Why Are CEOs Rarely Fired?
Evidence from Structural Estimation

LUCIAN A. TAYLOR*

Abstract

I evaluate the forced CEO turnover rate and quantify effects on shareholder value by estimating a dynamic model. The model features costly turnover and learning about CEO ability. To fit the observed forced turnover rate, the model needs the average board of directors to behave as if replacing the CEO costs shareholders at least $200 million. This cost mainly reflects CEO entrenchment rather than a real cost to shareholders. The model predicts shareholder value would rise 3% if we eliminated this perceived turnover cost, all else equal. In addition, the model helps explain the relation between CEO firings, tenure, and profitability.

On average, 2% of CEOs at large U.S. corporations are fired each year (e.g., Kaplan and Minton (2006), Huson, Parrino, and Starks (2001)). This rate seems low, and it is tempting to conclude that CEOs are entrenched and boards of directors are not acting in shareholders' interests. However, the literature provides little guidance for making such judgments. For example, it is not clear what rate of forced CEO turnover we should expect from a well functioning board. Therefore, it is difficult to judge whether the observed 2% rate is low or high. If it is indeed too low, it is not clear how much shareholder value is being destroyed.

This paper's goals are to provide a benchmark for the CEO firing rate and to quantify the amount of shareholder value at stake. The benchmark is a dynamic model featuring a rational board that maximizes shareholder value. In the model, the board decides at each point in time whether to fire or keep its current CEO. Some CEOs are more skilled than others, meaning they can produce higher average industry-adjusted profits. If CEO turnover entails a real cost for shareholders— for instance, because searching for a replacement is costly — then firing an unskilled CEO is not always in shareholders' interests. Complicating matters, the board cannot directly observe CEO ability, but instead learns about it over time. The board learns by observing profits and an additional signal. At each point in time, the board observes the two signals, assesses the CEO's ability, and then decides optimally whether to replace him or her with a new CEO of uncertain ability.

The model indicates three potential reasons why boards rarely fire CEOs. First, the cost to shareholders may be large. Second, if there is not much difference between a good and bad CEO, then there is not much reason to replace one CEO with another. Finally, boards may learn slowly about CEO ability, allowing untalented CEOs to survive several years before being fired.

A fourth potential reason is CEO entrenchment. In this paper, CEOs are said to be entrenched if the board retains some CEOs whom shareholders would rather see fired, ex

post. Entrenchment occurs through two channels in the model. First, CEO turnover entails a personal utility cost to the board. Unlike the real turnover cost discussed previously, the personal cost does not directly affect firm profits. The personal cost reflects a distaste for firing CEOs, which may result because the board has personal or professional ties to the CEO, or because firing the CEO puts the directors' own jobs at risk, requires uncompensated effort, or hurts directors' chances of being nominated to other boards (Hermalin and Weisbach (1998)). The model predicts CEOs are more entrenched if the board faces a larger personal cost, all else equal. The second channel for entrenchment is the weight boards place on shareholder value. The model predicts CEOs are more entrenched when the board cares less about shareholder value, all else equal. I cannot disentangle these two channels empirically, so I label their total impact the "effective personal turnover cost."

Measuring the relative importance of these four reasons for infrequent firings presents a challenge. The board's turnover decisions are endogenous, which generates endogenous patterns in firm performance. There are no obvious instruments. Several elements are unobservable, including a CEO's actual and perceived ability, the CEO talent pool, the board's additional signals of CEO ability, and the board's distaste for firing CEOs. While we can measure the reasons' directional effects using reduced-form empirical techniques, evaluating their magnitudes requires estimating or calibrating an economic model.

These challenges lend themselves to a structural estimation approach. The structural approach infers unobservable quantities from endogenous patterns in firing decisions and firm performance, without requiring an instrument. With a structural approach we can assess not only the reasons' directional effects, but also their magnitudes. Moreover, the structural model allows us to explore counterfactuals. For instance, how much would shareholder value rise if there were less entrenchment?

I estimate the model's parameters by applying the simulated method of moments to data on firm profitability and both forced and voluntary CEO turnover in large U.S. firms

from 1971 to 2006. The estimated parameters include the real cost of CEO turnover to shareholders, the variation in ability across new CEOs, the volatility and persistence of profitability, the precision of boards' additional information about CEO ability, and the effective personal turnover cost.

Estimates indicate the model needs a very large total CEO turnover cost to fit the observed forced turnover rate. Boards behave as if firing a CEO costs shareholders an estimated 5.9% of the firm's assets, or \$254 million (\$1.3 billion) for the median (mean) sample firm. This estimate provides a benchmark for the forced CEO turnover rate: the observed 2% rate is low, in the sense that a very large turnover cost is needed to explain it.

The total CEO turnover cost is the sum of the real cost to shareholders and the effective personal utility cost to the board. I disentangle these costs using two assumptions: the personal cost does not directly affect profits, and the real cost shows up in profits in the two years around the succession. The total 5.9% cost breaks down into a 4.6% effective personal cost and a 1.3% real cost to shareholders. In other words, the board behaves as if firing the CEO costs shareholders 5.9% of assets, whereas it really only costs shareholders 1.3%. Because of this gap between the perceived and actual turnover cost, the board retains many CEOs whom shareholders would rather see fired, implying CEOs are highly entrenched. The large effective personal cost indicates either that CEO turnover is extremely costly to directors (in a utility sense), or that directors do not care much about shareholder value. The degree of entrenchment is significantly lower in recent years (1990 to 2006) and is slightly negative (although statistically insignificant) in larger firms, indicating that boards in large firms actually fire *more* CEOs than is optimal for shareholders, ex post. One potential explanation for this last result is that directors in large, poorly performing companies use the CEO as a scapegoat to protect their own jobs and reputations.

The disadvantage of the structural approach is that it imposes strong assumptions on the data. Three assumptions crucial for identification are the following: within-industry

variation in long-run average profitability is due entirely to variation in CEO ability; the firm turnover cost shows up in realized profits in the two years around a succession; and boards take into account only shareholder value and personal turnover costs when making CEO firing decisions. In the robustness section I show that alternate models with a flat firing threshold, firm fixed effects, and costless voluntary turnover produce personal turnover costs of 5.3%, 5.1%, and 5.9%, respectively. While the exact estimates differ, the conclusion is the same: the alternate models all need a high degree of CEO entrenchment to fit the data. Of course, I have not explored all possible models, and future research may produce a model that fits the data without requiring a high degree of entrenchment.

Parrino's (1997) algorithm, which I use to classify successions in the data as forced or voluntary, is conservative and probably underestimates the rate of forced CEO turnover. As a result, the true effective personal turnover cost is likely lower than the 4.6% estimate. Using an alternate algorithm that arguably overestimates the empirical firing rate, the estimated effective personal turnover cost drops to 1.3%. The true value is likely somewhere between 1.3% and 4.6%. Even the 1.3% lower bound amounts to a $292 ($56) million effective personal turnover cost for the average (median) sample firm, still implying a high degree of CEO entrenchment.

I find that entrenchment (i.e. the effective personal turnover cost) is positively correlated with three proxies for weak governance: early half of the sample (1970 to 1989 vs. 1990 to 2006), small firm size, and low fraction of outsiders on the board. I find no correlation with stock ownership by directors. In other words, for 3 of 4 proxies, entrenchment is stronger when we expect it to be, which provides a useful consistency check. However, this analysis says nothing about whether CEO entrenchment is bad for shareholders ex ante. While entrenchment harms shareholders ex post, it may benefit them ex ante, for instance, by allowing them to hire a better CEO.

If we believe entrenchment is bad for shareholders even ex ante, then how bad is it?

When I eliminate entrenchment by reducing the effective personal turnover cost from 4.6% to zero, holding all other parameters constant, the simulated CEO firing rate rises from 2% to 13% per year, average profitability rises by half a percentage point per year, and shareholder value rises by 3%. In other words, eliminating entrenchment would increase shareholder value by 3%, assuming the model and baseline parameter estimates are true, and assuming we could eliminate the personal turnover cost without affecting the other model parameters. The effect varies considerably across subsamples. For instance, eliminating entrenchment increases shareholder value by 7.7% in smaller firms, but by only 1.4% in the 1990 to 2006 subsample. In a separate experiment, I find that hiring a very bad CEO (ability two standard deviations below the mean) destroys 6.3% of shareholder value when entrenchment is at its estimated level, but destroys only 4.6% of value in a world with no entrenchment. These counterfactual experiments are a first step towards quantifying the effect of CEO entrenchment on shareholder value.

The estimated model fits several features of the data. The model predicts that 2.2% of CEOs are fired per year on average, which is close to the 2.3% rate in the data. Because of its dynamic setup, the model can also explain the timing of forced CEO turnover, which depends on how fast boards learn about CEO ability. In both the model and the data, the median fired CEO spends a total of four years in office. The model produces a V-shaped pattern in average profitability around CEO dismissals, which closely matches the empirical pattern. Profitability predicts firings with an R-squared of just 2% in the model and 3% in the data.

While there is a large empirical[1] and theoretical[2] literature on CEO turnover, this is the first study that estimates a structural model of CEO turnover. Therefore, this represents one of the first attempts at using an economic model to evaluate the forced turnover rate and to quantify the potential effects of suboptimal turnover decisions on shareholder value. As I discuss later, my model is consistent with several existing empirical findings, and the model produces new, untested predictions. My goal is not to make a theory contribution,

however, but to adapt and estimate existing models.

Hermalin and Weisbach (1998) is the most closely related model in the CEO turnover literature. Unlike their model, mine has a fully dynamic setup and a persistent performance signal. My model is simpler in that it does not feature endogenous monitoring, board composition, or CEO pay.

Eisfeldt and Rampini (2008) also use a model to evaluate magnitudes in the CEO turnover data. Their model focuses on CEOs' incentives for revealing private information, whereas mine focuses on learning about manager ability with symmetric information. They calibrate their model to match business cycle variation in CEO turnover and compensation, which I do not attempt to explain.

Like my model, the model of Dangl, Wu, and Zechner (2007) features learning about a manager's skill from his output, optimal firing decisions, costly turnover, and a dynamic setup. Unlike my contribution, theirs is purely theoretical, and their focus is mutual fund managers.

Miller (1984) estimates a model similar to mine using data from multiple occupations, such as farm workers. Our models both feature optimal firings, a dynamic setup, and learning about skill from the worker's output. However, there are several differences between our models, and we use different data, identification strategies, and estimation methods.

I present the model in Section I. Section II describes the data and estimation method. Section III presents estimation results, and Section IV discusses robustness. Section V concludes.

I. Model

A. Assumptions

The model features a firm that lives for an infinite number of periods, a large pool of potential CEOs, and a board that makes CEO firing decisions. I set one period equal to a year in the empirical implementation. The board of directors can fire the CEO and hire a new one at the beginning of each period. In addition, a CEO who has already served τ periods voluntarily leaves the firm (he either quits or retires) with exogenous probability $f(\tau)$.[3] The firm's book value of assets equals B_t at the beginning of period t.[4] The firm generates dollar profits equal to $Y_t B_t$ at the end of period t, so Y_t is the rate of firm profitability.[5] Profitability has three components:

$$Y_t = v_t + y_t - \mathbf{1}(turn_t)c^{(firm)}. \tag{1}$$

Component v_t is industry profitability. Indicator $\mathbf{1}(turn_t)$ equals 1 if CEO turnover occurs in period t, and otherwise equals zero. Parameter $c^{(firm)}$ is the cost of CEO turnover to shareholders, which I define later. Firm-specific profitability y_t mean-reverts around α, the current CEO's ability level:

$$y_t = y_{t-1} + \phi(\alpha - y_{t-1}) + \epsilon_t. \tag{2}$$

The shock ϵ_t is independently and normally distributed with mean zero and variance σ_ϵ^2. To be precise, equation (2) should include a CEO-specific subscript on ability α, because different CEOs can have different ability levels. A CEO's ability is constant over time. Equation (2) defines the model's notion of CEO skill: A CEO is considered highly skilled (i.e. high α) if he can achieve profitability higher than the industry, on average and in the long run.

Firm-specific profitability y_t follows a random walk when persistence parameter $\phi = 0$,

is iid when $\phi = 1$, and is mean reverting for $0 < \phi < 1$. I allow persistence in firm-specific profitability for two reasons. First, there is empirical evidence of persistence (Fama and French (2000)). More importantly, persistence allows a CEO to have long-lasting effects on profitability, which is plausible and affects the firing decision. For instance, after a CEO is fired for poor earnings performance, earnings may continue to be low for a few years even if the replacement CEO is highly skilled. During those years, the board would not want to penalize the new CEO for the old CEO's mistakes.

There are two CEO turnover costs in the model: a real cost to shareholders ("firm cost"), denoted $c^{(firm)}$, and a cost to the board ("personal cost"), denoted $c^{(pers)}$. The firm cost includes severance or retirement packages, fees to executive search firms, disruption costs, and any other CEO turnover costs that affect profits. The personal cost does not directly affect the firm's profits, but does affect the board members. The personal cost includes the directors' loss of the CEO as an ally both within the firm and in their careers outside the firm, any uncompensated effort and stress from the succession process, and in the case of forced turnover, reputation costs or benefits from "rocking the boat" (Hermalin and Weisbach (1998)). I show later that if the personal cost is positive, then the board does not fire some CEOs whom shareholders would rather see fired. For this reason I interpret the personal cost as a measure of CEO entrenchment. I assume forced and voluntary CEO turnover are equally costly. For robustness, in Section IV I solve and estimate the model assuming voluntary turnover is costless. Both $c^{(firm)}$ and $c^{(pers)}$ are a constant fraction of the firm's book assets. This assumption is motivated from empirical evidence that costs such as separation pay and executive search fees increase with firm size.[6]

The board makes firing choices $d_t \in \{$fire CEO, keep CEO$\}$ that maximize the board's lifetime utility U_t :

$$\max_{\{d_{t+s}\}_{s=0}^{\infty}} U_t = E_t \left[\sum_{s=0}^{\infty} \beta^s u_{t+s} \right]. \tag{3}$$

Parameter β is both the board's and investors' discount factor, with $0 < \beta < 1$. Per-period

utility is

$$u_t = \kappa B_t Y_t - \mathbf{1}(turn_t) B_t c^{(pers)}. \tag{4}$$

The board prefers higher profits (the first term) and experiences a personal cost from CEO succession (the second term). The constant $\kappa > 0$ controls the degree to which the board internalizes shareholder value. For instance, we might believe κ is higher when directors have a greater sense of fiduciary responsibility, own more shares or options, or receive greater reputation benefits from the firm's success. I show later that κ affects the degree of CEO entrenchment. CEO firing choices affect the board's utility in two ways. They affect profitability Y_t via the acting CEO's ability and the firm turnover cost $c^{(firm)}$. Second, the board incurs an additional personal cost $c^{(pers)}$ each time it fires the CEO. Since utility is only defined up to a positive constant, there is an indeterminacy between κ and $c^{(pers)}$. I divide (4) by κ and call $c^{(pers)}/\kappa$ the effective personal cost. This cost is in units of board utility, so we can interpret its magnitude only in terms of indifference relations with shareholder profits, the other quantity in the utility function. Loosely speaking, the effective personal cost measures not how much the board cares about losing its own dollars, but rather how much the board cares about seeing shareholders lose their dollars.

Substituting (4) into (3), we can decompose the board's objective function into two terms:

$$U_t = \kappa E_t \left[\sum_{s=0}^{\infty} \beta^s B_{t+s} Y_{t+s} \right] - E_t \left[\sum_{s=0}^{\infty} \beta^s \mathbf{1}(turn_{t+s}) B_{t+s} c^{(pers)} \right] \tag{5}$$

The first term in equation (5) is κ times the board's assessment of shareholder value. The second term is the net present value of personal turnover costs. Therefore, the board maximizes shareholder value when and only when there is no personal cost ($c^{(pers)} = 0$), at least in an ex post sense[7].

