda_t = \Lambda_t \frac{\alpha_t - \hat{a}_t}{\sigma_W} dZ_t \quad (51)$$

$\Lambda_t$  is the Radon-Nikodym derivative relating the measures  $P_t^0$  and  $P_t^a$ , where  $\Lambda_t = 1$  in the case of no deviations. The relative density process  $\Lambda$  can be used to reformulate the CEO's problem:

$$\begin{aligned} & \max_a \mathbb{E}_t^a \left[ \int_t^T e^{-r(s-t)} [u(w_s, a_s) + \lambda C(\tilde{\theta}_s^a)] ds + e^{-r(T-t)} C(\tilde{\theta}_T^a) \right] \\ & = \max_a \mathbb{E}_t \left[ \int_t^T e^{-r(s-t)} \Lambda_t [u(w_s, a_s) + \lambda C(\tilde{\theta}_s^a)] ds + e^{-r(T-t)} \Lambda_T C(\tilde{\theta}_T^a) \right] \end{aligned} \quad (52)$$

Note that the inclusion of  $\Lambda_t$  allows the expectation operator to be interchanged. Using this representation, I solve the weak formulation of the CEO's problem, where the choice of  $a_t$  corresponds to a choice of distribution  $P_t^a$  over  $Y_t$  (Cvitanic and Zhang, 2012). For the Hamiltonian function  $H$  and pair of states  $(\Lambda_t, \alpha_t)$ , define the corresponding adjoint processes  $(p_t^\Lambda, p_t^\alpha)$  as the solutions to the backward SDEs:

$$dp_t^\Lambda = rp_t^\Lambda dt - H_\Lambda dt + q_t^\Lambda dZ_t \quad (53)$$

$$dp_t^\alpha = rp_t^\alpha dt - H_\alpha dt + q_t^\alpha dZ_t \quad (54)$$

where  $(q_t^\Lambda, q_t^\alpha)$  are the volatility processes corresponding to the adjoint pair  $(p_t^\Lambda, p_t^\alpha)$ . Conditions (53) and (54) are the stochastic counterparts of the deterministic Pontryagin's maximum principle. The CEO's (current value) Hamiltonian reads:

$$H(t, \Lambda, \alpha, a, p^\Lambda, p^\alpha, q^\Lambda, q^\alpha) = \Lambda(w + \phi a + \lambda C(\tilde{\theta}^a)) + p^\alpha(v(a - \alpha)) + q^\Lambda \Lambda \frac{\alpha - a}{\sigma_W} \quad (55)$$

Incentive compatibility of the first-best action  $a_t = 0$  requires  $\frac{\partial H}{\partial a} \leq 0$ . Differentiating the Hamiltonian yields:

$$\Lambda_t \phi + p_t^\alpha v_t - \frac{q_t^\Lambda}{\sigma_W} \Lambda_t \leq 0 \quad (56)$$

(56) is the necessary condition for the incentive compatibility of the efficient strategy  $a^*$ .

Observe that under  $\{a_t = 0\}$ ,  $\Lambda_t = 1$  and  $\alpha_t = 0$ :

$$dp_t^\Lambda = (rp_t^\Lambda - w_t - \lambda C(\tilde{\theta}_t))dt + q_t^\Lambda dZ_t \quad (57)$$

$$dp_t^\alpha = (r + \nu_t)p_t^\alpha dt - \Lambda \frac{q_t^\Lambda}{\sigma_W} dt + q_t^\alpha dZ_t \quad (58)$$

with terminal values  $p_T^\Lambda = C(\tilde{\theta}_T)$  and  $p_T^\alpha = C_\alpha(\tilde{\theta}_T)$ . Letting  $\beta_t \equiv \frac{q_t^\Lambda}{\sigma_W}$ , the pair  $(p_t^\Lambda, \beta_t)$  is a weak solution to (57) where:

$$p_t^\Lambda = \mathbb{E}_t \left[ \int_t^T e^{-r(s-t)} w_s ds + \lambda \int_t^T e^{-r(s-t)} C(\tilde{\theta}_s) ds + e^{-r(T-t)} C(\tilde{\theta}_T) \right] \quad (59)$$

Notice that  $p_t^\Lambda = U_t$ , so the CEO's continuation payoff  $U_t$  becomes a state variable of the firm's problem, preserving the history-dependence of the contract. However, as discussed in Williams (2011), additional information is needed in environments with persistent private information. This additional information is represented by the second co-state variable  $p_t^\alpha$ . The pair  $(p_t^\alpha, q_t^\alpha)$  is a weak solution to (58) where:

$$p_t^\alpha = \mathbb{E}_t \left[ \nu_t^{-1} \int_t^T e^{-r(s-t)} \beta_s \nu_s ds + \nu_t^{-1} \nu_T e^{-r(T-t)} C_\alpha(\tilde{\theta}_T) \right] \quad (60)$$

$p_t^\alpha \equiv \Gamma_t$  denotes the CEO's *information rent*. Substituting into the IC constraint, and noting again that  $\Lambda_t = 1$  under the recommended action path, the constraint becomes:

$$\beta_t \geq \Gamma_t \nu_t + \phi \quad (61)$$

(57) can then be rewritten as:

$$dU_t = (rU_t - w_t - \lambda C(\tilde{\theta}_t))dt + \beta_t \sigma_W dZ_t \quad (62)$$

where  $\frac{dU_t}{dY_t} = \beta_t$  is the sensitivity of  $U_t$  to profitability  $dY_t$ . The constraint (61) implies a lower bound on the sensitivity process  $\beta_t$  and information rent  $\Gamma_t$ . Substituting the constraint into (58) yields:

$$d\Gamma_t = (r + \nu_t)\Gamma_t dt - \beta_t dt + q_t^\alpha dZ_t \quad (63)$$

$$\geq (r\Gamma_t - \phi)dt + q_t^\alpha dZ_t \equiv d\Gamma_t^* \quad (64)$$

Solving (64) using the terminal value  $\Gamma_t^* = C_a(\tilde{\theta}_T) = \frac{\phi}{r}$  yields:

$$\Gamma_t^* = \mathbb{E} \left[ \int_t^T e^{-r(s-t)} \phi ds + e^{-r(T-t)} \frac{\phi}{r} \right] \quad (65)$$

$$= \frac{\phi}{r} \quad (66)$$

$\Gamma_t^*$  is the minimum value of  $\Gamma_t$  in equilibrium, which is attained when the IC binds. Under the assumed functional form of the CEO's outside option,  $\Gamma_t^*$  is a constant, so can be dispensed with as a state variable as long as the IC binds.  $\square$

*Proposition 3.* Given an arbitrary termination boundary  $\theta_f(t)$ , the CEO's equilibrium value of continuing employment satisfies the HJB equation:

$$rU(\tilde{\theta}_t, t, \theta_f(t)) = w_t + \frac{\partial U}{\partial t} + \theta_f'(t) \frac{\partial U}{\partial \theta_f} + \frac{\nu_t^2 \sigma_W^2}{2} \frac{\partial^2 U}{\partial \theta^2} + \lambda C(\tilde{\theta}_t) \quad (67)$$

The cost-minimizing compensation process ensures that  $U(\tilde{\theta}_t, t, \theta_f(t)) = U_t^*$  for all  $t$ . Substituting the expression for  $U_t^*$  into the HJB equation above and collecting terms yields the equilibrium pay process (34).  $\square$

*Proposition 4.* Applying similar arguments as in the first-best case, Ito's lemma implies that the firm's HJB equation is given by:

$$\begin{aligned} rV(\tilde{\theta}_t, t, U_t) &= b\tilde{\theta}_t - w_t + V_t(\tilde{\theta}_t, t, U_t) + (rU_t - w_t - \lambda C(\tilde{\theta}_t))V_U(\tilde{\theta}_t, t, U_t) \\ &\quad + \frac{\nu_t^2 \sigma_W^2}{2} V_{\theta\theta}(\tilde{\theta}_t, t, U_t) + \frac{\beta_t^2 \sigma_W^2}{2} V_{UU}(\tilde{\theta}_t, t, U_t) + \lambda V_T \end{aligned} \quad (68)$$

To derive the optimal firing boundary  $\theta_f(t)$ , I apply the following conditions:

$$V(\theta_f(t), t, U_t) = V_T - b\pi \quad (69)$$

$$V_t(\theta_f(t), t, U_t) = \frac{\partial V_T}{\partial t} = 0 \quad (70)$$

$$V_U(\theta_f(t), t, U_t) = \frac{\partial V_T}{\partial U} = 0 \quad (71)$$

Condition (69) is the value matching condition, while conditions (70) and (71) are smooth pasting conditions. Substituting these conditions into the HJB equation evaluated along

the firing boundary yields:

$$\theta_f(t) = -r\pi + b^{-1} \left( \rho V_T - \frac{\nu_t^2 \sigma_W^2}{2} V_{\theta\theta}(\theta_f(t), t, U_t) - \frac{\beta_t^2 \sigma_W^2}{2} V_{UU}(\theta_f(t), t, U_t) \right) \quad (72)$$

□

## A.2 Model Fit

Figure 8: Fit of Estimation Moments

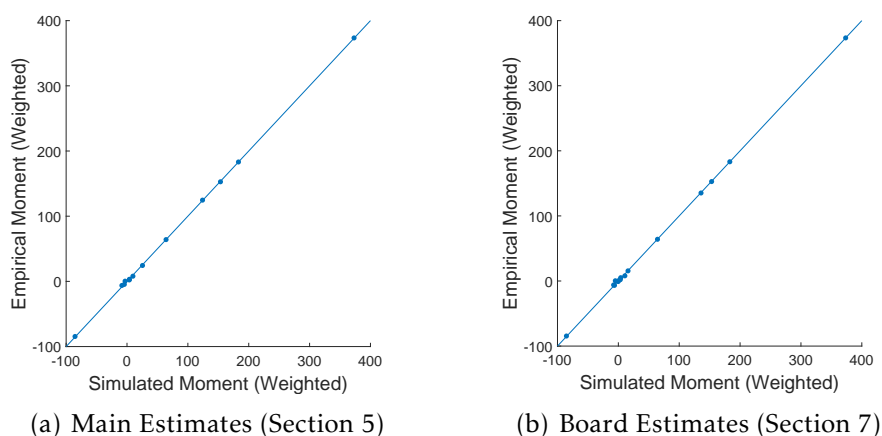


Figure 8 scatters the empirical moments over their simulated counterparts using the 45-degree line as a reference point. Panel (a) corresponds to the main estimation routine with results presented in Section 5, while Panel (b) does the same with the board estimates presented in Section 7. All moments are scaled by the corresponding empirical standard error.

## References

- Ai, Hengjie, Dana Kiku, and Rui Li**, “A Quantitative Model of Dynamic Moral Hazard,” *The Review of Financial Studies*, 08 2022, 36 (4), 1408–1463.
- Ali, Ashiq and Weining Zhang**, “CEO tenure and earnings management,” *Journal of Accounting and Economics*, 2015, 59 (1), 60–79.
- Allgood, Sam and Kathleen A. Farrell**, “The Match between CEO and Firm,” *The Journal of Business*, 2003, 76 (2), 317–341.



- Barry, John**, “Shareholder Voice and Executive Compensation,” *Working Paper*, 2023.
- Bebchuk, Lucian, Alma Cohen, and Allen Ferrell**, “What matters in corporate governance?,” *The Review of financial studies*, 2009, 22 (2), 783–827.
- Bennedsen, Morten, Francisco Perez-Gonzalez, and Daniel Wolfenzon**, “Do CEOs Matter? Evidence from Hospitalization Events,” *The Journal of Finance*, 2020, 75 (4), 1877–1911.
- , **Kasper Meisner Nielsen, Francisco Perez-Gonzalez, and Daniel Wolfenzon**, “Inside the Family Firm: The Role of Families in Succession Decisions and Performance,” *The Quarterly Journal of Economics*, 2007, 122 (2), 647–691.
- Berger, Philip G., Eli Ofek, and David L. Yermack**, “Managerial Entrenchment and Capital Structure Decisions,” *The Journal of Finance*, 1997, 52 (4), 1411–1438.
- Bertomeu, Jeremy, Edwige Cheynel, Edward Xuejun Li, and Ying Liang**, “How Pervasive Is Earnings Management? Evidence from a Structural Model,” *Management Science*, 2021, 67 (8), 5145–5162.
- Bertrand, Marianne and Antoinette Schoar**, “Managing with Style: The Effect of Managers on Firm Policies,” *The Quarterly Journal of Economics*, 2003, 118 (4), 1169–1208.
- **and Sendhil Mullainathan**, “Are CEOs Rewarded for Luck? The Ones Without Principals Are\*,” *The Quarterly Journal of Economics*, 08 2001, 116 (3), 901–932.
- Biais, Bruno, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve**, “Large Risks, Limited Liability, and Dynamic Moral Hazard,” *Econometrica*, 2010, 78 (1), 73–118.
- Bismut, Jean-Michel**, “Conjugate convex functions in optimal stochastic control,” *Journal of Mathematical Analysis and Applications*, 1973, 44 (2), 384–404.
- Black, Fischer and Myron Scholes**, “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 1973, 81 (3), 637–654.
- Brandimarte, P.**, *Numerical Methods in Finance and Economics: A MATLAB-Based Introduction* 04 2006.
- Capron, Andrew, Barton Hamilton, Andrés Hincapié, and Noah Lyman**, “Insider CEOs: Lucky or Good?,” *Working Paper*, 2024.