The board can observe all parameters but cannot observe a CEO's ability α. Therefore, when the board observes high firm-specific profitability, it cannot be sure whether this is due to CEO skill (i.e. high α) or luck (i.e. high ϵ_t). When the board hires a new CEO,

it starts with normally distributed prior beliefs about his ability: $\alpha \sim \mathcal{N}(\mu_0, \sigma_0^2)$. The board's prior beliefs match the distribution of ability in the CEO talent pool. Therefore, parameter σ_0 plays two roles. It is both the initial uncertainty about a newly hired CEO's ability, and also the dispersion in ability across replacement CEOs. Each period, the board uses Bayes' Rule to update its beliefs about CEO ability, using two signals: firm-specific profitability y_t, and z_t, which is an additional, latent, orthogonal signal of CEO ability. Signal z_t includes all additional information held privately in the firm (e.g. the CEO's specific actions and choices, the performance of individual projects, the CEO's strategic plan) as well as public information (e.g. growth prospects, stock prices, market share, discretionary earnings accruals). I assume this additional signal is independent of profitability, normally distributed, and centered at the CEO's ability: $z_t \sim \mathcal{N}(\alpha, \sigma_z^2,)$. The signal z is more precise when its volatility σ_z is lower.

Like all models, this model presents a simplified view of the world. The simplifications allow me to obtain predictions from the model and identify parameter values from the data. Section IV describes several elements which are missing from my model, including firm fixed effects, time-varying entrenchment, CEO learning on the job, fluctuating CEO skill, board risk aversion, and earnings manipulation.

B.   Model Solution and Predictions

To solve the model, I first use Bayes' Rule to derive the board's beliefs about CEO ability. I substitute these beliefs into the board's objective function to obtain the Bellman equation, which I solve numerically. I obtain additional predictions by simulation. Details are in the Internet Appendix*.

Next I discuss the model's predictions about the board's firing policy, the frequency and timing of firings, and the relation between turnover and profitability. These predictions hold for a wide range of plausible parameter values. However, since I solve the model numerically,

I do not present these predictions as formal propositions.

B.1. The Board's Firing Policy

The board's posterior mean belief about the CEO's ability fluctuates over time as signals arrive. The board fires the CEO as soon as the posterior mean drops below an endogenous threshold. The threshold depends on all model parameters and the CEO's tenure. Raising the total turnover cost $c = c^{(firm)} + c^{(pers)}/\kappa$ shifts the firing threshold down, making firings less likely. Intuitively, when firing the CEO is more costly, the CEO must have lower perceived ability to make firing him worth it. This result does not depend on whether turnover cost c is larger due to a higher firm cost $c^{(firm)}$ or higher effective personal turnover cost $c^{(pers)}/\kappa$.

B.2. The Frequency and Timing of Forced CEO Turnover

I illustrate the model's predictions using the following parameter values: discount factor $\beta = 0.9$, prior mean ability $\mu_0 = 1\%$, prior standard deviation of ability $\sigma_0 = 2\%$, volatility of profitability shocks $\sigma_\epsilon = 3\%$, total CEO turnover cost $c = 3\%$, persistence parameter $\phi = 0.12$, and volatility of the board's additional signal $\sigma_z = 7\%$. These parameter values are close to the empirical estimates in Section III. I assume CEOs retire if they complete 15 years, but not before then.

The top panel of Figure 1 plots firing hazard rates vs. tenure for three values of total turnover cost c. Not surprisingly, firing rates are lower and more CEOs survive to retirement when the total turnover cost is higher. Consistent with this prediction, Parrino (1997) finds that forced CEO turnover is more frequent in homogenous industries. As Parrino notes, the real costs of firing a CEO are probably lower when the firm can find a replacement in a similar firm. Turnover costs also affect the timing of forced turnover, with firings occurring

later when the turnover cost is higher. Hazard rates are downward sloping when $c = 0\%$, hump-shaped when $c = 3\%$, and upward sloping when $c = 5\%$. Intuitively, the board is cautious and waits for more information when firing the CEO is more costly.

INSERT FIGURE 1 NEAR HERE

The bottom panel of Figure 1 plots firing hazard rates for three values of prior uncertainty about CEO ability, σ_0. Forced turnover is more likely when σ_0 is higher, meaning there is more variation in ability across CEOs. To my knowledge, this prediction has never been tested. Intuitively, if CEOs are roughly alike in terms of their ability (i.e., low σ_0), then there is not much incentive to replace one CEO with another, especially when doing so is costly. Firings occur later when prior beliefs are tighter (i.e. lower σ_0), since the board learns more slowly. Supporting this intuition, the model also predicts that firings occur later when the additional signal's volatility σ_z is higher, profit volatility σ_ϵ is higher, and persistence parameter ϕ is lower, all of which cause the board to learn more slowly. To my knowledge, no one has tested whether firings occur later when boards learn more slowly about CEO ability.

B.3. Profitability around CEO Dismissals

Figure 2 plots predicted average firm-specific profitability in event time around CEO dismissals (solid line). The figure also shows the average across CEOs of $\mu_t = E_t[\alpha]$, the board's posterior mean of the CEO's ability α (dashed line). The posterior mean drops gradually leading up to forced turnover at time zero. This must occur in order for the posterior mean to drop below the firing threshold at time zero. The posterior mean jumps up to the prior mean μ_0 following dismissals, because the firm hires a new CEO and resets its beliefs.

INSERT FIGURE 2 NEAR HERE

Average realized excess profitability has a V shape around dismissals, consistent with the empirical evidence of Huson, Malatesta, and Parrino (2004). To cause downward revisions in posterior mean skill μ_t, either profitability y_t or the signal z_t must repeatedly be lower than expected. On average, excess profitability y_t drops before forced CEO turnover at time zero. Consistent with this prediction, Weisbach (1988), Huson, Parrino, and Starks (2001), and many others find that CEO turnover is more likely following low profitability. The model predicts a rise in profitability after forced turnover, for two reasons. First, the CEO's replacement has higher ability, on average. Second, realized profitability is typically below the fired CEO's perceived ability, a prediction similar to the one in Pástor, Taylor, and Veronesi (2007). Profitability rises gradually rather than abruptly because of the assumed persistence in profitability.

Profitability is not a perfect predictor of CEO dismissal in the model, because the board also uses the additional signal z_t to evaluate ability. Even if a CEO achieves very low profitability, he may avoid being fired if the z_t signal is high enough, and vice-versa. The top panel of Figure 3 shows average changes in excess profitability in event time around forced turnover for three values of σ_z, the additional signal's volatility. The drop in profitability is smaller when the signal z is more precise, i.e., its volatility σ_z is lower. The reason is that the board rationally relies more on z_t and less on profitability to evaluate CEO ability, so the board is less likely to fire the CEO when profitability is low, which produces a smaller average drop in profitability around CEO dismissals.

INSERT FIGURE 3 NEAR HERE

The firm turnover cost also affects the turnover-profitability relation. The bottom panel of Figure 3 shows that profitability drops more around forced turnover when firm turnover cost $c^{(firm)}$ is larger, holding constant the total turnover cost. This prediction follows mechanically from the assumption that the firm turnover cost reduces profitability by $c^{(firm)}$ in

the turnover period.

In sum, Figure 3 shows two reasons why profitability drops around CEO dismissals: (1) Low profitability causes a dismissal (top panel), and (2) dismissal causes low profitability (bottom panel). Reason (1) is due to learning, and reason (2) is due to the firm turnover cost.

To my knowledge, no one has tested Figure 3's prediction that the turnover-profitability relation is stronger when the board's additional signals of CEO ability are less precise. The model also predicts that the turnover-profitability relation is stronger when profitability is less volatile (low σ_ϵ) and less persistent (high ϕ), both of which make profitability a more precise signal. Consistent with these predictions, Engel, Hayes, and Wang (2003) find a stronger empirical turnover-profitability relation when profitability is less volatile and more sensitive to CEOs' actions. Also, Wu and Zhang (2008) find a stronger turnover-profitability relation after firms adopt international accounting standards that improve transparency.

II.   Estimation

A.   Data

Data come from Compustat and a CEO turnover database constructed using the method of Huson, Parrino, and Starks (2001). The sample consists of CEOs in the *Forbes* annual compensation surveys who left office from 1971 to 2006. To my knowledge, this sample spans more years than any sample from the existing CEO turnover literature. I am grateful to Robert Parrino for providing CEO turnover data from 1971 to 1994. CEO successions are classified as forced or unforced, as described in Huson, Parrino, and Starks (2001), p.2273:

First, if the *Wall Street Journal* reports that the CEO is fired, forced from the position, or departs due to unspecified policy differences, the succession is classi-

fied as forced. For the remaining cases, the succession is classified as forced if the departing CEO is under the age of 60 and the *Wall Street Journal* announcement of the succession (1) does not report the reason for the 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. The circumstances surrounding the departures of the second group are further investigated by searching the business and trade press for relevant articles in order to reduce the likelihood that a turnover is incorrectly classified as forced. 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.

I discuss this algorithm's limitations in the robustness section.

I interpret one model period as a year, and I assign successions to either the most recent past or soonest future fiscal year end, depending on which is closest to the actual turnover date. This timing convention eliminates from the sample fewer than 2% of CEOs who left office after 6 months or less. I use data from all years a CEO spent in office, including years before 1971. There are many ways one could sample the data, which raises the possibility of estimation bias. My simulation estimator solves this problem by sampling the real and simulated data in the same way. Additional details are in Section II.C. The sample does not include CEOs who left office due to a takeover, whose successions were not announced in the *Wall Street Journal*, or who have missing data in Compustat. I set firm profitability Y_t equal to the firm's return on assets (ROA) in year t.[8] Industry average profitability v_t is an equal-weighted average of ROA across firms in each of 12 industries defined on Kenneth French's website. When computing v_t I use each year's 1,000 largest Compustat firms (by lagged assets) to avoid bias from changes in Compustat's coverage.

Table I contains summary statistics. The full sample contains 981 CEOs and 7,325 firm/year observations. The Forbes surveys contain at least 2,500 CEOs from 1971 to 2005; I eliminate many of these CEOs because I cannot classify their succession as forced or voluntary, or because Compustat data are missing[9]. Out of the 981 successions in my sample, 168 (17.1%) are forced. On average, 2.29% of sample firms fire their CEOs in a given year. The annual forced CEO turnover rate varies across industries, e.g. 3.8% in the business equipment industry versus 0.7% in the chemicals industry. Forced turnover became more common from 1970 to 2006, increasing from 7.8% of successions from 1970 to 1974 to 24.7% of successions from 2005 to 2006. Kaplan and Minton (2006) also document an increase in forced turnover rates. I define a CEO spell as all the years a CEO served in office. The median CEO spell lasted 6 years (Panel B). Median spell length was shorter for CEOs forced out of office (4 years) compared to CEOs who left voluntarily (7 years).

INSERT TABLE 1 NEAR HERE

B. Identification and Additional Restrictions

This subsection discusses additional assumptions which deliver an identified model and improve precision. I also provide intuition for how the model is identified.

Solving the board's optimization problem requires parameters to be constant over a firm's lifetime, an assumption I maintain here. I also assume parameters are constant across firms. This same assumption is implicit in much of the reduced-form empirical work on CEO turnover. For instance, when Kaplan and Minton (2006) estimate a probit model of CEO turnover on lagged firm performance, they implicitly assume the probit slopes are constant across all firms and years. Later, I estimate the model in subsamples to allow some heterogeneity. Introducing heterogeneity across firms is challenging, because there are only 168 forced successions in my sample.

I set the discount factor β to 0.9, a plausible value given investors' annual discount factor. I try other β values later for robustness.

Next I provide intuition for how the remaining parameters are identified. Time-series autocorrelation in firm-specific profitability helps identify ϕ, the profitability persistence parameter. After removing its persistent component, profitability volatility within CEO spells helps identify σ_ϵ, the volatility of profitability shocks. Dispersion in average profitability across CEOs helps identify dispersion in CEO ability, σ_0. Intuitively, the more variation there is in ability across CEOs, the more variation we should see in average profitability across CEOs. To see this point more clearly, imagine we work with the subsample containing only CEOs' first year in office. (I consider the first year only to provide intuition, not for actual estimation.) In this subsample the board has made no turnover decisions yet, so we avoid complications from endogenous dismissals. For each CEO i in this subsample we can compute persistence-adjusted profitability $X_{i,0}$:

$$X_{i,0} = \left(y_{i,0} - y_{i,-1}\right)/\phi + y_{i,-1} = \alpha_i + \epsilon_{i,0}/\phi. \tag{6}$$

The variance of $X_{i,0}$ across CEOs equals $\sigma_0^2 + \sigma_\epsilon^2/\phi^2$. Knowing ϕ and σ_ϵ, we can back out σ_0, the dispersion in ability across CEOs. We also see that the average level of profitability helps identify prior mean ability μ_0: the mean of $X_{i,0}$ across CEOs equals μ_0, since shocks $\epsilon_{i,0}$ have mean zero. These estimators for μ_0 and σ_0 would be consistent but not efficient, since they only use data from CEOs' first year in office. The full estimation procedure described in the next section uses data from all years a CEO spends in office.

As we saw above, the variation in ability across CEOs (σ_0) is identified off variation in profitability across CEOs. In the model and estimation procedure, there are three reasons why average profitability varies across CEOs. First, CEOs have different ability levels. This is the variation we want to isolate when measuring σ_0. Second, CEOs may belong to industries with different profit margins or accounting rules. As prescribed by the model,

the estimation procedure uses industry-adjusted profitability data in order to remove this variation. Third, persistence in profitability generates differences in average profitability across CEOs, at least in the short term. If two firms have different past profitability, then they will have different expected future profitability, even if their CEOs have equal ability. The estimation procedure takes this persistence into account. Of course, there are additional reasons outside the model why average profitability varies across CEOs. Two robustness tests in Section IV address this concern, but I cannot completely solve this problem.

Next I discuss identification of the cost parameters and σ_z, the additional signal's volatility. The frequency of forced turnover at different tenures helps identify the total turnover cost $c = c^{(firm)} + c^{(pers)}/\kappa$. As we saw in Figure 1, forced turnover is less frequent when the cost c is higher. Changes in average firm-specific profitability around CEO dismissals help disentangle the firm turnover cost $c^{(firm)}$ and effective personal cost $c^{(pers)}/\kappa$, and also help identify signal z's volatility, σ_z. As Figure 3 suggests, we can pin down these parameters by measuring how much profitability drops around CEO dismissals, and decomposing the drop into a gradual component due to σ_z and an abrupt component due to firm cost $c^{(firm)}$. This identification strategy exploits two assumptions embedded in equation (1): first, the firm turnover cost has a direct effect on profits, but the personal turnover cost does not; and second, the firm turnover cost is realized in the period when the CEO is fired. Since successions do not perfectly line up with fiscal year ends in my data, I assign half the firm cost to the outgoing CEO's last year and half to the new CEO's first year. Once we know the firm cost $c^{(firm)}$ and the total cost c, we can back out the effective personal cost: $c^{(pers)}/\kappa = c - c^{(firm)}$. The ratio $c^{(pers)}/\kappa$ is identified, but $c^{(pers)}$ and κ are not. In other words, the model cannot distinguish between a board with a strong distaste for firing the CEO (large $c^{(pers)}$) and a board that cares little about shareholder value (low κ).

I constrain the firm turnover cost $c^{(firm)}$ to be non-negative. This constraint binds in 4 out of 15 specifications below. A negative firm turnover cost would imply that CEO turnover increases firm profits, for reasons unrelated to CEO ability. Given all the reasons

18

why turnover decreases profits (e.g. severance or retirement packages, executive search costs, turnover in lower ranks, general disruption), a negative firm turnover cost seems unlikely. As I discuss later, a negative effective personal turnover cost may be more reasonable, so I do not constrain this parameter.