- Chang, Xin and Hong Feng Zhang**, “Managerial Entrenchment and Firm Value: A Dynamic Perspective,” *The Journal of Financial and Quantitative Analysis*, 2015, 50 (5), 1083–1103.
- Chetty, Raj, John N. Friedman, and Jonah E. Rockoff**, “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates,” *American Economic Review*, September 2014, 104 (9), 2593–2632.
- Coles, Jeffrey L., Naveen D. Daniel, and Lalitha Naveen**, “Managerial incentives and risk-taking,” *Journal of Financial Economics*, 2006, 79 (2), 431–468.
- , —, and —, “Co-opted Boards,” *The Review of Financial Studies*, 04 2014, 27 (6), 1751–1796.
- Core, John and Wayne Guay**, “Estimating the Value of Employee Stock Option Portfolios and Their Sensitivities to Price and Volatility,” *Journal of Accounting Research*, 2002, 40 (3), 613–630.
- Cvitanic, Jakša and Jianfeng Zhang**, *Contract theory in continuous-time models*, Springer Science & Business Media, 2012.
- Cziraki, Peter and Dirk Jenter**, “The Market for CEOs,” *European Corporate Governance Institute - Finance Working Paper No. 831/2022*, 2024.
- Daniel, Naveen D, Yuanzhi Li, and Lalitha Naveen**, “Symmetry in Pay for Luck,” *The Review of Financial Studies*, 09 2019, 33 (7), 3174–3204.
- DeMarzo, Peter M. and Michael J. Fishman**, “Optimal Long-Term Financial Contracting,” *The Review of Financial Studies*, 09 2007, 20 (6), 2079–2128.
- Demarzo, Peter M. and Yuliy Sannikov**, “Learning, Termination, and Payout Policy in Dynamic Incentive Contracts,” *The Review of Economic Studies*, 07 2017, 84 (1), 182–236.
- Duffie, Darrell and Kenneth J. Singleton**, “Simulated Moments Estimation of Markov Models of Asset Prices,” *Econometrica*, 1993, 61 (4), 929–952.
- Edmans, Alex, Xavier Gabaix, and Augustin Landier**, “A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium,” *The Review of Financial Studies*, 12 2008, 22 (12), 4881–4917.

- , – , and **Dirk Jenter**, “Executive Compensation: A Survey of Theory and Evidence,” 2017, 1, 383–539.
- Fama, Eugene**, “Agency Problems and the Theory of the Firm,” *Journal of Political Economy*, 1980, 88 (2), 288–307.
- Fee, C. Edward and Charles J. Hadlock**, “Raids, Rewards, and Reputations in the Market for Managerial Talent,” *The Review of Financial Studies*, 04 2015, 16 (4), 1315–1357.
- Ferraro, Valeria**, “Media Focus and Executive Turnover: Consequences for Female Leadership,” *Working Paper*, 2021.
- Gabaix, Xavier and Augustin Landier**, “Why has CEO Pay Increased So Much?\*,” *The Quarterly Journal of Economics*, 02 2008, 123 (1), 49–100.
- Garvey, Gerald T. and Todd T. Milbourn**, “Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad,” *Journal of Financial Economics*, 2006, 82 (1), 197–225.
- Gayle, George-Levi and Robert A. Miller**, “Has Moral Hazard Become a More Important Factor in Managerial Compensation?,” *American Economic Review*, December 2009, 99 (5), 1740–69.
- , **Limor Golan, and Robert A. Miller**, “Promotion, Turnover, and Compensation in the Executive Labor Market,” *Econometrica*, 2015, 83 (6), 2293–2369.
- Gibbons, Robert and Kevin Murphy**, “Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence,” *Journal of Political Economy*, 02 1992, 100, 468–505.
- Gompers, Paul, Joy Ishii, and Andrew Metrick**, “Corporate Governance and Equity Prices,” *The Quarterly Journal of Economics*, 02 2003, 118 (1), 107–156.
- Grochulski, Borys and Yuzhe Zhang**, “Termination as an incentive device,” *Theoretical Economics*, 2023, 18 (1), 381–419.
- Hamilton, Barton, Andrés Hincapié, and Noah Lyman**, “The Bright Side to Nepotism? Family CEOs, Turnover, and Firm Performance,” *Working Paper*, 2024.
- Hermalin, Benjamin and Michael Weisbach**, “Endogenously Chosen Boards of Directors and Their Monitoring of the CEO,” *The American Economic Review*, 1998, 88 (1), 96–118.

- Hermalin, Benjamin E.**, “Trends in Corporate Governance,” *The Journal of Finance*, 2005, 60 (5), 2351–2384.
- Holmström, Bengt**, “Managerial Incentive Problems: A Dynamic Perspective,” *The Review of Economic Studies*, 01 1999, 66 (1), 169–182.
- Hu, Aidong and Praveen Kumar**, “Managerial Entrenchment and Payout Policy,” *Journal of Financial and Quantitative Analysis*, 2004, 39 (4), 759–790.
- Jenter, Dirk and Fadi Kanaan**, “CEO Turnover and Relative Performance Evaluation,” *The Journal of Finance*, 2015, 70 (5), 2155–2184.
- Jovanovic, Boyan**, “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 1979, 87 (5), 972–90.
- **and Julien Prat**, “Dynamic contracts when the agent’s quality is unknown,” *Theoretical Economics*, 2014, 9 (3), 865–914.
- Kaplan, Steven N., Mark M. Klebanov, and Morten Sorensen**, “Which CEO Characteristics and Abilities Matter?,” *The Journal of Finance*, 2012, 67 (3), 973–1007.
- Laffont, Jean-Jacques and Jean Tirole**, “The Dynamics of Incentive Contracts,” *Econometrica*, 1988, 56 (5), 1153–1175.
- Lehn, Kenneth, Sukesh Patro, and Mengxin Zhao**, “Governance indexes and valuation: Which causes which?,” *Journal of Corporate Finance*, 2007, 13 (5), 907–928.
- Lippi, Francesco and Fabiano Schivardi**, “Corporate control and executive selection,” *Quantitative Economics*, 2014, 5 (2), 417–456.
- Mackey, Alison**, “The Effect of CEOs on Firm Performance,” *Strategic Management Journal*, 2008, 29 (12), 1357–1367.
- Margiotta, Mary M and Robert A Miller**, “Managerial Compensation and the Cost of Moral Hazard,” *International Economic Review*, 2000, 41 (3), 669–719.
- Merton, Robert C.**, “Theory of Rational Option Pricing,” *The Bell Journal of Economics and Management Science*, 1973, 4 (1), 141–183.
- Milbourn, Todd T.**, “CEO reputation and stock-based compensation,” *Journal of Financial Economics*, 2003, 68 (2), 233–262.