C. Estimation Method

I estimate the parameters $\theta = \left\{ \mu_0, \sigma_0, \sigma_z, \sigma_\epsilon, \phi, c^{(firm)}, c^{(pers)}/\kappa \right\}$ using the simulated method of moments (SMM)[10]. Like GMM, SMM estimates parameter values by matching certain data moments and model-implied moments as closely as possible. Whereas GMM uses closed-form expressions for the model-implied moments, SMM estimates the model-implied moments using simulations. This project lends itself to SMM estimation, because few closed-form expressions are available. The SMM estimator $\widehat{\theta}$ is

$$\widehat{\theta} = \arg \min_\theta \left(\widehat{M} - \frac{1}{S} \sum_{s=1}^{S} \widehat{m}^s \left(\theta \right) \right)' W \left(\widehat{M} - \frac{1}{S} \sum_{s=1}^{S} \widehat{m}^s \left(\theta \right) \right). \qquad (7)$$

I set W equal to the efficient weighting matrix, which is the inverse of the estimated covariance of moments M. $\widehat{M}$ is a vector of moments estimated from the empirical data, and $\widehat{m}^s \left(\theta \right)$ is the corresponding vector of moments estimated from the sth sample simulated using parameters θ. Since my empirical sample contains 981 CEO spells, each simulated sample contains 981 CEO spells as well. Michaelides and Ng (2000) find that using a simulated sample 10 times as large as the empirical sample generates good small-sample performance. I use $S = 20$ simulated samples to be conservative. I obtain a simulated sample by simulating enough firms until I have 981 CEO spells. I must sample the simulated data the same way I sample the actual data, otherwise the empirical and simulated moments will measure different quantities. Following the empirical sampling method described in Section II.A, I collect all CEOs who left the simulated firm between simulation years 1971 and 2006, and I bring in profitability data from all years the CEOs spent in office, including years before

1971. I begin simulations in year 1900 to ensure the model has reached a steady state by 1971. Following Rust (1994) and Hennessy and Whited (2005, 2007), I use a simulated annealing optimization algorithm to avoid local minima of (7). Additional details are in the Internet Appendix.

I estimate the 7 parameters using 14 moments in vectors M and m. The first seven moments are coefficients from the pooled regression

$$y_{it}^* = \lambda_0 + \lambda_1 y_{i,t-1}^* + \Delta^{(-2)} + \Delta^{(-1)} + \Delta^{(0)} + \Delta^{(1)} + \Delta^{(2)} + \delta_{it}. \tag{8}$$

y_{it}^* is firm i's profitability Y_{it} minus industry profitability v_t. Coefficient $\Delta^{(k)}$ is a fixed effect for whether the CEO of firm i was fired in year $t + k$. The intercept λ_0 helps pin down the prior mean skill μ_0, and the AR1 coefficient λ_1 helps pin down the persistence parameter ϕ. The fixed effects $\Delta^{(k)}$ measure changes in average profitability around CEO dismissals. As discussed in Section II.B, these changes help identify z_t's volatility σ_z and firm turnover cost $c^{(firm)}$. The eighth moment is $Var(\delta_{it})$, the variance of the residual from equation (8), which is informative about the time series volatility of profitability, σ_ϵ.

The next four moments are forced turnover hazard rates. I define $h^{(k)}$ to be the % of CEOs fired per year in tenure category (k) years, conditional on the CEO reaching (k). I use hazard rates $h^{(1-2)}$, $h^{(3-4)}$, $h^{(5-7)}$, and $h^{(8+)}$. These four rates help identify the total turnover cost.

The last two moments disentangle σ_ϵ and σ_0. Both moments use data on persistence-adjusted profitability $\widehat{X}_{it} = \left(y_{it}^* - \widehat{\lambda}_1 y_{it-1}^*\right) / \left(1 - \widehat{\lambda}_1\right)$, where $\widehat{\lambda}_1$ is estimated in regression (8). First, for each CEO j I compute $E_j[X_{it}]$ and $Var_j(X_{it})$, respectively, the mean and variance of $\widehat{X}_{it}$ across all the years CEO j spent in office. The 13th moment is $E\left[Var_j\left(X_{it}\right)\right]$, the mean of CEOs' variances. Since this moment removes the effect of each CEO's ability, it is most informative about σ_ϵ, the time-series volatility of profitability. The 14th moment is $Var(E_j[X_{it}])$, the variance of CEOs' means. This moment is most informative about σ_0, the

20

dispersion in ability across CEOs, because it measures the cross-CEO dispersion in a proxy for ability, namely, the CEO's average realized profitability.

Finally, the hazard function for voluntary turnover, $f(\tau)$, is an input to the model. I pool all CEO spells in the sample and estimate $f(\tau)$ as the frequency of voluntary turnover after τ years conditional on the CEO surviving $\tau - 1$ years.

III.   Empirical Results

A.   Parameter Estimates

Parameter estimates from the baseline model are in Table II. Section IV describes how estimates change when I alter the model's assumptions. The estimated turnover cost to the firm, $c^{(firm)}$, is 1.33% of the firm's assets, with 95% confidence interval [0.1%, 2.5%]. The firm turnover cost is $57 ($292) million in 2009 dollars for the median (mean) sample firm. We can compare this estimated cost to known succession costs. Yermack (2006) reports average CEO separation payments of $18 million and $2 million for forced and voluntary successions, respectively. Fees to executive search firms are on the order of $1 million[11]. Other costs are difficult to measure. Succession at the CEO level may result in costly succession at lower levels. It may take time for the new management team to learn the business, develop relationships inside and outside the company, and so on. If we add up all these succession costs, the estimated $57 million median firm cost is not unreasonable.

INSERT TABLE II NEAR HERE

The estimated effective personal turnover cost, $c^{(pers)}/\kappa$, is 4.61% of the firm's assets (95% confidence interval: [3.5%, 5.7%]), or $197 million ($1.03 billion) for the median (mean) firm. This cost is in units of board utility, so we can interpret its magnitude only in terms of indifference relations with shareholder profits, the other quantity in the board's utility

function. In other words, a \$1.03 billion personal cost should not be interpreted as \$1.03 billion leaving directors' pockets. The correct interpretation is as follows. Adding the firm cost and effective personal cost, the board behaves as if firing the CEO costs *shareholders* $c^{(firm)} + c^{(pers)}/\kappa = 5.9\%$ of the firm's assets. However, firing the CEO really only costs shareholders $c^{(firm)} = 1.3\%$ of assets. The gap, which equals the effective personal cost, makes the board retain some CEOs whom shareholders would rather see fired. In this sense, the large effective personal cost implies a high degree of CEO entrenchment. Consistent with severe entrenchment, CEOs had considerable influence on the choice of directors during much of the period I investigate[12]. Since I can only estimate the ratio $c^{(pers)}/\kappa$, I cannot determine whether this entrenchment is due to a strong distaste for firing CEOs (large $c^{(pers)}$) or a board that does not care much about shareholder value (low κ).

Parameter σ_0 is both the standard deviation of ability across new CEOs, and also the uncertainty about a newly hired CEO's ability. The estimate of σ_0 is 2.42% of assets per year. Using this estimate to compare new CEOs at the 5th and 95th percentiles of ability, the difference in average firm profitability they generate is $2 \times 1.65 \times \sigma_0 = 8.0\%$ per year. This difference suggests that CEO skill matters greatly. For comparison, Bertrand and Schoar (2003) estimate manager-specific fixed effects in annual ROA. They find a 7% standard deviation in fixed effects across managers, implying even greater dispersion in ability than reported here.

Parameter σ_0 likely understates the dispersion in CEOs' total contribution to firm profits. A CEO's total contribution is split between the CEO (in the form of compensation) and shareholders (in the form of profits net of compensation). Since profits in my model and data are net of CEO compensation, parameter σ_0 measures dispersion only in shareholders' portion of CEOs' contribution. Graham, Li, and Qiu (2009) estimate the dispersion in manager fixed effects in CEO compensation, which is the remaining portion of CEOs' contribution. Taylor (2009) estimates how CEOs and shareholders split the CEO's contribution to firm profits.

The estimated persistence parameter ϕ is 0.125, indicating that firm-specific profitability nearly follows a random walk. One reason for this low ϕ value is that I do not include firm fixed effects in profitability. For robustness, in Section IV I introduce firm fixed effects in profitability, which results in a higher estimate of ϕ.

To interpret the estimate of σ_z, the volatility of the board's additional signal, I compare the influence of the profitability signal and additional signal z_t on the board's beliefs about CEO ability. Specifically, I compare the change in posterior beliefs resulting from a one standard deviation z shock and a one standard deviation profitability shock. In the Internet Appendix I show that the response to the z shock is $P = \sigma_\epsilon/(\phi\sigma_z)$ times larger than the response to the profitability signal shock. The P ratio indicates that the additional z signal is more influential when it is more precise (σ_z lower), and when profits are noisier (σ_ϵ higher) or more persistent (ϕ lower). Applying the delta method, I obtain an estimate of P equal to 5.3, with a standard error of 0.3. In other words, the additional signal z has a 5.3 times larger influence on the board's beliefs, compared to the profitability signal. This result implies boards rely heavily on non-earnings information when evaluating the CEO. Consistent with this result, Bushman, Indjejikian, and Smith (1996) find that boards give considerable weight to information besides earnings and stock performance when determining a CEO's bonus. It is plausible that boards use this additional information in firing decisions as well.

B.   Model Fit

Next I assess how well the estimated model fits empirical patterns in forced CEO turnover and firm profitability. The first test is a formal test of the overall model. Since I estimate 7 parameters using 14 moments, the SMM procedure delivers a χ^2 test of over-identifying restrictions (bottom of Table III). The p-value rejects at the 1% confidence level the hypothesis that all 14 simulated moments equal the empirical moments. In other words, the data reject the model. I do not consider this result particularly damning, since we can reject

any model with enough data.

Next I examine the 14 moments individually to gauge where the model fails. Each row in Table III shows a moment's empirical estimate, simulated value, standard error, and a p-value that tests whether the empirical and simulated moments are equal. For 4 out of 14 moments we can reject equality at the 5% confidence level; only two of these are reliably different at the 1% level. The model matches the intercept and AR slope from the profitability regression very closely, indicating the model can fit the long-run mean and persistence of firm-specific profitability. The fixed effects Δ measure average changes in profitability around CEO dismissals. The model matches all the Δ's except $\Delta^{(1)}$ fairly closely. I examine this pattern in detail later.

Turning to the forced turnover hazard rates $h^{(k)}$, the model produces too few firings in the first two years, too many firings in years 3 and 4, and not enough firings after year 7. However, the gap between simulated and empirical hazard rates is less than 1% per year for all four moments, and the model successfully produces the hump-shaped empirical relation between tenure and firings, which Allgood and Farrell (2003) also document. The top panel of Figure 4 examines this pattern more carefully, plotting the % of CEOs fired vs. tenure. The empirical turnover rates are within the model's grey 95% confidence region at almost all tenure levels.

Returning to Table III, there are no significant differences between the empirical and simulated values for any of the second moments of profitability. The model appears to match time-series volatility in profitability for a given CEO ($Var(\delta)$ and $E[Var(X)]$), and also the variation in average profitability across CEOs, $Var(E[X])$.

More measures of model fit are in Table IV. The model exactly matches the median

tenure of CEOs forced out of office (4 years) and CEOs who leave voluntarily (7 years). In the real data, 17.1% of successions are forced, compared to 16.2% in the model. The model matches the forced turnover rate quite closely, predicting that 2.16% of CEOs are fired each year on overage, compared to 2.29% in the real data. On these dimensions, the model fits the data quite well.

INSERT TABLE IV NEAR HERE

Next I examine the relation between profitability and forced CEO turnover. First, I estimate a probit model which use lagged firm-specific profitability to forecast whether a CEO is fired. I use one year of lagged profitability, although results are similar using three lags. The last columns in Table IV compare results when I estimate the probit model using real and simulated data. Profitability has a statistically significant negative slope in both the real and simulated data, indicating that CEOs are more likely to be fired after a year of low profitability. The estimated slopes from the real and simulated data are within two standard errors. The low pseudo R-squared values (2% from the simulated data, 3% from the real data) indicate profitability poorly predicts forced turnover. The model generates a weak profitability-firing relation because profitability has a relatively small influence on the board's beliefs about CEO ability, as discussed in Section III.A.

The bottom panel of Figure 4 shows average firm-specific profitability in event time around CEO dismissals, comparing real and simulated data. As expected, there is a V-shaped pattern, both in the real and simulated data. The empirical pattern is within the model's grey 95% confidence region in all event years except year –5. Overall, the model can closely match the level of profitability, the magnitude of the changes in profitability, and the timing of the changes. The model has the most difficulty matching the change in profitability in the year after the CEO is fired (event years 0 to 1). In that year the model predicts a rise in profitability, whereas profitability drops in the data.

C. CEO Entrenchment and Shareholder Value

A central result so far is that the model needs a high degree of CEO entrenchment (i.e. a large effective personal turnover cost to the board) to fit the data. This section examines the relation between CEO entrenchment and shareholder value. I start by providing possible interpretations of the personal turnover cost. I then examine the relation between the personal turnover cost and measures of governance quality, and finally I quantify the effect of entrenchment on shareholder value.

I offer two extreme interpretations of the large effective personal turnover cost. The bad-governance interpretation is that the personal cost prevents boards from acting in share-holders' interests, even ex ante. Shareholders prefer but for some reason fail to obtain a board that is more willing to fire the CEO. The good-governance interpretation is that the personal cost and resulting CEO entrenchment are optimal for shareholders ex ante. By electing a board with a large personal turnover cost, shareholders commit up front to a low probability of firing the CEO. This commitment may benefit shareholders by allowing them to pay the CEO less (Almazan and Suarez (2003), Hermalin and Weisbach (2008)), hire from a better CEO talent pool, or provide the CEO with incentives to take risks (Manso (2007)). According to this view, the estimated 4.6% effective personal cost delivers the level of entrenchment that is optimal for shareholders ex ante.

According to both the good- and bad-governance interpretations, the effective personal cost should be larger in firms with weaker governance. To test this hypothesis, I examine whether the effective personal cost is related to measures of governance quality. I split the sample using a measure of governance quality, estimate the model independently in each subsample, and then test whether personal costs are equal across the sub-samples[13].

I use four governance measures. First, I create large- and small-firm subsamples by comparing firms' inflation-adjusted assets to the sample median. If shareholders face a fixed cost of monitoring a board, then they have a stronger incentive to monitor boards of larger

firms, which make up more of their portfolio. Next, I create early and late subsamples based on whether the CEO left office between 1971 and 1989 or 1990-and 2006. Pointing to time trends in board composition and size, director compensation, and institutional stock ownership and activism, Huson, Parrino, and Starks (2001) argue that monitoring became stronger from 1970 to 2000. Monitoring continued to intensify due to the Sarbanes-Oxley Act in 2002. Third, I form two subsamples based on whether the % of the firm's shares owned by non-CEO officers and directors is above or below the median value, 1.31%.[14] Boards owning more shares arguably care more about shareholder value, which in the model means they have a higher value of κ and hence a lower effective personal cost, $c^{(pers)}/\kappa$. Finally, I form two subsamples based on whether the % of directors who are not officers of the firm is above or below the median value, 72.7%. Fama and Jensen (1983) and Weisbach (1988) argue that outsider-dominated boards monitor management more closely.

Parameter estimates for the subsamples are in Panel A of Table V. Panel B compares CEO tenures, firing rates, and firing sensitivities across subsamples and also between the real and simulated data.

INSERT TABLE V NEAR HERE

Consistent with the hypothesis that CEO entrenchment is less severe in firms with stronger governance, the effective personal cost is significantly lower in large firms than in small firms (–0.29% compared to 8.53%, difference[15] has $t = -6.8$). The effective personal cost estimate for large firms is slightly negative but not statistically different from zero ($t = -0.4$). An effective personal cost of zero implies that firing decisions within the model are optimal for shareholders, ex post. Section IV explains that the personal cost estimates are likely biased upward due to the forced/voluntary classification algorithm. Adjusting for this bias would make the personal cost for large firms even more negative than reported here. A negative effective personal cost implies that boards fire *more* CEOs than is optimal for shareholders, ex post. First I explain the mechanics of this surprising result, then I offer a

potential economic explanation.