- Miller, Robert A.**, “Job Matching and Occupational Choice,” *Journal of Political Economy*, 1984, 92 (6), 1086–1120.
- Moscarini, Giuseppe**, “Job Matching and the Wage Distribution,” *Econometrica*, 2005, 73 (2), 481–516.
- Page, T. Beau**, “CEO attributes, compensation, and firm value: Evidence from a structural estimation,” *Journal of Financial Economics*, 2018, 128 (2), 378–401.
- Parrino, Robert**, “CEO Turnover and Outside Succession, A Cross-Sectional Analysis,” *Journal of Financial Economics*, 1997, 46 (2), 165–197.
- Peréz-González, Francisco**, “Inherited Control and Firm Performance,” *American Economic Review*, 2006, 96 (5), 1559–1588.
- Peters, Florian S. and Alexander F. Wagner**, “The Executive Turnover Risk Premium,” *The Journal of Finance*, 2014, 69 (4), 1529–1563.
- Sannikov, Yuliy**, “A Continuous- Time Version of the Principal: Agent Problem,” *The Review of Economic Studies*, 2008, 75 (3), 957–984.
- Shleifer, Andrei and Robert W. Vishny**, “Managerial Entrenchment: The Case of Manager-Specific Investments,” *Journal of Financial Economics*, 1989, 25 (1), 123–139.
- Spear, Stephen and Cheng Wang**, “When to fire a CEO: optimal termination in dynamic contracts,” *Journal of Economic Theory*, 2005, 120 (2), 239–256.
- Stiglitz, Joseph E. and Andrew Weiss**, “Incentive Effects of Terminations: Applications to the Credit and Labor Markets,” *The American Economic Review*, 1983, 73 (5), 912–927.
- Taylor, Lucian**, “Why Are CEOs Rarely Fired? Evidence from Structural Estimation,” *The Journal of Finance*, 2010, 65 (6), 2051–2087.
- Williams, Noah**, “Persistent Private Information,” *Econometrica*, 2011, 79 (4), 1233–1275.
- Zhu, John Y.**, “Optimal Contracts with Shirking,” *The Review of Economic Studies*, 2013, 80 (2 (283)), 812–839.
- Zwiebel, Jeffrey**, “Dynamic Capital Structure under Managerial Entrenchment,” *The American Economic Review*, 1996, 86 (5), 1197–1215.

## B Online Appendix

### B.1 Data Appendix

**Execucomp** I obtain data on CEO pay and tenure from Execucomp. Each CEO-firm match is uniquely identified by the variable *co\_per\_rol*. The key compensation variable I use in estimation is *tdc1*, defined as “Total Compensation (Salary + Bonus + Other Annual + Restricted Stock Grants + LTIP Payouts + All Other + Value of Option Grants).” I convert *tdc1* to millions of dollars in estimation; all nominal variables are denoted in 2015 dollars. I winsorize the distribution of *tdc1* at its 1st and 99th percentiles.

**Compustat** I obtain company fundamentals data from Compustat North America, which contains a rich set of financial information on publicly held companies in Canada and the U.S. Each firm is uniquely identified by the variable *gvkey*. Using operating income before depreciation (item *oibdp*) and total assets (item *at*) I compute return on assets ( $ROA_{ijt}$ ) for each firm-year as:

$$ROA_{ijt} = \frac{oibdp_{ijt}}{at_{ijt}}$$

I drop firms with values of  $ROA_{ijt}$  outside of the range  $[-1, 1]$  (70 observations). Industries are defined using the Global Industry Classification Standard (GICS) codes, corresponding to the Compustat variable *gind*.

**Forced turnover data** Data on forced CEO turnover was graciously shared by Florian Peters. He and a team of researchers gathered these data for CEOs listed in Execucomp from years 1995 to 2015. The criteria used to classify turnover as forced are described in detail in Peters and Wagner (2014) and Jenter and Kanaan (2015). Both methodologies follow the three-step criteria to classify successions as forced from Parrino (1997):

1. “All successions for which the *Wall Street Journal* reports that the CEO is fired, forced from the position, or departs due to unspecified policy differences are classified as forced.”
2. “All other successions in which the departing CEO is under age 60 are reviewed to identify cases in which the *Wall Street Journal* announcement of the succession either (1) does not report the reason for departure as involving death, poor health, or the acceptance of another position (elsewhere or within the firm), or (2) reports that the CEO is retiring, but does

not announce the retirement at least six months before the succession. These cases are also classified as forced successions.”

3. “The circumstances surrounding departures that are classified as forced in the previous step are further investigated by searching the business and trade press for relevant articles. These successions are reclassified as voluntary if the incumbent takes a comparable position elsewhere or departs for previously undisclosed personal or business reasons that are unrelated to the firm’s activities.”

If turnover is not classified as forced in Florian Peters’ data, it is assumed to be voluntary. For a small number of cases, forced turnover is reported in year  $t$ , but the executive is still listed as CEO in year  $t + 1$ . To avoid inconsistencies, all indicators of turnover are moved to the last year of the CEOs tenure as reported in Execucomp. In my final sample, I observe 908 instances of forced turnover and 2,667 instances of voluntary turnover.

### B.1.1 Profitability and CEO Tenure

Table 7: Profitability and CEO Tenure

	Profitability and CEO Tenure			
	(1)	SE	(2)	SE
	Coefficient		Coefficient	
<i>CEO Characteristics:</i>				
Tenure	.059***	(.008)	.181	(.183)
Age	-.009	(.009)	-.036	(.080)
Female	.085	(.273)	-	-
<i>Firm Characteristics:</i>				
Log(Assets)	-4.61***	(.062)	-8.71***	(.534)
Log(Revenue)	5.94***	(.087)	11.7***	(.654)
<i>Fixed Effects:</i>				
Year		✓		✓
CEO-Firm Match				✓
Observations	41,415		41,415	

Notes: Column (1) reports pooled OLS estimates while column (2) reports within-match estimates. CEO gender is omitted from column (2) since this is fixed within match. The tenure effect disappears within match.

In principle, profitability may rise with tenure as a result of learning by doing on part of the CEO. To test this, I report in Table 7 estimates from an ROA regression with and without CEO-firm match effects. Across matches, there is a positive and significant relationship between CEO tenure and firm performance. However, the tenure effect vanishes

within match. This is evidence favoring selection on CEO quality as the key determinant of the tenure-profitability relationship as opposed to learning by doing.

### B.1.2 Dismissal, Retirement, and CEO Performance Histories

Table 8: Turnover and Cumulative Performance

	Forced Turnover		Voluntary Turnover	
	Marginal Effect	SE	Marginal Effect	SE
<i>CEO Reputation:</i>				
$\tilde{\theta}_{ijt}$ (Standardized)	-.003***	(.001)	$-3.5e^{-4}$	(.002)
<i>CEO Characteristics:</i>				
Tenure	-.001***	( $1.4e^{-4}$ )	-.001***	( $1.7e^{-4}$ )
Age	$-4.1e^{-4}$ ***	( $1.1e^{-4}$ )	.007***	( $2.2e^{-4}$ )
Female	.008**	(.004)	-.016**	(.008)
<i>Firm Characteristics:</i>				
Log(Assets)	-.002**	(.001)	-.005***	(.001)
Log(Revenue)	.002	(.001)	.006***	(.002)
Observations	41,202		41,202	

Notes: Columns (1) and (2) respectively report marginal effects obtained from a logit regression of forced and voluntary turnover indicators on vectors of CEO and firm characteristics.

Taylor (2010) and Hamilton et al. (2024) suggest that firms make CEO replacement decisions in response to new information about CEO quality. To proxy for firms' evolving information set, I define the *adaptive* shrinkage estimator  $\tilde{\theta}_{ijt}$  as the best estimate of  $\theta_{ij}$  given information at time  $t$ . Concretely,  $\tilde{\theta}_{ijt}$  is defined as:

$$\tilde{\theta}_{ijt} = \frac{\tilde{\theta}_{ijt-1} + \omega_j \hat{\epsilon}_{ijt}}{1 + \omega_j} \quad (73)$$

$\tilde{\theta}_{ijt}$  is the cumulative weighted average of the performance residuals  $\hat{\epsilon}_{ijt}$  implied by estimating Equation (1). I refer to the quantity  $\tilde{\theta}_{ijt}$  as CEO  $i$ 's *reputation* with firm  $j$  at time  $t$ .<sup>45</sup>

Table 8 reports marginal effects obtained from regressing forced and voluntary turnover indicators on sets of CEO and firm characteristics, along with the reputation proxy  $\tilde{\theta}_{ijt}$ . The probability of forced turnover significantly declines in response to positive cumulative performance, as proxied by  $\tilde{\theta}_{ijt}$ . On the other hand, the probability of voluntary

<sup>45</sup> $\tilde{\theta}_{ijt}$  only conditions on performance residuals up to time  $t$ , whereas the James-Stein estimator  $\hat{\theta}_{ij}$  conditions on the CEO's complete performance history. The terminal value of  $\tilde{\theta}_{ijt}$  is the baseline James-Stein estimate  $\hat{\theta}_{ij}$ .

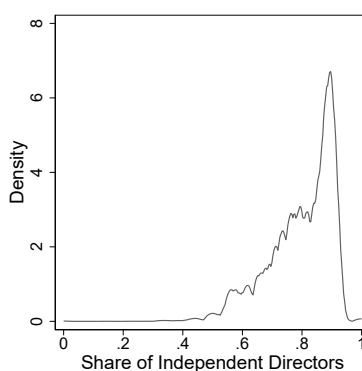


turnover is unaffected by cumulative performance. In light of these results, I assume in the model section that voluntary turnover is independent of CEOs' reputation.

### B.1.3 ISS Director Data

To assess the impact of board independence on CEO entrenchment, I augment my main sample with director information obtained from Institutional Shareholder Services (ISS). The director sample ranges from 2007-2019 and is therefore missing for all firm-year observations prior to 2007. Furthermore, ISS covers only a subset of firms in my full sample. In total, I am missing director information for 65% of observations in my main sample. For firms with non-missing director information, I compute the share of directors classified as independent in each firm year. The median board consists of roughly 83% independent directors, which I use as my threshold to classify boards as either "high" or "low." The density of this variable is plotted in Figure 9.

Figure 9: Distribution of Board Independence



Notes: Figure 9 plots the kernel density of the share of independent directors on the board for all observations contained within the ISS Director Data sample. Its median is given by .83.

## B.2 Additional Model Results and Discussion

### B.2.1 Belief Manipulation

Suppose that at tenure  $t$  the firm recommends  $a_{ijt}^* = 0$ , but the CEO deviates to  $\hat{a}_{ijt} > 0$ . The firm, assuming the CEO took the recommended action, updates beliefs according to:

$$\begin{aligned} d\tilde{\theta}_{ijt} &= \nu_{ijt}(dY_{ijt} - \tilde{\theta}_{ijt}dt)d\tilde{\theta}_{ijt} \\ &= \nu_{ijt}\sigma_W dZ_{ijt} \end{aligned} \tag{74}$$

The CEO, knowing their true action choice, will update according to:

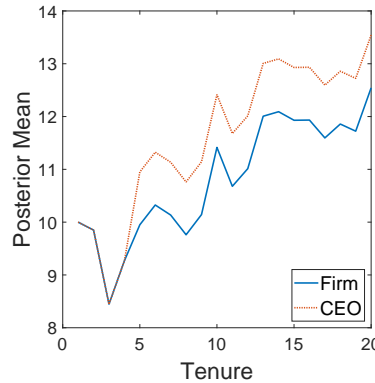
$$\begin{aligned} d\tilde{\theta}_{ijt}^a &= v_{ijt}(dY_{ijt} - (\tilde{\theta}_{ijt} - \hat{a}_{ijt})dt) \\ d\tilde{\theta}_{ijt}^a &= v_{ijt}\sigma_W dZ_{ijt}^a \end{aligned} \quad (75)$$

where  $dZ_{ijt}^a = \sigma_W^{-1}(dY_{ijt} - (\tilde{\theta}_{ijt} - \hat{a}_{ijt})dt)$  is the innovation process observed by the CEO. The asymmetry in information induces a discrepancy in the incremental belief update:

$$d\tilde{\theta}_{ijt}^a - d\tilde{\theta}_{ijt} = \hat{a}_{ijt}v_{ijt}dt > 0 \quad (76)$$

The CEO, who is perfectly informed of their action choices, can always update beliefs “correctly” in the sense that their estimate of  $\theta_i$  is unbiased. The firm on the other hand lacks knowledge of the CEO’s action choices, and instead assumes that the recommended action  $a_{ijt}^*$  has been selected. The firm thus misinterprets the information generated under a deviation. In particular, positive deviations decrease realized performance relative to the firm’s expectations  $(\tilde{\theta}_{ijt} - a_{ijt}^*)dt$ , leading the firm’s beliefs to drift downwards relative to those of the CEO. Figure 10 illustrates an example.

Figure 10: Belief Manipulation



Notes: Figure 10 illustrates the gap in beliefs induced following a deviation by the CEO. I simulate a sample path of profitability and record the corresponding beliefs. In this simulation, the firm recommends  $a_{ijt} = 0$  for all  $t$ . The CEO deviates to  $\hat{a}_{ijt} = .1$  at  $t = 5$ . This leads performance to fall short (on average) of the firm’s expectations, decreasing the firm’s beliefs relative to the CEO’s. The CEO is informed of their action choice, so updates beliefs accurately.

In essence, the firm faces an identification problem in which the effects of CEO quality and actions on performance cannot be disentangled, since neither are observable. The CEO can take advantage of this by deviating and inducing a gap in beliefs between firm and CEO. This benefits the CEO because the optimal contract will reward them if performance surpasses the firm’s expectations. The lower the firm’s expectations relative to the CEO’s, the higher the probability that realized performance exceeds expectations. Hence,

manipulating the firm's beliefs downwards increases the CEO's likelihood of accumulating rewards later in their employment spell. The gap in beliefs  $\alpha_{ijt} = \tilde{\theta}_{ijt}^a - \tilde{\theta}_{ijt}$  has law of motion:

$$d\alpha_{ijt} = v_{ijt}((\hat{a}_{ijt} - a_{ijt}^*) - \alpha_{ijt})dt \quad (77)$$

Note that in the absence of further deviations, the gap in beliefs converges to zero at rate  $v_{ijt}$ .