There are two main reasons for the negative personal cost estimate in large firms. First, dispersion in average profitability across CEOs is quite low in large firms ($\sigma(E[X])$=10.3%, compared to 14.9% in the full sample), which results in a low estimate of variation in CEO ability (σ_0 =1.23%, compared to 2.42% in the full sample). A lower σ_0 requires a lower total turnover cost (only 1.42%) to fit the CEO firing rate, which is just slightly higher in larger firms (2.45% per year, versus 2.29% in the full sample). Second, the model needs a firm turnover cost that is even larger than the total cost in order to fit a large, unexpected drop in profitability around forced turnovers in large firms. Since the firm cost exceeds the total cost, the estimated effective personal turnover cost is slightly negative.

A negative effective personal cost implies that boards derive a personal net benefit from firing CEOs. This is not unreasonable, since firing the CEO may protect directors' reputations or positions on the board. For instance, imagine a firm that has experienced very low earnings. If additional signals tell the board that the low earnings were not due to low CEO ability, then firing the CEO may not be in shareholders' interests. However, if this additional signal is private or unverifiable, or if shareholders irrationally "demand blood" as in Fisman, Khurana, and Rhodes-Kropf (2005), then the board may use the CEO as a scapegoat by firing him or her. It is possible that, in larger firms, directors care more about their reputations, and shareholders (or politicians) are more likely to demand a scapegoat.

As noted above, the estimated variation in CEO ability is lower in larger firms (σ_0=1.23% in large firms vs. 3.26% in small firms). One potential explanation is that CEOs matter less (as a % of assets) in larger firms, where CEOs necessarily delegate more to other senior managers. Another explanation is that CEOs are more closely vetted in larger firms. The difference may also be due to noisier accounting data in smaller firms, although the parameter σ_ϵ (time-series volatility of profitability) already captures much of this difference.

The estimated effective personal cost drops from 8.32% in the 1971 to 1989 subsample

to 2.28% in the 1990 to 2006 subsample. This difference is statistically significant at the 1% level. While lower, the estimated cost is still significantly positive, indicating CEOs were still entrenched from 1990 to 2006. The total turnover cost drops from 8.4% to 4.0% of assets. A lower turnover cost should result in more forced successions, all else equal. Indeed, the model predicts an increase from 10% to 23% in the % of successions that are forced. The empirical change is from 12% to 23%, a close match.

As expected, forced turnover is more common in the subsample with high stock ownership by the board (17.6% compared to 12.6%, Panel B). However, the estimated personal cost is also higher in the subsample with higher stock ownership (7.99% vs. 6.41% of assets), although this difference is not statistically significant ($t = 0.83$). This result is unexpected. The model attributes the higher turnover rate not to lower turnover costs, but to more dispersion in ability across CEOs in the high-ownership subsample: σ_0 increases from 3.10% to 3.99%. This result is partially due to smaller firm size in the high-ownership subsample (median assets are $1.1B vs. $2.2B), although results are quite different when I sort directly on firm size above. Unfortunately, I cannot simultaneously control for size and ownership, as I explain in footnote 13.

The personal turnover cost is significantly lower in the subsample with more outsiders on the board (3.00% compared to 8.25%; difference has t-statistic 2.7). The lower turnover cost explains why more successions are forced in the subsample with more outsiders (16.0% vs. 11.6%). Part of this result is due to a size effect; the subsample with few outsiders contains smaller firms (median assets are $0.8B vs. $1.7 B).

The estimated noise of the board's additional signal is lower in the subsample with more insiders on the board (σ_z=5.85% vs. 10.80%). In other words, insider-dominated boards appear to rely more signals besides profitability when evaluating CEO ability. This difference makes sense if we believe insiders have better access to non-earnings signals such as the CEO's day-to-day decisions and their outcomes. Panel B shows that the difference

in σ_z allows the model explain why profitability is a worse predictor of CEO dismissal in insider-dominated boards (pseudo R^2 of 2% vs. 5%). The estimate of σ_z is also lower in the subsample with higher rather than lower stock ownership by boards (8.62% vs. 12.92%). A potential explanation is that boards owning more stock have a stronger incentive to gather additional signals about CEO ability. The difference in σ_z helps the model explain why profitability is a worse predictor of forced turnover in the high-ownership subsample (pseudo R^2 of 3% vs. 6% in Panel B).

To summarize, three out of four split-sample tests (year, outsiders, size) go in the expected direction: weaker governance is associated with a higher effective personal turnover cost. The fourth test (board stock ownership) finds no significant relation. These results provide a useful consistency check that hopefully builds confidence in the paper's interpretation of the personal turnover cost. As noted above, this analysis has nothing to say about whether weak governance is good or bad. It simply tests a predicted consequence of weak governance, namely, that boards will be less willing to fire bad CEOs.

If weak governance is indeed bad for shareholders, then how bad is it? How much shareholder value does CEO entrenchment (i.e. the effective personal turnover cost) destroy, and how might the world look with no entrenchment? I use the model to explore this counterfactual scenario. I create a control sample by simulating the model using parameter values estimated from the full sample. I then create a treatment sample by setting the effective personal turnover cost to zero and repeating the simulations. Results are in the first row of Table VI. When I lower the personal turnover cost from its estimate of 4.6% to 0%, the % of CEOs fired per year rises from 2% to 12%, and mean CEO tenure drops from 7.5 to 4.5 years. Average profitability rises, because CEOs with low ability are more likely to be fired, which raises surviving CEOs' average ability. However, average profitability rises by only half a percentage point per year. Entrenchment's effect on profitability is fairly small, for three reasons: even with no entrenchment, bad CEOs would still spend at least one year in office before being identified and fired; firms would pay the 1.3% firm cost ($c^{(firm)}$) more

often since succession is more frequent; and the very worst CEOs were already being fired even when entrenchment was severe.

INSERT TABLE VI NEAR HERE

I compute an average market-to-book ratio by discounting back the average future simulated profits. When I set the personal turnover cost to zero, the market/book ratio increases by 3.1%. In other words, eliminating entrenchment would raise shareholder value by 3.1%, assuming the model and baseline parameter estimates are true, and assuming we could eliminate the personal cost without affecting the other model parameters.

I may now be pushing the model beyond its limitations. Holding other parameters constant while varying the personal cost is particularly aggressive. For instance, CEOs may exit the labor market if we eliminate entrenchment, which in turn may change the model parameters μ_0 and σ_0 that govern the CEO talent pool. As I explain earlier in this section, it is not even clear that reducing entrenchment is optimal for shareholders. These concerns also apply to the experiments I describe later in this section. These exercises represent a first step towards quantifying the effect of CEO entrenchment on shareholder value. Hopefully future research will produce more refined estimates.

The next rows in Table VI repeat the same experiment using the parameter estimates from the eight subsamples examined in Table V. The predicted increase in shareholder value from eliminating entrenchment varies considerably across subsamples. The effect is much larger in the subsamples with higher estimated personal costs, including the 1971 to 1989 subsample (a 6.9% increase in shareholder value), the subsample with high stock ownership by boards (an 8.8% increase), the subsample with fewer outsiders on the board (a 6.5% increase), and in small firms (a 7.7% increase). In the 1990 to 2006 subsample, eliminating entrenchment increases shareholder value by just 1.4%, because the estimated personal cost is only 2.28%. Since the estimated personal cost is negative for large firms, setting this cost to zero results in fewer CEO firings (down from 2.6% to 1.8% per year), but the resulting

increase in shareholder value is less than 0.05%.

The next experiment in Table VI examines the effect on shareholder value of hiring a very bad CEO, specifically, a CEO with ability two standard deviations below the mean. First I run this experiment with all parameters at their estimated values from the full sample. Compared to a randomly drawn CEO, the bad CEO faces a higher annual firing rate (17.3% vs 2.1%), survives less time in office (4.0 vs. 7.5 years), and achieves lower average annual profitability (13.9% vs 15.5%). The net present value of hiring this very bad CEO is 6.3% lower than the NPV from randomly drawing a CEO from the prior distribution. The next row in the table repeats this experiment assuming no entrenchment (i.e. personal costs equal zero). Comparing the experiment with and without entrenchment, I find that eliminating entrenchment mitigates the negative effects of a bad CEO: bad CEOs get fired sooner (1.6 years versus 4.0 years) and therefore destroy less shareholder value (–4.6% versus –6.3%).

IV.   Robustness

All the results depend on the model's specific assumptions. This section describes how results change when I use different assumptions. I estimate the model using a more aggressive classification of CEO successions, firm fixed effects in profitability, a different industry classification, a flat firing threshold, a different assumption about voluntary turnover costs, and alternate discount factors. I also discuss earnings manipulation. Results are in Table VII.

A.   Alternate Forced/Voluntary Classification

Parrino's (1997) algorithm, which I use to classify successions in the data as forced or voluntary, arguably produces more false negatives than false positives. For example, if a CEO abruptly "retires to spend more time with the family" at age 60, the algorithm classifies the

succession as voluntary, whereas it may have really been forced[16]. As a result of these false negatives, the true forced turnover rate is likely higher than the estimated 2% per year. The model would need a lower turnover cost to fit this higher firing rate. In other words, we should interpret the original 4.6% estimated personal cost as an upper bound.

To establish a lower bound, I try an alternate classification that goes aggressively in the opposite direction, arguably producing more false positives than false negatives. I assume any CEO who left office at age 64 or younger was forced out. The forced turnover rate rises from 2% to 9% per year, and the estimated effective personal cost drops from 4.6% to 1.3% (Table VII, row "Alt. fired def'n"). Since many of these pre-65 successions were probably voluntary, I interpret 9% as an upper bound on the firing rate and 1.3% as a lower bound on the personal turnover cost. Although far below the 4.6% upper bound, the 1.3% lower bound is still quite large in dollar terms: $57 ($292) million for the median (mean) sample firm. For this reason, I still conclude that the model needs a large personal cost to fit the data, implying a high degree of entrenchment. My prior is that the true personal cost is closer to the 4.6% upper bound than the 1.3% lower bound, for two reasons. First, this alternate classification is crude compared to Parrino's classification, which uses information from news articles, not just CEO age. Second, the model has a much harder time fitting data from the alternate classification (the χ^2 statistic increases from 33 to 399), suggesting the classification may be flawed.

B.   Heterogeneity in Average Profitability

A key identification assumption is that all within-industry, long-run variation in average profitability is due to variation in CEO skill. If some industry subsectors were more profitable than others, then I would over-estimate the variation in CEO skill, σ_0. Since higher values of σ_0 may require higher turnover costs to fit firing rates, my estimated turnover costs may also be too large. I perform two additional tests to address this concern.

First, I use a finer industry classification, which makes it less likely that some industry subsectors are more profitable than others. I re-estimate the model using the 30- rather than 12-industry classification on Kenneth French's website. Results are in Table VII, row "30 industries." As expected, the estimated dispersion in ability drops slightly (from 2.42% to 2.30%), and so does the total estimated turnover cost (from 5.94% to 4.91%). However, the total turnover cost is still quite large and mainly reflects a personal cost to the board (down from 4.61% to 4.05%), so the main conclusions do not change.

Second, I introduce firm fixed effects in profitability, which allows some firms to be more profitable than others for reasons unrelated to CEO ability. If the fixed effects are independent of CEO ability, then we can measure each firm's fixed effect by averaging profitability across multiple CEOs in the firm. After subtracting the firm's average profitability from each CEO's average profitability, any remaining variation in average profitability across CEOs is due only to variation in CEO ability. Following this logic, I demean excess profitability y_{it}^* at the firm level and re-estimate the model.

Results are in Table VII, row "Fixed effects." As expected, the estimated dispersion in ability across CEOs (σ_0) is lower with fixed effects (down from 2.42% to 1.61%). Contrary to the logic above, fixed effects result in a higher estimated total turnover cost (up from 5.9% to 6.1%), although the increase is not statistically significant. In short, including fixed effects actually strengthens the conclusion that the model needs large turnover costs to fit observed firing rates. To understand why the estimated total turnover cost increases, note that predicted firing rates depend not just on turnover costs and dispersion in ability, but also on the speed of learning. Since less dispersion (lower σ_0) and higher turnover costs both push the predicted firing rate down, the estimated speed of learning must have increased in order to pull the predicted firing rate back up to its empirical level. This is indeed what happened. Firm fixed effects increase the persistence parameter ϕ from 0.12 to 0.26, indicating profitability is less persistent. As explained in Section I.B, the board learns faster when profitability is less persistent. The lower estimates of σ_ϵ (profitability volatility) and σ_z

(volatility of the additional signal) also imply that the board learns faster, which increases the predicted firing rate.

C.   Tenure and the Firing Threshold

The model assumes the board is risk neutral and both turnover costs and a CEO's ability are constant over time. As a result, the endogenous firing threshold increases with tenure, meaning the board is more willing to fire CEOs the longer they have been in office. Intuitively, there is less uncertainty about long-tenured CEOs, and the board prefers higher uncertainty since it can fire CEOs and thereby minimize downside risk. The board holds an option to fire the CEO, and this option is more valuable when there is more uncertainty.

Four model extensions could change this prediction. First, CEO entrenchment may increase with tenure as the CEO appoints more allies to the board or gains bargaining power, as in Hermalin and Weisbach (1998). Translating into my model, the effective personal cost $c^{(pers)}/\kappa$ may increase with tenure, making the board less willing to fire long-tenured CEOs. Second, if the board is risk averse, then it will be less willing to fire long-tenured CEOs, whose ability is less uncertain. Third, if CEOs gain human capital from learning on the job, and if shareholders receive at least part of the surplus, then boards should be less willing to fire CEOs the longer they have spent in office. Fourth, CEOs' ability may fluctuate randomly over time as their human capital gains or loses productivity. Dangl, Wu, and Zechner (2007) show that random fluctuations in ability cause uncertainty to drop less with tenure, because old signals lose relevance. As they show, the gains in option value from replacing an old manager with a new one are smaller, which reduces the slope of the firing threshold.

All four of these extensions would make the firing threshold increase less with tenure than my model predicts. To see how a flatter threshold affects my results, I estimate the model forcing the firing threshold to be perfectly flat and equal to its value in CEOs' first year in office. A flat firing threshold is admittedly *ad hoc*. The four factors above may result in a

threshold with a slightly positive slope, or even a negative slope. I impose a flat threshold and use the same learning dynamics as before to keep this robustness exercise simple.

The χ^2 statistic indicates the model with a flat threshold fits the data roughly as well as the main model (Table VII). The central conclusions do not change: the total turnover cost is huge (5.99%, up from 5.94%) and mainly reflects a personal cost (5.28%, up from 4.61%). These conclusions are even stronger in the story where entrenchment increases with tenure, since flattening the firing threshold requires the total turnover cost to start at 5.99% and then increase with tenure, as I show in the Internet Appendix.

D.   Costless Voluntary Turnover

The model assumes forced and voluntary CEO turnover are equally costly. For robustness, I re-solve and re-estimate the model assuming voluntary turnover is costless to both shareholders and the board. In other words, I set $c^{(firm)} = c^{(pers)} = 0$ for voluntary turnover, and then I estimate $c^{(firm)}$ and $c^{(pers)}$ for forced turnover, along with the other model parameters. When voluntary turnover is costless, boards have an incentive to wait until bad CEOs retire at no cost. This feature changes the model predictions in two main ways: it reduces the overall firing rate, and it makes firings at late tenures relatively less common. Parameter values must therefore change to raise the predicted firing rate back to its empirical level, and to shift more firings to later tenures. Figure 1 shows that this is a challenge: reducing turnover costs or increasing dispersion in ability (σ_0) would increase the predicted firing rate (as desired) but would shift firings in the wrong direction, i.e. toward earlier tenures. The model finds a clever, albeit imperfect[17], solution (Table VII, row "$c_{(retire)} = 0$"). The estimated total turnover costs are almost unchanged (up from 5.94% to 5.95%), dispersion in ability (σ_0) increases from 2.42% to 3.34% (which raises the overall firing rate back toward its empirical level), and σ_z (the noise in the additional signal) increases from 5.15% to 9.71% (which makes the board learn slower and shifts firings to later tenures). The key is that an

increase in σ_0 *requires* an increase in σ_z in order to continue fitting the V-shape pattern in average profitability around firings. In fact, the predicted V shape is now slightly deeper than before, so there is less of an unexpected drop in profitability in the firing year, which means the model does not need such a large firm turnover cost to fit the data ($c^{(firm)}$ drops from 1.33% to 0.01%). Since the total turnover cost is unchanged and the firm cost drops, the estimated personal cost increases from 4.61% to 5.94%. In sum, the paper's two main conclusions are robust: the model still needs large turnover costs to fit the data, and these costs mainly reflect entrenchment rather than a real cost to shareholders.