**Theorem 1.** *Given any IC contract  $\mathcal{C} = (a, c, T)$  recommending  $a_t > 0$  for some  $t$ , there is an alternative contract  $\mathcal{C}'$  recommending  $\{a'_t = 0\}_{t \geq 0}$  in which the CEO's payoff is unchanged and the firm's payoff is weakly greater.*

*Proof.* This proof is essentially a restatement of Lemma A in Demarzo and Sannikov (2017). Let  $\omega_t$  denote an arbitrary sample path of profitability  $Y_{ijt}$  up until tenure  $t$ . The compensation process  $c_t$  and stopping time  $T$  under the original contract  $\mathcal{C}$  map from sample paths to  $\mathbb{R}_+$ :

$$\begin{aligned} c_t &\equiv c_t\left(Y_s; s \in [0, t]\right) : \omega_t \rightarrow \mathbb{R}_+ \\ T &\equiv T\left(Y_s; s \in [0, t]\right) : \omega_t \rightarrow \mathbb{R}_+ \end{aligned}$$

Consider the alternative contract  $\mathcal{C}' = (a', c', T')$  with compensation and stopping time defined by:

$$\begin{aligned} c'_t &\equiv c_t\left(Y_s - \int_0^s a_l dl; s \in [0, t]\right) + \phi \int_0^t a_s ds \\ T' &\equiv T\left(Y_s - \int_0^s a_l dl; s \in [0, t]\right) \end{aligned}$$

which adjust compensation and the stopping time according to the cash flows the CEO would have diverted given original action recommendation  $\{a_t > 0\}$ . If the CEO selected strategy  $\{a'_t \geq 0\}$  under contract  $\mathcal{C}'$ , their flow payoff would be identical to their payoff under  $\mathcal{C}$  and  $\{a'_t + a_t\}$  for a given path  $\omega_t$ . Furthermore, under contract  $\mathcal{C}$  and strategy  $\{a'_t + a_t\}$ , the path of:

$$Y_t = Y_0 + \int_0^t (\tilde{\theta}_{ijt}^a - a'_s - a_s) ds + \sigma_W \int_0^t dZ_s \quad (78)$$

coincides with the path of:

$$Y_t = Y_0 + \int_0^t (\tilde{\theta}_{ijt}^a - a'_s) ds + \sigma_W \int_0^t dZ_s \quad (79)$$

under contract  $\mathcal{C}'$  and strategy  $\{a'_t\}$ . Given the definition of  $c'_t$ , the CEO's flow payoff under  $\mathcal{C}'$  is greater by  $\phi a_t dt$ , the amount they would have diverted, so their payoff under both contracts and respective strategies are identical. The incentive compatibility of original contract  $\mathcal{C}$  implies that  $\{a'_t = 0\}$  is optimal for the CEO under  $\mathcal{C}'$ . Additionally, given  $\{a_t > 0\}$  under  $\mathcal{C}$  the firm's payoff strictly increases under  $\mathcal{C}'$  as long as  $\phi = b^\alpha < b$ , which holds when  $\alpha < 1$ .  $\square$

## B.3 Numerical Solution and Estimation Routine

### B.3.1 Numerical Solution for First-Best Case

The retirement process ensures that  $T < \infty$ . Thus, to approximate the value function, I assume that  $t^* = T$  for some arbitrarily large value of  $t^*$ . I solve the model backwards from this point. I'll first introduce some notation largely following Brandimarte (2006). Define the discrete grids of state variables:

$$\Theta = \{\mu + \Delta\theta, \mu + 2\Delta\theta, \dots, \mu + (M-1)\Delta\theta, \mu + M\Delta\theta\} \quad (80)$$

$$\mathbb{T} = \{\Delta t, 2\Delta t, \dots, (N-1)\Delta t, N\Delta t\} \quad (81)$$

where  $\mu + \Delta\theta$  is the smallest value of  $\theta_{ijt}$  contained in its discrete grid. Define  $V_{i,j} \equiv V(\mu + i\Delta\theta, j\Delta t)$  as the discretized counterpart of the firm's value function evaluated at grid points  $\mu + i\Delta\theta$  and  $j\Delta t$ . I use a finite difference approach to approximate the derivatives in the HJB equation (43). I use the backward and standard approximations of  $V_t$  and  $V_{\theta\theta}$  respectively:

$$\frac{\partial V(\tilde{\theta}_t, t)}{\partial t} \approx \frac{V_{i,j} - V_{i,j-1}}{\Delta t} \quad (82)$$

$$\frac{\partial^2 V(\tilde{\theta}_t, t)}{\partial \theta^2} \approx \frac{V_{i+1,j-1} - 2V_{i,j-1} + V_{i-1,j-1}}{(\Delta\theta)^2} \quad (83)$$

Plugging these approximations into (43) and doing some algebra yields the discretized HJB equation:

$$V_{i,j} = A_{i,j}V_{i-1,j-1} + B_{i,j}V_{i,j-1} + C_{i,j}V_{i+1,j-1} + D_i \quad (84)$$

$$A_{i,j} = \frac{v_j^2 \sigma^2}{2} \rho \quad (85)$$

$$B_{i,j} = v_j^2 \sigma^2 \rho + r\Delta t + 1 \quad (86)$$

$$C_{i,j} = -\frac{v_j^2 \sigma^2}{2} \rho \quad (87)$$

$$D_i = -(\mu + i\Delta\theta)\Delta t \quad (88)$$

where  $\rho = \frac{\Delta t}{(\Delta\theta)^2}$ . This can be represented as an  $M - 1 \times M - 1$  system of linear equations:

$$\begin{pmatrix} V_{1,j} \\ V_{2,j} \\ V_{3,j} \\ \vdots \\ V_{M-1,j} \\ V_{M,j} \end{pmatrix} = \begin{pmatrix} B_{1,j} & C_{1,j} & 0 & 0 & 0 & 0 & 0 & 0 \\ A_{2,j} & B_{2,j} & C_{2,j} & 0 & 0 & 0 & 0 & 0 \\ 0 & A_{3,j} & B_{3,j} & C_{3,j} & 0 & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & A_{M-1,j} & B_{M-1,j} & C_{M-1,j} & \\ 0 & 0 & 0 & 0 & 0 & A_{M,j} & B_{M,j} & \end{pmatrix} \begin{pmatrix} V_{1,j-1} \\ V_{2,j-1} \\ V_{3,j-1} \\ \vdots \\ V_{M-1,j-1} \\ V_{M,j-1} \end{pmatrix} + \begin{pmatrix} D_1 + A_{1,j}V_{0,j-1} \\ D_2 \\ D_3 \\ \vdots \\ D_{M-1} \\ D_M + C_{M,j}V_{M+1,j-1} \end{pmatrix}$$

Note that the values  $V_{0,j-1}$  and  $V_{M+1,j-1}$  are not defined. These are instead given by the boundary conditions  $V_{0,j-1} = V_T$  and  $V_{M+1,j-1} = V(M\Delta\theta, (j-1)\Delta t)$ . Rewriting the system slightly:

$$\underbrace{\begin{pmatrix} V_{1,j} - (D_1 + A_{1,j}V_{0,j-1}) \\ V_{2,j} - D_2 \\ V_{3,j} - D_3 \\ \vdots \\ V_{M-1,j} - D_{M-1} \\ V_{M,j} - (D_M + C_{M,j}V_{M+1,j-1}) \end{pmatrix}}_{\mathbf{D}} = \underbrace{\begin{pmatrix} B_{1,j} & C_{1,j} & 0 & 0 & 0 & 0 & 0 & 0 \\ A_{2,j} & B_{2,j} & C_{2,j} & 0 & 0 & 0 & 0 & 0 \\ 0 & A_{3,j} & B_{3,j} & C_{3,j} & 0 & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & A_{M-1,j} & B_{M-1,j} & C_{M-1,j} & \\ 0 & 0 & 0 & 0 & 0 & A_{M,j} & B_{M,j} & \end{pmatrix}}_{\mathbf{Q}} \underbrace{\begin{pmatrix} V_{1,j-1} \\ V_{2,j-1} \\ V_{3,j-1} \\ \vdots \\ V_{M-1,j-1} \\ V_{M,j-1} \end{pmatrix}}_{\mathbf{V}}$$

The optimal values at each point in time are then obtained as  $\mathbf{V} = \mathbf{Q}^{-1}\mathbf{D}$ .

### B.3.2 Numerical Solution for Second-Best Case

The solution method for the second-best case works largely the same as in the first-best, with the addition of  $U_t$  as a state variable. As before, I assume some arbitrarily large  $t^*$  such that  $t^* = T$ .

For  $t < T$ , the firm's HJB equation is given by:

$$rV(\tilde{\theta}_t, t, U_t) = \tilde{\theta}_{ijt} - w_{ijt} + V_t + (rU_t - w_t - \lambda C(\tilde{\theta}_t))V_U \quad (89)$$

$$+ \frac{v_t^2 \sigma^2}{2} V_{\theta\theta} + \frac{\beta_t^2 \sigma^2}{2} V_{UU} + \lambda V_T \quad (90)$$

I approximate this with a linear system of equations using an upwind finite-difference scheme. The approximations of the relevant derivatives are given by:

$$\frac{\partial V(\tilde{\theta}_t, t, U_t)}{\partial t} \approx \frac{V_{i,j}^t - V_{i,j}^{t-1}}{\Delta t} \quad (91)$$

$$\frac{\partial^2 V(\tilde{\theta}_t, t, U_t)}{\partial \theta^2} \approx \frac{V_{i+1,j}^{t-1} - 2V_{i,j}^{t-1} + V_{i-1,j}^{t-1}}{(\Delta\theta)^2} \quad (92)$$

$$\frac{\partial^2 V(\tilde{\theta}_t, t, U_t)}{\partial U^2} \approx \frac{V_{i,j+1}^{t-1} - 2V_{i,j}^{t-1} + V_{i,j-1}^{t-1}}{(\Delta U)^2} \quad (93)$$

Consistent with the upwind scheme, I approximate the partial derivative  $V_U$  by:

$$\frac{\partial V(\tilde{\theta}_t, t, U_t)}{\partial U} \approx \begin{cases} \frac{V_{i,j+1}^{t-1} - V_{i,j}^{t-1}}{\Delta U} & \text{if } rU_{ijt} \geq w_{ijt} + \lambda C(\tilde{\theta}_{ijt}) \\ \frac{V_{i,j}^{t-1} - V_{i,j-1}^{t-1}}{\Delta U} & \text{if } rU_{ijt} < w_{ijt} + \lambda C(\tilde{\theta}_{ijt}) \end{cases} \quad (94)$$

Substituting these approximations into the firm's HJB yields the implicit scheme:

$$V_{i,j}^t = A^t V_{i-1,j}^{t-1} + B_{i,j}^t V_{i,j}^{t-1} + C^t V_{i+1,j}^{t-1} + D_{i,j}^t V_{i,j-1}^{t-1} + E_{i,j}^t V_{i,j+1}^{t-1} + F_{i,j}^t \quad (95)$$

$$A^t = -\frac{1}{2} v_t^2 \sigma^2 \rho_{\theta\theta} \quad (96)$$

$$B_{i,j}^t = 1 + r\Delta t + (rU_{i,j}^{t-1} - w_{i,j}^{t-1} - \lambda C(i\Delta\theta)) \rho_u \left( \mathbf{1}[drift \geq 0] - \mathbf{1}[drift < 0] \right) + \rho_{\theta\theta} v_t^2 \sigma^2 + \rho_{uu} \beta_t^2 \sigma_t^2 \quad (97)$$

$$C^t = -\frac{1}{2} v_t^2 \sigma^2 \rho_{\theta\theta} \quad (98)$$

$$D_{i,j}^t = (rU_{i,j}^{t-1} - w_{i,j}^{t-1} - \lambda C(i\Delta\theta)) \rho_u \mathbf{1}[drift < 0] - \frac{1}{2} \beta_t^2 \sigma^2 \rho_{uu} \quad (99)$$

$$E_{i,j}^t = -(rU_{i,j}^{t-1} - w_{i,j}^{t-1} - \lambda C(i\Delta\theta)) \rho_u \mathbf{1}[drift \geq 0] - \frac{1}{2} \beta_t^2 \sigma^2 \rho_{uu} \quad (100)$$

$$F_{i,j}^t = -(i\Delta\theta - w_{i,j}^{t-1} + \lambda V_T) \quad (101)$$

where  $\rho_u = \frac{\Delta t}{\Delta U}$ ,  $\rho_{uu} = \frac{(\Delta t)^2}{(\Delta U)^2}$ , and  $\rho_{\theta\theta} = \frac{(\Delta t)^2}{(\Delta\theta)^2}$ . This is represented as an  $(M-1)(N-1) \times (M-1)(N-1)$  linear system:

$$\begin{pmatrix} V_{1,1}^t \\ V_{2,1}^t \\ \vdots \\ V_{M,1}^t \\ V_{1,2}^t \\ \vdots \\ V_{M-1,N}^t \\ V_{M,N}^t \end{pmatrix} = \begin{pmatrix} B_{1,j} & C_{1,j} & 0 & 0 & \dots & E_{1,1}^t & 0 & \dots & \dots \\ A_{2,j} & B_{2,j} & C_{2,j} & 0 & \dots & 0 & E_{2,2}^t & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & A_{3,j} & B_{3,j} & C_{3,j} & \dots & E_{M,2}^t & \dots & 0 & \dots \\ D_{1,1}^t & \dots & A_{3,j} & B_{3,j} & C_{3,j} & \dots & \dots & E_{1,3}^t & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & D_{M-1,N-1}^t & \dots & \dots & 0 & 0 & A_{M-1,j} & B_{M-1,j} & C_{M-1,j} \\ 0 & 0 & D_{M,N-1}^t & \dots & 0 & 0 & 0 & A_{M,j} & B_{M,j} \end{pmatrix} \begin{pmatrix} V_{1,1}^{t-1} \\ V_{2,1}^{t-1} \\ \vdots \\ V_{M,1}^{t-1} \\ V_{1,2}^{t-1} \\ \vdots \\ V_{M-1,N}^{t-1} \\ V_{M,N}^{t-1} \end{pmatrix} + \begin{pmatrix} F_{1,1}^t + A_{1,j}V_{0,j-1} \\ F_{2,1}^t \\ \vdots \\ F_{M,1}^t \\ F_{1,2}^t \\ \vdots \\ F_{M-1,N}^t \\ F_{M,N}^t + C_{M,j}V_{M+1,j-1} \end{pmatrix}$$

With this representation in hand, I proceed by using the same method as in the first-best case.