E. Alternate Discount Factors

Next, I estimate the model using different assumed values of β, the board and investors' shared discount factor. The main results use $\beta = 0.9$, and in these robustness tests I use β=0.85 and 0.95. The estimated total turnover cost is increasing in the discount factor (Table VII). This result is expected, because increasing β increases the present value of future benefits from firing the CEO, which makes the board more willing to fire the CEO. To offset this effect, the model makes the board less willing to fire the CEO by increasing the total turnover cost. Even when $\beta = 0.85$, the total turnover cost is still quite large (3.89%) and mainly reflects a personal cost (2.68%).

F. Earnings Manipulation

The model assumes reported earnings equal true earnings. In reality, CEOs have incentives to manipulate earnings in at least three ways.

First, if a CEO believes he is close to being fired, then he may try to inflate reported earnings. Murphy and Zimmerman (1993) find no empirical evidence of such manipulation.

Weisbach (1995) and Murphy and Zimmerman (1993) document that CEOs take an

earnings bath when they first enter office, possibly to boost future compensation and chances of staying in office. Since my model will mistakenly attribute these earnings baths to the firm turnover cost, my estimated firm turnover cost $c^{(firm)}$ (1.3% in the full sample) may be too high. The estimated effective personal turnover cost $c^{(pers)}/\kappa$ may therefore be too low, since the total turnover cost is unlikely to change. Earnings baths may also explain why the model fails to explain the drop in profitability in the year after forced successions (Figure 4).

Finally, CEOs may engage in "signal jamming," injecting noise into earnings to make it harder for the board to learn the CEO's ability (e.g. Fudenberg and Tirole (1986), Hermalin and Weisbach (2008)). While signal jamming may help explain some of my estimates— for instance, why the estimated volatility of profits σ_ϵ is 3.4% instead of some lower number— signal jamming does not necessarily imply any bias in these estimates. Signal jamming may also help explain my finding that boards rely heavily on non-earnings signals when evaluating the CEO.

V. Conclusion

This study establishes a benchmark for the forced CEO turnover rate and quantifies the amount of shareholder value at stake. I solve and estimate a dynamic model of the CEO firing decision, and I find three main results. First, the empirical forced turnover rate is low, in the sense that the model needs large turnover costs to fit the data. Second, these costs mainly reflect CEO entrenchment rather than a real cost to shareholders. According to the model, eliminating this entrenchment would raise shareholder value by 3%, assuming we could hold all else constant.

CEO entrenchment is bad for shareholders ex post, because it means that boards retain some CEOs whom shareholders would rather see fired. The model says nothing about

whether this entrenchment is bad for shareholders ex ante. For instance, shareholders may rationally allow some degree of entrenchment, e.g. by appointing a CEO-friendly board, in order to attract a talented CEO. To understand what level of entrenchment is optimal ex ante, we need to extend the model to include the initial choice of governance structures. Developing and estimating such a model is an interesting area for future research.

REFERENCES

Allgood, Sam, and Kathleen A. Farrell, 2003, The match between CEO and firm, *Journal of Business* 76(2), 317–341.

Almazan, Andres, and Javier Suarez, 2003, Entrenchment and severance pay in optimal governance structures, *Journal of Finance* 58, 519–547.

Bertrand, Marianne, and Antoinette Schoar, 2003, Managing with style: The effects of managers on firm policies, *Quarterly Journal of Economics* 118(4), 1169–1208.

Brickley, James A., 2003, Empirical research on CEO turnover and firm-performance: A discussion, *Journal of Accounting and Economics* 36, 227–233.

Bushman, Robert M., Raffi J. Indjejikian, and Abbie Smith, 1996, CEO compensation: The role of individual performance evaluation, *Journal of Accounting and Economics* 21, 161–193.

Chevalier, Judith and Glenn Eillison, 1999, Career concerns of mutual fund managers, *Quarterly Journal of Economics*, 114(2), 389–432.

Coughlan, Anne T., and Ronald M. Schmidt, 1985, Executive compensation, management turnover, and firm performance, *Journal of Accounting and Economics* 7, 43–66.

Dangl, Thomas, Youchang Wu, and Josef Zechner, 2006, Market discipline and internal governance in the mutual fund industry, Forthcoming, *Review of Financial Studies.*

DeAngelo, Harry and Linda DeAngelo, 1989, Proxy contests and the governance of publicly held corporations, *Journal of Financial Economics* 23(1), 29–60.

Demb, Ada and F. Friedrich Neubauer, 1992, *The Corporate Board* (Oxford University Press, Oxford).

Eisfeldt, Andrea L. and Adriano A. Rampini, 2008, Managerial incentives, capital realloca-

tion, and the business cycle, *Journal of Financial Economics* 87, 177–199.

Engel, Ellen, Rachel M. Hayes, and Xue Wang, 2003, CEO turnover and properties of accounting information, *Journal of Accounting and Economics* 36, 197–226.

Fama, Eugene F., and Kenneth R. French, 2000, Forecasting profitability and earnings, *Journal of Business* 73, 161–175.

Fama, Eugene F., and Michael C. Jensen, 1983, Separation of ownership and control, *Journal of Law and Economics* 26, 301–325.

Fee, Edward C. and Charles J. Hadlock, 2004, Management turnover across the corporate hierarchy, *Journal of Accounting and Economics*, 37, 3–38.

Fisman, Ray, Rakesh Khurana, and Matthew Rhodes-Kropf, 2005, Governance and CEO turnover: Do something or do the right thing?, Working paper, Columbia University.

Fudenberg, Drew and Jean Tirole, 1986, A 'Signal-Jamming' Theory of Predation, *Rand Journal of Economics* 17, 366–376.

Gabaix, Xavier, and Augustin Landier, 2008, Why Has CEO Pay Increased So Much?, *Quarterly Journal of Economics 123*, 49–100.

Graham, John R., Si Li, and Jiaping Qiu, 2009, Managerial attributes and executive compensation, Working paper, Duke University.

Hennessy, Christopher A., and Toni M. Whited, 2005, Debt dynamics, *Journal of Finance* 60(3), 1129–1165.

Hennessy, Christopher A., and Toni M. Whited, 2007, How costly is external financing? Evidence from a structural estimation, *Journal of Finance* 62, 1705–1745.

Hermalin, Benjamin E., and Michael S. Weisbach, 2008, Information disclosure and corporate governance, Working paper, University of California Berkeley.

Hermalin, Benjamin E., and Michael S. Weisbach, 1998, Endogenously chosen boards of directors and their monitoring of the CEO, *American Economic Review* 88(1), 96–118.

Hermalin, Benjamin E., and Michael S. Weisbach, 1991, The effects of board composition and direct incentives on firm performance, *Financial Management* 20(4), 101–112.

Herschleifer, David, and Anjan V. Thakor, 1994, Managerial performance, boards of directors, and takeover bidding, *Journal of Corporate Finance* 1, 63–90.

Herschleifer, David, and Anjan V. Thakor, 1998, Corporate control through board dismissals and takeovers, *Journal of Economics and Management Strategy* 7(4), 489–520.

Huson, Mark R., Paul H. Malatesta, and Robert Parrino, 2004, Managerial succession and firm performance, *Journal of Financial Economics,* 237–275.

Huson, Mark R., Robert Parrino, and Laura T. Starks, 2001, Internal monitoring mechanisms and CEO turnover: A long-term perspective, *Journal of Finance* 56, 2265–2297.

Jenter, Dirk, and Fadi Kanaan, 2008, CEO turnover and relative performance evaluation, Working Paper, Stanford University and MIT.

Kaplan, Steven N. and Bernadette A. Minton, 2006, How has CEO turnover changed? Increasingly performance sensitive boards and increasingly uneasy CEOs, Working paper, University of Chicago.

Kim, Yungsan, 1996, Long-term firm performance and chief executive turnover: An empirical study of the dynamics, *Journal of Law, Economics, and Organization* 12(2), 480–496.

Lehn, Kenneth M., and Mengxin Zhao, 2006, CEO turnover after acquisitions: Are bad bidders fired?, *Journal of Finance* 61(4), 1759–1811.

Lorsch, Jay W., and Elizabeth MacIver, 1989. *Pawns or potentates: The reality of America's corporate boards* (Harvard Business School Press, Boston).

Manso, Gustavo, 2007, Motivating innovation, Working paper, MIT Sloan School of Management.

McFadden, Daniel, 1989, A method of simulated moments for estimation of discrete response models without numerical integration, *Econometrica* 57(5), 995–1026.

Michaelides, Alexander, and Serena Ng, 2000, Estimating the rational expectations model of speculative storage: A Monte Carlo comparison of three simulation estimators, *Journal of Econometrics* 96, 231–266.

Miller, Robert A., 1984, Job matching and occupational choice, *Journal of Political Economy* 92(6), 1086–1120.

Murphy, Kevin J., and Jerold L. Zimmerman, 1993, Financial performance surrounding CEO turnover, *Journal of Accounting and Economics* 16, 273–315.

Pakes, Ariel, and David Pollard, 1989, Simulation and the asymptotics of optimization estimators, *Econometrica* 57(5), 1027–1057.

Parrino, Robert, 1997, CEO turnover and outside succession: A cross-sectional analysis, *Journal of Financial Economics* 46, 165–197.

Parrino, Robert, Richard W. Sias, and Laura T. Starks, 2003, Voting with their feet: Institutional ownership changes around forced CEO turnover, *Journal of Financial Economics* 68, 3–46.

Pástor, Ľuboš, Lucian A. Taylor, and Pietro Veronesi, 2007, Entrepreneurial learning, the IPO decision, and the post-IPO drop in firm profitability, *Review of Financial Studies* 22, 3005–3046.

Rust, John, 1994, Structural estimation of Markov decision processes, in R.F. Engle and D.L. McFadden, ed.: *Handbook of Econometrics, Volume IV* (Elsevier Science B.V.).

Shivdasani, Anil, and David Yermack, 1999, CEO involvement in the selection of new board

members: An empirical analysis, *Journal of Finance* 54, 1829–1854.

Taylor, Lucian A., 2009, CEO pay and CEO power: Evidence from a dynamic learning model, Working paper, University of Pennsylvania.

Warner, Jerold B., Ross L. Watts, and Karen H. Wruck, 1988, Stock prices and top management changes, *Journal of Financial Economics* 20, 461–492.

Weisbach, Michael S., 1995, CEO turnover and the firm's investment decisions, *Journal of Financial Economics* 37, 159–188.

Weisbach, Michael S., 1988, Outside directors and CEO turnover, *Journal of Financial Economics* 20, 431–460.

Wu, Joanna S., and Ivy X. Zhang, 2008, The adoption of international accounting standards and firm internal performance evaluation, Working paper, University of Rochester.

Yermack, David, 2006, Golden handshakes: Separation pay for retired and dismissed CEOs, *Journal of Accounting and Economics* 41, 237–256.

Footnotes

1. E.g., Coughlan and Schmidt (1985); Warner, Watts, and Wruck (1988); Weisbach (1988); Murphy and Zimmerman (1993); Weisbach (1995); Kim (1996); Parrino (1997); Huson, Parrino, and Starks (2001); Parrino, Sias, and Starks (2003); Engel, Hayes, and Wang (2003); Huson, Malatesta, and Parrino (2004); Kaplan and Minton (2006); Jenter and Kanaan (2006); Lehn and Zhao (2006).

2. E.g., Herschleifer and Thakor (1994, 1998); Hermalin and Weisbach (1998, 2007); Fisman, Khurana, and Rhodes-Kropf (2007); Eisfeldt and Rampini (2008).

3. Making voluntary turnover exogenous and only dependent on tenure improves tractability, is consistent with the lack of an empirical correlation between firm performance and voluntary turnover (Huson, Malatesta, and Parrino (2004)), and still allows the timing of voluntary turnover to affect the firing decision.

4. For tractability, I assume all profits (including potential negative profits) are immediately paid out as dividends. It follows that assets B_t are constant over time.

5. Profitability is net of CEO pay, which is outside the model.

6. Yermack (2006) finds that separation pay to CEOs is increasing in firm size. Executive search fees are proportional to CEO compensation, which increases in firm size, (Gabaix and Landier (2008)).

7. Later I discuss whether the personal turnover cost is optimal for shareholders ex ante.

8. Annual ROA equals operating income before depreciation (Compustat item OIBDP) divided by the midpoint of assets (item AT) from the current and previous fiscal years. I eliminate ROA observations outside [-100%,100%].

9. The 2,500 count is based on the 1,316 successions in Parrino's 1971 to 1994 sample, plus the 883 turnovers in the 1995 to 2006 Forbes surveys, plus 438 CEOs who had not left office by the end of 2006. Out of the 883 Forbes successions from 1995 to 2006, only 514 (58%) could be classified as forced or voluntary, and only 312 (35%) also had complete Compustat data.

10. For example, McFadden (1989), Pakes and Pollard (1989), Rust (1994), Hennessy and Whited (2005, 2007).

11. The industry standard CEO search fee is one-third of the CEO's total cash compensation in his first year in office, so the average search fee in my sample is roughly $1 million.

12. The CEO approves and often proposes the slate of directors (e.g. Lorsch and MacIver (1989), Demb and Neubauer (1992)). DeAngelo and DeAngelo (1989) show that shareholders almost always approve the slate proposed by management. Shivdasani and Yermack (1999) provide additional empirical evidence.

13. I perform simple split sample tests for two reasons. Including several control variables at once, as in a multiple regression, would be too expensive computationally. Splitting the sample along several dimensions (e.g. by industry and size) is problematic because the full sample contains only 168 forced successions, and the resulting subsamples would be extremely small.

14. I thank Robert Parrino for providing data on board share ownership data (originally from proxy statements) and board composition (originally from the Million Dollar Directory). Both measures are available only up to 1994. I exclude CEO spells with missing governance measures.

15. I conduct inference by assuming estimators from the two subsamples are uncorrelated with each other. This assumption is plausible under the model's assumption that draws from the CEO talent pool, profitability shocks, and realizations of signal z are all independently distributed across both firms and time.

16. If the CEO were 59 then the algorithm would classify this succession as forced.

17. The model fits worse in this robustness exercise; χ^2 increases from 33.2 to 36.0.

Table Captions

Table I: Panel A contains summary statistics on CEO spells in various subsamples. The full sample consists of CEO spells for firms in the 1970-2007 *Forbes* annual compensation surveys. I include complete CEO spells that ended between 1971-2006. Additional details are in Section II.A. The 12 industries are defined on Kenneth French's website. "1970-1974" is the subset of full CEO spells ending in 1970-1974, and so on. Firm/years is the number of firm/year observations in the given subsample. Total spells is the number of CEO spells. "% forced" is the % of CEO spells that ended in forced succession. "% forced per year" is the % of sample firm/years that ended in a CEO dismissal. Panel B contains additional statistics for the full sample. ROA is the % firm-level annual return on assets. $y_t^* = ROA - v_t$ equals firm ROA minus average industry ROA. "Assets" is Compustat item "AT" in historical \$B. Spell length is the number of years the CEO completed in office before leaving his position. Statistics for ROA, v_t, and assets are computed pooling all firm/years, and statistics for spell length are computed across all CEO spells.