## B.4 Estimation Appendix

### B.4.1 Weighting Matrix

From the empirical sample I obtain a  $K \times 1$  vector of moments  $\hat{M}$ . Let  $\Psi$  denote the corresponding  $N \times K$  matrix of influence functions,  $N$  being the number of observations in the sample. Each element  $\Psi_{nk}$  is the influence function describing observation  $n$ 's contribution to moment  $k$ . The covariance matrix of the vector of moments can then be estimated as:

$$a\hat{v}ar(\hat{M}) = \Psi' \Psi \quad (102)$$

The weighting matrix  $\hat{W}$  is then obtained as the inverse of matrix 102. Let  $\Theta \in \mathbb{R}^P$  denote an arbitrary vector of structural parameters. Define the moment residual  $g : \mathbb{R}^P \rightarrow \mathbb{R}^M$  as:

$$g(\Theta) = \hat{M} - \frac{1}{S} \sum_{s=1}^S \hat{m}^s(\Theta) \quad (103)$$

Where  $\hat{M}$  is the vector of empirical moments,  $\hat{m}^s(\Theta)$  is the vector of simulated moments given parameter values  $\Theta$  in simulation  $s$ , and  $S$  is the total number of simulations. The vector of estimates  $\hat{\Theta}$  minimizes the SMM objective function:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} g(\Theta) \hat{W} g(\Theta)' \quad (104)$$

### B.4.2 Model Estimation Algorithm

I use the particle swarm algorithm to minimize the SMM objective function (104). The model is estimated as follows:

1. *Set initial guesses for model parameters:* I set initial values for the structural parameters  $\Theta$ . The initial guess is chosen manually, while subsequent guesses are selected

by the particle swarm algorithm.

2. *Compute the firm's value function:* Given a vector of parameters  $\Theta$ , I compute the value function  $V$  using the procedure outlined in Appendix B.3.2, from which the optimal firing boundary and compensation process can be computed.
3. *Simulate model:* Given the optimal firing and compensation policies, I simulate 5000 firms 20 times each. Firms draw an initial CEO from distribution  $N(\theta_0, \delta_0^2)$ . Performance evolves according to Equation (5) and beliefs evolve according to Equation (20). Firms make compensation and firing decisions based upon the optimal policies outlined in the second-best case of the model.
4. *Construct simulated panel and compute moments:* Using the simulated data, I construct a panel resembling the empirical sample and compute the same moments as described in Section 4.
5. *Evaluate objective function:* Given the set of simulated moments, I evaluate the SMM objective function (104). If the objective function value satisfies the particle swarm stopping criterion, the algorithm halts. Otherwise, a new candidate parameter vector  $\Theta'$  is selected and steps 2-5 repeat. This continues until the algorithm halts.

### B.4.3 Standard Errors for Parameter Estimates

For true parameter vector  $\Theta$  and consistent estimate  $\hat{\Theta}$ , we have the following asymptotic distribution (Duffie and Singleton, 1993):

$$\sqrt{n}(\hat{\Theta} - \Theta) \rightarrow^d N(0, avar(\hat{\Theta})) \quad (105)$$

$avar(\hat{\Theta})$  can be expressed as:

$$avar(\hat{\Theta}) = \left(1 + \frac{1}{S}\right) \left( \frac{\partial g(\Theta)}{\partial \Theta} W \frac{\partial g(\Theta)}{\partial \Theta'} \right)^{-1} \quad (106)$$

where  $\frac{\partial g(\Theta)}{\partial \Theta}$  is the Jacobian of the moment residual (103) with respect to the structural parameters,  $W$  is the optimal weighting matrix, and  $S$  is the number of simulations. I approximate the Jacobian using:

$$\frac{\partial \hat{g}_m(\Theta)}{\partial \Theta_p} = \frac{g_p(\hat{\Theta} + h_p) - g_p(\hat{\Theta})}{h_p} \quad (107)$$



for each moment  $m$  and parameter  $p$ .  $h_p$  is the perturbation size for parameter which I set to 1% of the absolute value of the parameter estimate. The standard errors are then obtained as the square root of the diagonal elements of the matrix:

$$\left(1 + \frac{1}{S}\right) \left( \frac{\partial \hat{g}(\Theta)}{\partial \Theta} \hat{W} \frac{\partial \hat{g}(\Theta)}{\partial \Theta'} \right)^{-1} \quad (108)$$

where  $\hat{W}$  is the sample counterpart of the optimal weighting matrix.

#### B.4.4 Re-Estimation with Director Data

I merge the ISS Director Data with my main sample to analyze the impact of board independence on entrenchment. The key challenge is that the director data covers only a subset of firm-years in my full sample; this information is missing for 65% of observations in my full sample. An easy route forward could be to simply estimate the model using the subset of data for which director information is not missing. However, the observations with missing director data still contain useful identifying information for other model parameters, and as such taking such an approach would discard valuable information. The alternative approach I take is as follows. First, I randomly classify 65% of firms in the simulation as “missing;” these correspond to the portion of the sample with no director information and will be subject to the entrenchment parameter  $\pi_{miss}$ . Conditional on not belonging to the “missing” group, I draw for each firm a share of independent directors  $s_j$  from its empirical distribution. As in the main model, I assume that board characteristics and therefore the level of entrenchment are fixed over time. I classify these firms into either the “high” or “low” group depending on whether their value of  $s_j$  is above or below the empirical median; “high” and “low” firms are respectively subject to the entrenchment parameters  $\pi_{high}$  and  $\pi_{low}$ . Once all firms have been assigned to a group, I solve and simulate the model as usual with the addition of separately solving the HJB equation for each of the three possible values of  $\pi_k$  for  $k \in \{\text{missing, high, low}\}$ .

To identify this model, I use the same moments as in the main estimation routine, but augment the forced turnover regression (Equation (109)) with indicators corresponding to each of the three firm-groups. In particular, I target the coefficients of the following equation:

$$d_{ijt} = \lambda_0 + \lambda_1 tenure_{ijt} + \lambda_2 tenure_{ijt}^2 + \lambda_3 missing_{ijt} + \lambda_4 low_{ijt} + \omega_{ijt} \quad (109)$$

where  $missing_{ijt}$  is an indicator = 1 if observation  $ijt$  lacks information on board independence and  $low_{ijt}$  is an indicator = 1 if the observation  $ijt$ 's share of independent directors

lies below the median of .83. Adding the coefficients  $\lambda_3$  and  $\lambda_4$  leaves me with 15 moments to identify the 11 structural parameters of the extended model. The associated weight matrix is computed using the same method described in the previous section. I provide a summary of model fit in the next section.