Table II: This table contains estimates of the parameters from the model in Section I. Estimation uses data on a sample of 981 CEOs, described in Section II.A. Parameters are estimated using the simulated method of moments, as described in Section III. $c^{(firm)}$ is the CEO turnover cost to the firm, and $c^{(pers)}$ is the personal CEO turnover cost to the board. κ controls the degree to which the board internalizes shareholder value. μ_0 and σ_0 are the mean and standard deviation, respectively, of boards' prior beliefs about a newly hired CEO's ability. ϕ controls the persistence in profitability. σ_ϵ is the conditional time-series volatility of profitability. σ_z is the standard deviation of the board's additional signal about CEO ability. All parameters are in units of % of assets per year, except ϕ (unitless), and the cost c (% of assets). Standard errors are in parentheses.

Table III: Panel A shows the 14 moments used in the SMM estimation described in Section II. Empirical moments are computed from the sample of 981 CEOs described in

Section II.A. Simulated moments are computed from data simulated from the model using parameter values in Table II. Moments' standard errors and p-values are computed by Monte Carlo, as follows. I create 10,000 sets of 14 moments, each from a simulation of 981 CEOs (to match the empirical sample size) using parameter values in Table II. The standard error is the standard deviation of the 10,000 simulated moments, and the p-value is fraction of simulated moments which are at least as far from the mean simulated moment as the empirical moment is. The first 8 moments come from the regression $y_{it}^* = \lambda_0 + \lambda_1 y_{i,t-1}^* + \Delta^{(-2)} + \Delta^{(-1)} + \Delta^{(0)} + \Delta^{(1)} + \Delta^{(2)} + \delta_{it}$, estimated pooling all years and firms. $y_{i,t}^*$ is firm i's annual % ROA in year t minus industry average profitability. $\Delta^{(k)}$ is a fixed effect for whether forced CEO turnover occurred at the end of period $t + k$. $E[Var(X)]$ is the mean across CEOs of the within-CEO variance of X_{it}, which is persistence-adjusted firm profitability. $Var(E[X])$ is the variance across CEOs of within-CEO average X_{it}. The hazard rates $h^{(j)}$ equal the % of CEOs forced out of office per year during tenure period (j), conditional on the CEO reaching (j). Panel B shows the χ^2 statistic and corresponding p-value for SMM's test of over-identifying restrictions, which jointly tests whether the empirical and simulated moments are equal. This test is defined in the Internet Appendix.

Table IV: This table compares statistics computed from the empirical sample (containing 981 CEOs, described in Section II.A) and a sample of 100,000 CEOs simulated from the model using parameter estimates in Table II. "Median tenure" is the number of years the CEO completed in office before leaving, computed separately for CEOs whose successions were forced and unforced. "% forced" is the % of CEOs who were forced out of office. "% forced per year" is the % of firm/years which end in forced CEO turnover. In the probit model, the dependent variable is an indicator for whether year t ended in a CEO dismissal, and the independent variable is y_t^*, the firm's annual ROA minus the industry's ROA in year t. "Slope" is the estimated slope on y_t^*, and "Stderr." is the associated robust standard error.

Table V: Each pair of rows contains estimation results using a subset of the 981 CEOs

from the full empirical sample. "Small (large) firms" have assets below (above) $6.6B, in CPI-adjusted 2007 dollars. Subsample 1971-1989 and 1990-2006 contain all CEO spells that ended between the given years. "Low (high) ownership" firms are those where the median % of shares owned by officers and directors, not including the CEO, is below (above) 1.31%. "Less (more) outsiders" are firms where the median % of directors who are not also officers in the firm is below (above) 72.7%. Panel A contains parameter estimates with standard errors in parentheses, and also the number of CEO spells used and the χ^2 and p-value from the test of over-identifying restrictions. Variable definitions are in Tables II and IV. Panel B compares statistics computed from the empirical subsample and a sample simulated from the model using parameters from Panel A. $\sigma(E[X])$ is the standard deviation across CEOs of each CEO's average persistence-adjusted profitability X_{it} during his time in office, in % per year.

Table VI: Each row presents results from a different counterfactual experiment in which I change one or more model input. All results are from simulations of the model described in Section I. "Treat." denotes the treatment sample. Experiment 1: The control uses the denoted subsample's parameter estimates from Tables II or V, and treatment uses those same parameters but sets the effective personal turnover cost $c^{(pers)}/\kappa$ to zero. Subsamples are defined in Table V. Experiment 2: Both the control and treatment use full-sample parameter estimates from Table 2 with personal cost as noted below. In the control, all CEOs' abilities are drawn from the prior distribution. In the treatment for experiment 2, the firm hires a CEO with ability 2 standard deviations below the mean, so $\alpha = \mu_0 - 2\sigma_0$, and subsequent CEOs' abilities are drawn from the prior distribution. "% effect on shareholder value" is the % difference in NPV between treatment and control, averaging across all years (for experiment 1), or measured when the first CEO is hired (experiment 2). Other column values are averages across all simulation years (experiment 1), or across the first CEOs' years (experiment 2).

Table VII: Each pair of rows contains estimation results (parameter estimates, standard

errors in parentheses, number of CEO spells used, and statistics for the test of over-identifying restrictions) from a different specification of the model. "Main results" are the same as in Tables II and III. 'Alt. forced def'n" assumes CEO successions are forced if and only if the CEO is aged 64 or less at the time of succession. "Fixed effects" uses profitability data de-meaned at the firm level. "30 industries" computes industry-adjusted profitability using the 30 industries defined on Kenneth French's website. "Flat threshold" forces the firing threshold to stay constant with tenure. "$c_{(retire)} = 0$" sets the turnover cost of voluntary succession to zero. The next rows use different values for β, the board's discount factor.

Figure 1: **Predicted CEO Dismissal Hazard Rates**
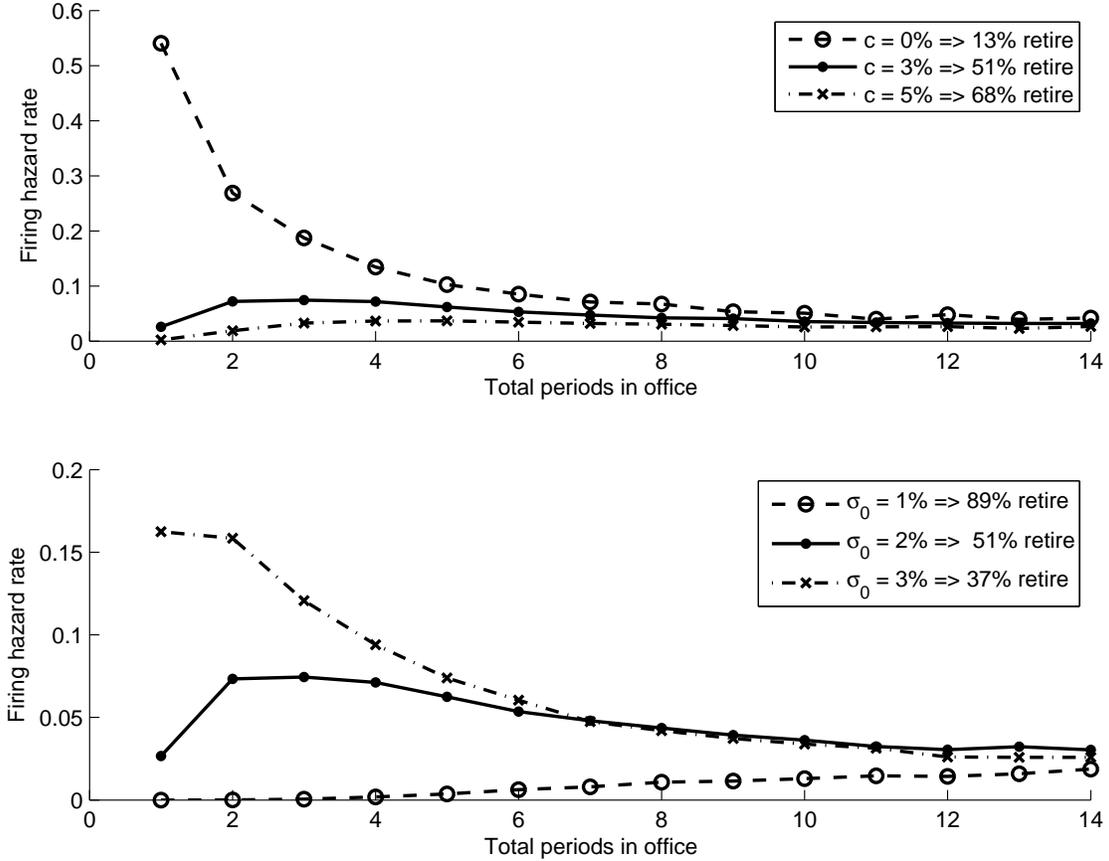
This figure shows the hazard rates for CEO dismissals at different tenure levels τ. The hazard rate for tenure τ is the probability that the CEO will be fired after his τth period in office, conditional on him surviving to period τ. The legends indicate the % of CEOs who reach retirement after 15 periods without being fired. The top panel shows hazard rates for three different values of the total turnover cost, $c = c^{(firm)} + c^{(pers)}/\kappa$. The bottom panel shows hazard rates for three values of σ_0, the standard deviation of CEO ability α in the pool of replacement CEOs. These results are from simulations of the model using parameter values $\beta = 0.9$, $\mu_0 = 1\%$, $\sigma_0 = 2\%$, $\sigma_\epsilon = 3\%$, $c = 3\%$, $\phi = 0.12$, and $\sigma_z = 7\%$; voluntary turnover occurs after (and only after) completing 15 periods in office, so $f(\tau) = 0$ for $\tau = 0, 1, ..., 13$, $f(14) = 1$.

Figure 2: **Profitability around CEO Dismissals**
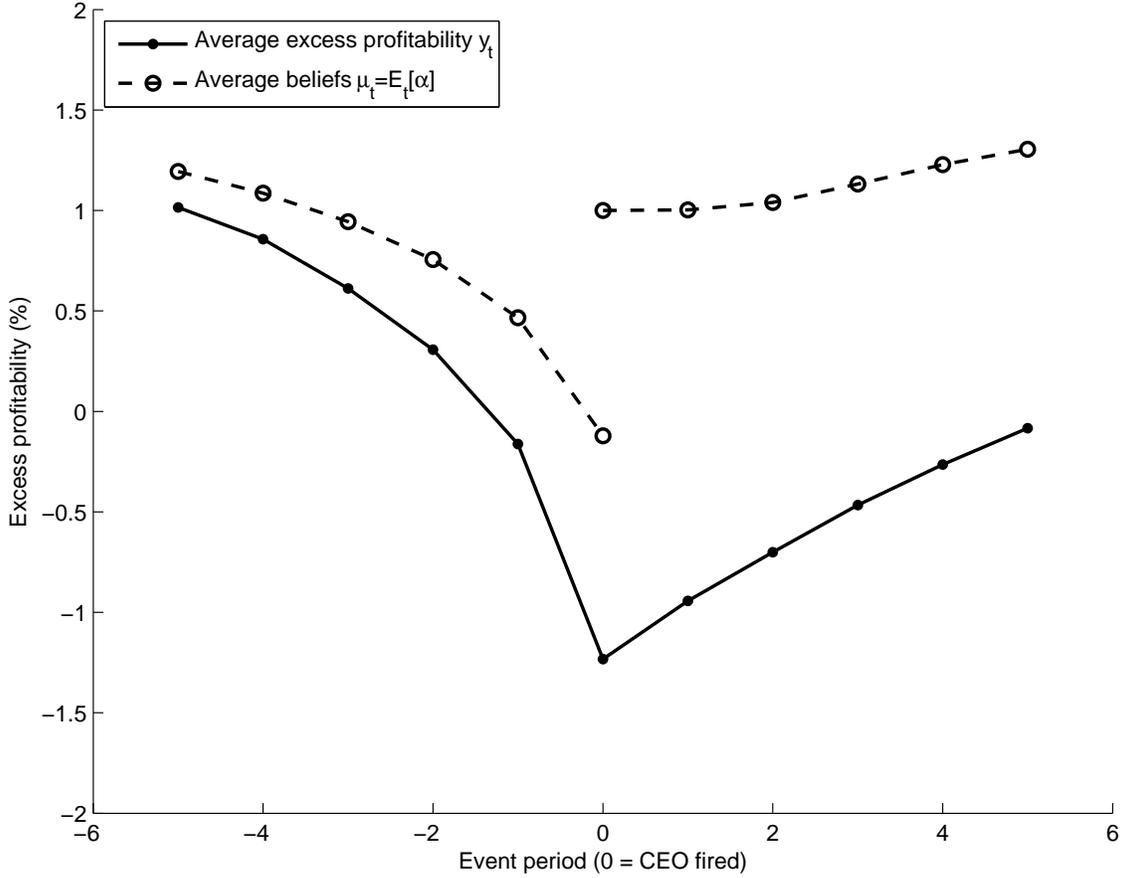
This figure plots average excess profitability in event time around CEO dismissals. Excess profitability y_t equals firm profitability Y_t minus industry profitability v_t. μ_t is the average across CEOs of the board's posterior mean beliefs about the CEO's ability α. Results are from 100,000 consecutive CEO spells in a simulated firm. Simulations use the following parameter values: $\beta = 0.9$, $\mu_0 = 1\%$, $\sigma_0 = 2\%$, $\sigma_\epsilon = 3\%$, $c = 3\%$, $\phi = 0.12$, and $\sigma_z = 7\%$; voluntary turnover occurs after (and only after) completing 15 periods in office, so $f(\tau) = 0$ for $\tau = 0, 1, ..., 13$, $f(14) = 1$. Also, there are is no firm turnover cost, so $c^{(firm)} = 0$ and hence $c = c^{(pers)}/\kappa$.

Figure 3: **Profitability Around CEO Dismissals: Comparative Statics**
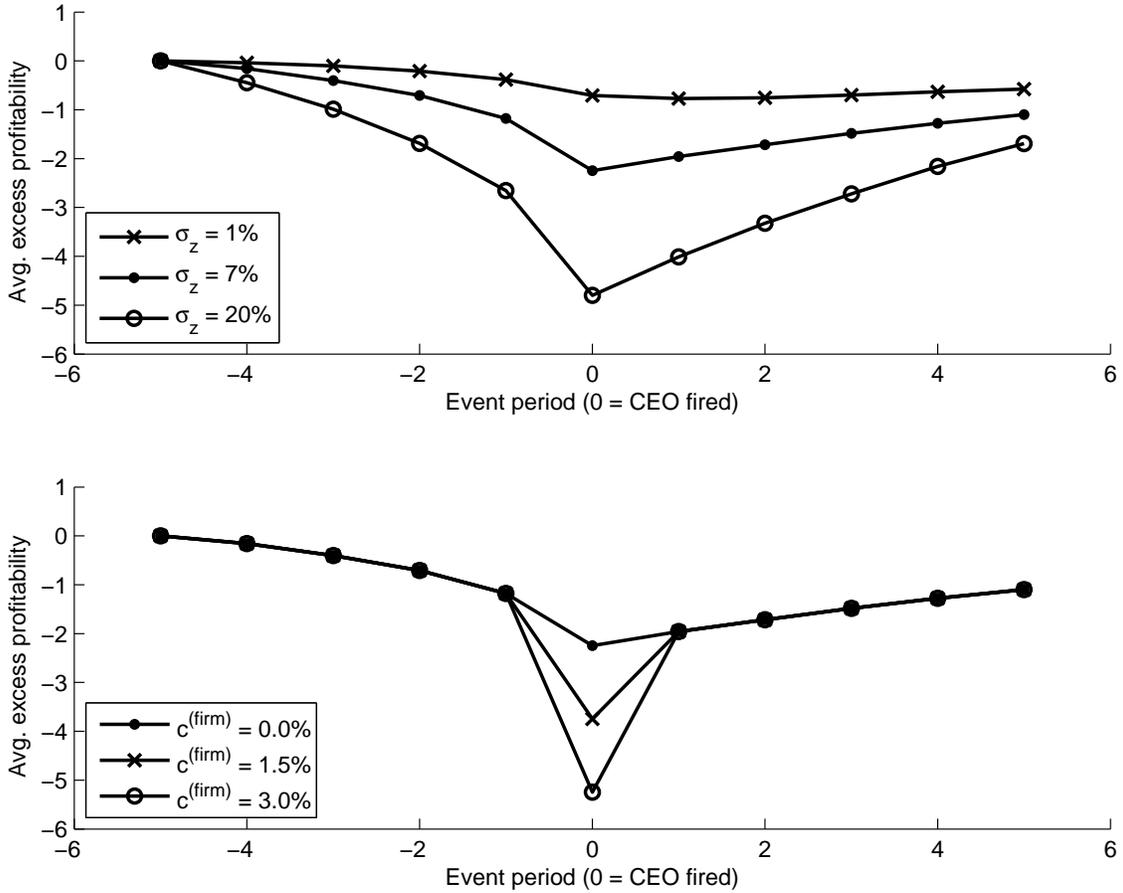
This figure shows average excess profitability in event time around CEO dismissals. Excess profitability y_t equals firm profitability Y_t minus industry profitability v_t. Results are from 100,000 consecutive CEO spells in a simulated firm. Unless otherwise noted in the legend, simulations use parameter values $\beta = 0.9$, $\mu_0 = 1\%$, $\sigma_0 = 2\%$, $\sigma_\epsilon = 3\%$, $c = 3\%$, $\phi = 0.12$, and $\sigma_z = 7\%$; voluntary turnover occurs after (and only after) completing 15 periods in office, so $f(\tau) = 0$ for $\tau = 0, 1, ..., 13$, $f(14) = 1$. In the top panel there is no firm turnover cost, so $c^{(firm)} = 0$, and $c = c^{(pers)}/\kappa$. In the bottom panel, I hold constant total costs $c = c^{(firm)} + c^{(pers)}/\kappa$ constant, and I vary the firm cost $c^{(firm)}$. To allow easier comparisons, I shift all lines vertically so that they match perfectly at period -5, where I set excess profitability equal to 0%.

Figure 4: **Empirical and Predicted Patterns in CEO Dismissals and Profitability**
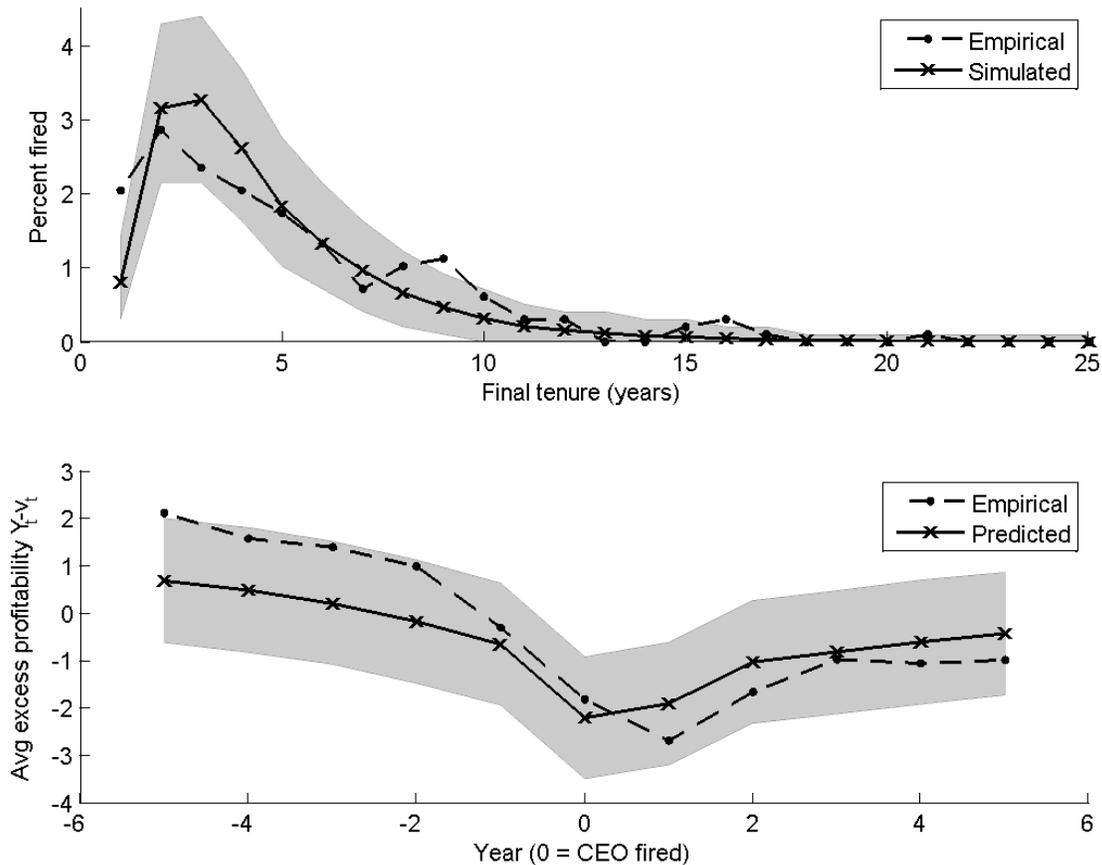
The top panel shows the unconditional % of CEOs fired at different tenure levels. The bottom panel shows average excess profitability in event time around dismissals. In both panels, the empirical pattern is computed from the sample of 981 CEO spells described in Section II.A. The predicted pattern is the average across 10,000 simulated samples, each of which contains 981 artificial CEOs. Simulations use parameter values from Table II. The grey 95% confidence region covers the area between the 2.5% and 97.5% percentiles from the 10,000 simulated samples. Average excess profitability $Y_t - v_t$ equals firm profitability Y_t minus industry profitability v_t. In the empirical sample, profitability is annual ROA.

Table I: **Summary Statistics**

Table captions on separate page.

			Panel A: CEO Spells			
Subsample	Firm/years	Total spells	Forced successions	Unforced successions	% forced	% forced per year
Full sample	7,325	981	168	813	17.1	2.29
Consumer nondurables	757	87	13	74	14.9	1.72
Consumer durables	325	42	8	34	19.0	2.46
Manufacturing	1,543	204	28	176	13.7	1.81
Energy	315	40	7	33	17.5	2.22
Chemicals	436	63	3	60	4.8	0.69
Business equipment	605	90	23	67	25.6	3.80
Telecom	153	23	5	18	21.7	3.27
Utilities	662	82	7	75	8.5	1.06
Wholesale and retail	517	88	19	69	21.6	3.68
Health	481	57	8	49	14.0	1.66
Finance	920	127	27	100	21.3	2.93
Other	611	78	20	58	25.6	3.27
1970-1974	580	102	8	94	7.8	1.38
1975-1979	886	132	15	117	11.4	1.69
1980-1984	1,182	146	21	125	14.4	1.78
1985-1989	1,340	163	23	140	14.1	1.72
1990-1994	1,113	126	23	103	18.3	2.07
1995-1999	713	90	20	70	22.2	2.81
2000-2004	989	149	40	109	26.8	4.04
2005-2006	522	73	18	55	24.7	3.45

		Panel B: Additional Statistics				
Variable	Observations	Mean	Std. dev.	Median	Min	Max
ROA	7325	16.0	9.07	15.5	-23.8	85.6
$y_t^* = ROA-v_t$	7325	2.00	7.37	0.75	-35.7	68.0
Assets ($billion)	7325	12.5	55.0	2.38	0.015	1264
Spell length:						
All	981	7.5	4.9	6	1	29
Forced	168	5.1	3.8	4	1	21
Unforced	813	8.0	4.9	7	1	29

Table II: **Parameter Estimates**

Table captions on separate page.

Firm cost $c^{(firm)}$	Personal cost $c^{(pers)}/\kappa$	Prior mean μ_0	Prior stdev. σ_0	Persis- tence ϕ	Profit. stdev. σ_ϵ	z signal stdev. σ_z
1.33	4.61	0.88	2.42	0.125	3.43	5.15
(0.61)	(0.58)	(0.34)	(0.06)	(0.004)	(0.09)	(0.33)

Table III: **Moments used in SMM Estimation**

Table captions on separate page.

Panel A: 14 Moments from SMM Estimation					
Moment	Notation	Empirical value	Simulated value	Standard error	p-value
Profitability intercept	λ_0	0.22	0.20	0.04	0.66
Profitability AR1 coefficient	λ_1	0.87	0.88	0.01	0.38
Δ profitability 2 yrs before	$\Delta^{(-2)}$	-0.37	-0.62	0.27	0.29
Δ profitability 1 yr before	$\Delta^{(-1)}$	-1.36	-1.06	0.27	0.14
Δ profitability 0 yrs before	$\Delta^{(0)}$	-1.73	-1.99	0.26	0.20
Δ profitability 1 yr after	$\Delta^{(1)}$	-1.15	-0.02	0.27	0.00
Δ profitability 2 yrs after	$\Delta^{(2)}$	0.60	0.42	0.27	0.38
Var(Profitability residuals)	$Var(\delta)$	11.99	11.92	0.20	0.60
Forced turn. hazard rate, yrs 1-2	$h^{(1-2)}$	2.52	1.76	0.29	0.00
Forced turn. hazard rate, yrs 3-4	$h^{(3-4)}$	2.66	3.46	0.46	0.03
Forced turn. hazard rate, yrs 5-7	$h^{(5-7)}$	2.10	2.71	0.39	0.10
Forced turn. hazard rate, yrs 8+	$h^{(8+)}$	1.96	1.47	0.26	0.04
Cross-CEO avg. of within-CEO profitability variance	$E[Var(X)]$	799.3	782.1	73.5	0.77
Cross-CEO variance of within-CEO avg. profitability	$Var[E(X)]$	221.5	183.5	20.9	0.11
Panel B: Test of over-identifying restrictions					
$\chi^2 =$	33.2		p-value=	0.000	

58

Table IV: **Additional Statistics on Model Fit**

Table captions on separate page.

	Median tenure		%	% forced	Probit model		
	Unforced	Forced	forced	per year	Slope	Stderr.	Pseudo-R^2
Empirical	7	4	17.1	2.29	-0.168	[0.026]	0.03
Simulated	7	4	16.2	2.16	-0.125	[0.002]	0.02

Table V: **Estimation in Subsamples**

Table captions on separate page.

	Panel A: Parameter Estimates in Subsamples								
	Firm cost	Personal cost	Prior mean	Prior stdev.	Persis-tence	Profit. stdev.	z signal stdev.		χ^2
Subsample	$c^{(firm)}$	$c^{(pers)}/\kappa$	μ_0	σ_0	ϕ	σ_ϵ	σ_z	CEOs	(p-val.)
Small firms	0.01	8.53	1.38	3.26	0.120	4.01	5.34	543	18.7
	(1.05)	(1.06)	(0.69)	(0.05)	(0.004)	(0.14)	(0.24)		(0.01)
Large firms	1.71	-0.29	0.35	1.23	0.135	2.89	7.65	438	29.9
	(0.64)	(0.74)	(0.37)	(0.03)	(0.002)	(0.10)	(0.27)		(0.00)
1971-1989	0.05	8.32	0.49	2.52	0.125	3.30	4.37	222	21.3
	(0.91)	(1.02)	(0.44)	(0.05)	(0.003)	(0.10)	(0.20)		(0.00)
1990-2006	1.67	2.28	1.24	2.72	0.123	3.61	9.51	222	27.5
	(0.70)	(0.76)	(0.67)	(0.07)	(0.002)	(0.14)	(0.46)		(0.00)
Low ownership	0.04	6.41	-0.76	3.10	0.126	3.10	12.92	327	16.1
	(1.07)	(1.09)	(0.60)	(0.11)	(0.003)	(0.11)	(0.93)		(0.02)
High ownership	0.00	7.99	0.14	3.99	0.088	3.35	8.62	325	8.6
	(1.19)	(1.54)	(0.94)	(0.18)	(0.004)	(0.15)	(1.03)		(0.29)
Less outsiders	0.00	8.25	1.57	2.93	0.117	3.69	5.84	491	9.7
	(1.26)	(1.69)	(0.67)	(0.11)	(0.004)	(0.18)	(0.94)		(0.21)
More outsiders	0.00	3.00	-0.56	1.98	0.116	2.87	10.80	490	14.7
	(0.85)	(0.99)	(0.47)	(0.07)	(0.002)	(0.10)	(0.65)		(0.04)

	Panel B: Model Fit in Subsamples							
		Median tenure		%	%		Probit model		
Subsample		Unforced	Forced	forced	forced/yr	$\sigma(E[X])$	Slope	Stderr.	Pseudo-R^2
Small firms	Data	7	4	15.7	2.14	19.1	-0.143	0.029	0.04
	Model	7	3	16.6	2.23	17.1	-0.089	0.002	0.01
Large firms	Data	7	4	18.6	2.45	10.3	-0.207	0.050	0.03
	Model	7	5	19.1	2.59	10.0	-0.252	0.003	0.04
1971-1989	Data	7	5	12.3	1.68	14.6	-0.121	0.031	0.03
	Model	7	4	10.2	1.30	12.7	-0.072	0.002	0.02
1990-2006	Data	7	4	23.1	3.03	16.1	-0.242	0.045	0.04
	Model	7	4	23.2	3.27	15.4	-0.225	0.002	0.04
Low ownership	Data	8	5	12.6	1.51	12.3	-0.161	0.040	0.06
	Model	7	5	12.1	1.54	12.1	-0.161	0.002	0.08
High ownership	Data	8	4	17.6	2.14	24.6	-0.121	0.031	0.03
	Model	7	4	16.6	2.22	19.8	-0.072	0.002	0.02
Less outsiders	Data	7	4.5	11.6	1.58	17.9	-0.080	0.038	0.02
	Model	7	4	12.0	1.54	14.8	-0.074	0.002	0.02
More outsiders	Data	8	5	16.0	2.05	10.4	-0.200	0.043	0.05
	Model	7	5	13.7	1.77	10.7	-0.171	0.002	0.05

Table VI: **The Effect of Entrenchment on Shareholder Value**
Table captions on separate page.

Experiment	% of CEOs fired per year		Mean CEO tenure (years)		Mean profitability (%/year)		% effect on shareholder value
	Control	Treat.	Control	Treat.	Control	Treat.	
1. Eliminate entrenchment							
Full sample	2.1	12.3	7.5	4.5	15.5	16.0	+3.1
1971-1989	1.3	25.7	7.9	2.9	15.2	16.2	+6.9
1990-2006	3.3	8.7	7.1	5.4	15.8	16.0	+1.4
Low ownership	1.6	25.1	7.8	3.0	13.9	14.6	+5.1
High ownership	2.2	26.3	7.5	2.9	15.1	16.4	+8.8
Few outsiders	1.6	26.4	7.8	2.9	16.3	17.4	+6.5
More outsiders	1.8	25.4	7.7	3.0	14.0	14.3	+2.5
Small firms	2.3	26.6	7.4	2.8	16.3	17.6	+7.7
Large firms	2.6	1.8	7.4	7.7	14.6	14.6	+0.0
2. Hire a very bad CEO							
Pers. cost = estimate	2.1	17.3	7.5	4.0	15.5	13.9	−6.3
Pers. cost = zero	12.3	57.1	4.5	1.6	16.0	14.4	−4.6

Table VII: **Alternate Specifications**

Table captions on separate page.

Specification	Firm cost $c^{(firm)}$	Personal cost $c^{(pers)}/\kappa$	Prior mean μ_0	Prior stdev. σ_0	Persis-tence ϕ	Profit. stdev. σ_ϵ	z signal stdev. σ_z	CEOs	χ^2 (p-val.)
Main results	1.33	4.61	0.88	2.42	0.125	3.43	5.15	981	33.2
	(0.61)	(0.58)	(0.34)	(0.06)	(0.004)	(0.09)	(0.33)		(0.00)
Alt. fired def'n	0.00	1.34	1.46	1.24	0.125	3.45	6.08	884	398.6
	(0.41)	(0.34)	(0.38)	(0.03)	(0.003)	(0.08)	(0.33)		(0.00)
Fixed effects	1.00	5.13	-0.10	1.61	0.262	3.31	2.90	981	39.1
	(0.53)	(0.55)	(0.19)	(0.03)	(0.004)	(0.08)	(0.15)		(0.00)
30 industries	0.86	4.05	0.89	2.30	0.134	3.46	6.29	981	38.7
	(0.59)	(0.52)	(0.40)	(0.05)	(0.002)	(0.09)	(0.19)		(0.00)
Flat threshold	0.71	5.28	0.87	2.94	0.132	3.32	6.74	981	33.1
	(0.67)	(0.52)	(0.34)	(0.05)	(0.004)	(0.11)	(0.47)		(0.00)
$c_{(retire)}=0$	0.01	5.94	0.67	3.34	0.128	3.42	9.71	981	36.0
	(0.57)	(0.53)	(0.32)	(0.04)	(0.001)	(0.08)	(0.35)		(0.00)
$\beta = 0.85$	1.21	2.68	0.81	2.47	0.129	3.45	6.30	981	37.2
	(0.57)	(0.60)	(0.38)	(0.04)	(0.002)	(0.08)	(0.23)		(0.00)
$\beta = 0.95$	1.60	5.33	1.00	2.01	0.129	3.37	5.06	981	43.3
	(0.52)	(0.56)	(0.39)	(0.05)	(0.003)	(0.11)	(0.37)		(0.00)

Internet Appendix to
"Why Are CEOs Rarely Fired?
Evidence from Structural Estimation"

LUCIAN A. TAYLOR*

Contents:

* The Wharton School, University of Pennsylvania, luket@wharton.upenn.edu.

Appendix A: The Board's Learning Problem

This Appendix solves the board's learning problem, which is a Kalman filtering problem. I use the notation $\kappa_\epsilon \equiv \sigma_\epsilon^2 / \left(\phi^2 \sigma_0^2\right)$, and $\kappa_z \equiv \sigma_z^2 / \sigma_0^2$. The surprises in the additional signal and persistence-adjusted profitability equal

$$
\begin{align}
\delta_{z,t} &\equiv z_t - \mu_t \tag{A1} \\
\delta_{y,t} &\equiv \frac{1}{\phi}\left(y_t - y_{t-1}\right) + y_{t-1} - \mu_t = \alpha + \frac{1}{\phi}\epsilon_t - \mu_t. \tag{A2}
\end{align}
$$

Standard results on Bayesian learning (e.g. Zellner 1971) imply that $\sigma^2(\tau)$, the board's variance of ability α after τ periods of learning has occurred, decays monotonically and deterministically with tenure according to

$$
\sigma^2(\tau) = \sigma_0^2 \left[1 + \tau\left(\kappa_\epsilon^{-1} + \kappa_z^{-1}\right)\right]^{-1}. \tag{A3}
$$

Then we have

$$
\begin{align}
\mu_{t+1} &= \mu_t + \delta_{y,t}\theta_y(\tau_t) + \delta_{z,t}\theta_z(\tau_t) \tag{A4} \\
\theta_y(\tau) &\equiv \frac{\sigma^2(\tau)\phi^2}{\sigma_\epsilon^2}\left(1 + \sigma^2(\tau)\phi^2/\sigma_\epsilon^2 + \sigma^2(\tau)/\sigma_z^2\right)^{-1} \tag{A5} \\
&= \kappa_\epsilon^{-1}\left(1 + (\tau+1)\left(\kappa_\epsilon^{-1} + \kappa_z^{-1}\right)\right)^{-1} \tag{A6} \\
\theta_z(\tau) &= \kappa_z^{-1}\left(1 + (\tau+1)\left(\kappa_\epsilon^{-1} + \kappa_z^{-1}\right)\right).^{-1} \tag{A7}
\end{align}
$$

The posterior mean follows a random walk with no drift. The board rationally ignores the industry component of profitability, v_t, which contains no information about the CEO's skill. Also, the board adjusts for persistence in profitability (Equation (A2)).

Next I compare the influence of the profitability signal and additional z signal on the board's beliefs about CEO skill. Specifically, I compare the change in posterior beliefs resulting from a one standard deviation z shock and a one standard deviation profitability signal shock. The model predicts that the response to the z shock is $P \equiv \sigma_\epsilon/(\phi\sigma_z)$ times larger than the response to the profitability signal shock. This result follows from equations (A4)-(A7). A one standard deviation z shock corresponds to $\delta_z = \sigma_z$, which moves beliefs by $\theta_z(\tau)\sigma_z$. A one standard deviation X shock corresponds to $\delta_X = \sigma_\epsilon/\phi$, which moves beliefs by $\theta_X(\tau)\sigma_\epsilon/\phi$. Taking ratios,

$$\frac{\theta_z(\tau)\sigma_z}{\theta_X(\tau)\sigma_\epsilon/\phi} = \frac{\kappa_z^{-1}\sigma_z}{\kappa_X^{-1}\sigma_\epsilon/\phi} = \frac{\sigma_\epsilon}{\sigma_z\phi} \equiv P. \tag{A8}$$

Appendix B: Bellman Equation for the Board's Optimization Problem

This Appendix provides the Bellman equation for the board's optimization problem. I introduce notation to distinguish between μ_t^{inc}, the posterior mean of the incumbent CEO's skill α going into period t, and μ_t, the prior mean of the CEO chosen to serve in period t. If the firm decides not to fire the incumbent, then $\mu_t = \mu_t^{inc}$, otherwise $\mu_t = \mu_0$.

Proposition 1 *(Bellman equation): The board's objective function can be simplified as*

$$\frac{U_t}{\kappa B_t} = E_t\left[\sum_{s=0}^{\infty}\beta^s v_{t+s}\right] + \left(\frac{1-\phi}{1-\beta(1-\phi)}\right)y_{t-1} + \tag{A9}$$

$$\left(\frac{\phi}{1-\beta(1-\phi)}\right)\left(\frac{1}{1-\beta}\right)\mu_0 + V\left(\eta_t^{inc}, \tau_t, b_t\right) \tag{A10}$$

where $\eta_t^{inc} = \mu_t^{inc} - \mu_0$, and the value function $V(\eta, \tau, 0)$ solves the Bellman equation

$$V(\eta, \tau, 0) = \max\{V_{fire}, V_{keep}(\eta, \tau)\}, \tag{A11}$$

$$V_{fire} = V(0, 0, 0) - c \tag{A12}$$

$$c \equiv c^{(firm)} + c^{(pers)}/\kappa \tag{A13}$$

$$V_{keep}(\eta, \tau) = \left(\frac{\phi}{1 - \beta(1 - \phi)}\right)\eta + \beta f(\tau) V(\eta, \tau, 1) + \tag{A14}$$

$$\beta(1 - f(\tau)) E\left[V(\eta + \theta_X(\tau)\delta_X + \theta_z(\tau)\delta_z, \tau + 1, 0)\right] \tag{A15}$$

$$\begin{pmatrix} \delta_y \\ \delta_z \end{pmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\epsilon^2/\phi^2 + \sigma^2(\tau) & 0 \\ 0 & \sigma_z^2 + \sigma^2(\tau) \end{bmatrix}\right),$$

subject to a boundary condition if the CEO has just retired:

$$V(\eta, \tau, 1) = V(0, 0, 0) - c. \tag{A16}$$

Proof: I distinguish between total turnover costs from forced turnover (c_{fire}) and total turnover costs from voluntary turnover (c_{retire}). In my main model results and estimation, I set $c_{fire} = c_{retire} = c$. In the robustness section, I allow $c_{fire} \neq c_{retire}$, so separating the two here is useful. Substituting equation (A13) into (4), and then substituting the result into (3), the board's optimization problem is

$$\max_{\{d_{t+s}\}_{s=0}^{\infty}} U_t = \max_{\{d_{t+s}\}_{s=0}^{\infty}} \kappa E_t\left[\sum_{s=0}^{\infty} \beta^s B_{t+s}(v_{t+s} + y_{t+s} - d_{t+s}c_{fire} - b_{t+s}c_{retire})\right], \tag{A17}$$

where d_t and b_t are indicator variables equal to 1 if the CEO is fired or retired, respectively, in period t. Since the firm pays out profits immediately as dividends, the firm's book value

is constant over time, so $B_{t+s} = B_t$ and

$$\max_{\{d_{t+s}\}_{s=0}^{\infty}} \frac{U_t}{\kappa B_t} = \max_{\{d_{t+s}\}_{s=0}^{\infty}} E_t \left[\sum_{s=0}^{\infty} \beta^s \left(v_{t+s} + y_{t+s} - d_{t+s} c_{fire} - b_{t+s} c_{retire} \right) \right] \quad \text{(A18)}$$

$$= E_t \left[\sum_{s=0}^{\infty} \beta^s v_{t+s} \right] + VF_t, \quad \text{(A19)}$$

$$VF_t = \max_{\{d_{t+s}\}_{s=0}^{\infty}} E_t \left[\sum_{s=0}^{\infty} \beta^s \left(y_{t+s} - d_{t+s} c_{fire} - b_{t+s} c_{retire} \right) \right]. \quad \text{(A20)}$$

Next I write y_{t+s} as a function of y_{t-1}, shocks, and future posterior means:

$$y_t = y_{t-1} (1 - \phi) + \phi \mu_t + \phi \delta_{y,t} \quad \text{(A21)}$$

$$y_{t+1} = \left[y_{t-1} (1 - \phi) + \phi \mu_t + \phi \delta_{y,t} \right] (1 - \phi) + \phi \mu_{t+1} + \phi \delta_{y,t+1} \quad \text{(A22)}$$

$$\vdots \quad \text{(A23)}$$

$$y_{t+s} = y_{t-1} (1 - \phi)^{s+1} + \phi \sum_{\tau=0}^{s} \mu_{t+\tau} (1 - \phi)^{s-\tau} + \phi \sum_{\tau=0}^{s} \delta_{y,t+\tau} (1 - \phi)^{s-\tau} \quad \text{(A24)}$$

$$E_t \left[y_{t+s} \right] = y_{t-1} (1 - \phi)^{s+1} + E_t \left[\phi \sum_{\tau=0}^{s} \mu_{t+\tau} (1 - \phi)^{s-\tau} \right], \quad \text{(A25)}$$

since $E_t [\delta_{y,t+\tau}] = E_t [E_{t+\tau} [\delta_{y,t+\tau}]]$ and $E_{t+\tau} [\delta_{y,t+\tau}] = 0$. Next, we have

$$E_t \left[\sum_{s=0}^{\infty} \beta^s y_{t+s} \right] = \sum_{s=0}^{\infty} \beta^s E_t [y_{t+s}] \quad \text{(A26)}$$

$$= \sum_{s=0}^{\infty} \beta^s \left[y_{t-1} (1 - \phi)^{s+1} + E_t \left[\phi \sum_{\tau=0}^{s} \mu_{t+\tau} (1 - \phi)^{s-\tau} \right] \right] \quad \text{(A27)}$$

$$= y_{t-1} (1 - \phi) \sum_{s=0}^{\infty} \beta^s (1 - \phi)^s + \phi \sum_{s=0}^{\infty} \sum_{\tau=0}^{s} \beta^s (1 - \phi)^{s-\tau} E_t [\mu_{t+\tau}] \quad \text{(A28)}$$

$$= \left(\frac{1 - \phi}{1 - \beta (1 - \phi)} \right) y_{t-1} + \left(\frac{\phi}{1 - \beta (1 - \phi)} \right) \sum_{s=0}^{\infty} \beta^s E_t [\mu_{t+s}] \quad \text{(A29)}$$

$$= \left(\frac{1 - \phi}{1 - \beta (1 - \phi)} \right) y_{t-1} + \quad \text{(A30)}$$

$$\left(\frac{\phi}{1 - \beta (1 - \phi)} \right) \sum_{s=0}^{\infty} \beta^s \left(\mu_0 + E_t [\eta_{t+s}] \right), \quad \text{(A31)}$$

where I have used the relation:

$$\mu_t = \mu_0 + \eta_t. \tag{A32}$$

In sum, we have

$$E_t \left[\sum_{s=0}^{\infty} \beta^s y_{t+s} \right] = \left(\frac{1-\phi}{1-\beta(1-\phi)} \right) y_{t-1} + \left(\frac{\phi}{1-\beta(1-\phi)} \right) \left(\frac{1}{1-\beta} \right) \mu_0 \tag{A33}$$

$$+ \left(\frac{\phi}{1-\beta(1-\phi)} \right) \sum_{s=0}^{\infty} \beta^s E_t \left[\eta_{t+s} \right] \tag{A34}$$

Plugging this into the expression for VF,

$$VF_t \equiv \left(\frac{1-\phi}{1-\beta(1-\phi)} \right) y_{t-1} + \left(\frac{\phi}{1-\beta(1-\phi)} \right) \left(\frac{1}{1-\beta} \right) \mu_0 + V_t^*, \tag{A35}$$

$$V_t^* = \max_{d_t} \left\{ \left(\frac{\phi}{1-\beta(1-\phi)} \right) \eta_t - d_t c_{fire} - b_t c_{retire} + \beta E_t \left[V_{t+1}^* \right] \right\}, \tag{A36}$$

so

$$V \left(\eta_t^{inc}, \tau_t, b_t \right) = \max_{d_t} \left\{ \left(\frac{\phi}{1-\beta(1-\phi)} \right) \eta_t - d_t c_{fire} - b_t c_{retire} + \beta E_t \left[V \left(\eta_{t+1}^{inc}, \tau_{t+1}, b_{t+1} \right) \right] \right\}. \tag{A37}$$

If the incumbent CEO has just retired, the firm hires a new CEO ($\eta = 0$) and pays the retirement cost:

$$V_{retire} = V \left(\eta_t^{inc}, \tau_t, 1 \right) = V(0, 0, 0) - c_{retire}. \tag{A38}$$

Otherwise, if $b_t = 0$ and $d_t = 1$ (the firm fires its CEO), then the firm hires a new CEO and pays the firing cost:

$$V_{fire} \left(\eta_t^{inc}, \tau_t, 0 \right) = V(0, 0, 0) - c_{fire}. \tag{A39}$$

If $b_t = 0$ and $d_t = 0$ (the firm keeps its CEO), then

$$V_{keep}\left(\eta_t^{inc}, \tau_t, 0\right) = \left(\frac{\phi}{1 - \beta\left(1 - \phi\right)}\right)\eta_t^{inc} + \beta E_t\left[V\left(\eta_{t+1}^{inc}, \tau_{t+1}, b_{t+1}\right)\right] \quad \text{(A40)}$$

$$= \left(\frac{\phi}{1 - \beta\left(1 - \phi\right)}\right)\eta_t^{inc} + \beta f\left(\tau_t\right) V^{retire} + \quad \text{(A41)}$$

$$\beta\left(1 - f\left(\tau_t\right)\right) E_t\left[V\left(\eta_{t+1}^{inc}, \tau_{t+1}, 0\right)\right]. \quad \text{(A42)}$$

The firm chooses d_t (fire or keep CEO) according to

$$V\left(\eta_t^{inc}, \tau_t, 0\right) = \max\left\{V^{fire}\left(\eta_t^{inc}, \tau_t, 0\right), V^{keep}\left(\eta_t^{inc}, \tau_t, 0\right)\right\}. \quad \text{(A43)}$$

Recalling from equation (A4) that

$$\mu_{t+1}^{inc} = \mu_t^{inc} + \theta_y\left(\tau_t\right)\delta_{y,t} + \theta_z\left(\tau_t\right)\delta_{z,t} \quad \text{(A44)}$$

$$\mu_0 + \eta_{t+1}^{inc} = \mu_0 + \eta_t^{inc} + \theta_y\left(\tau_t\right)\delta_{y,t} + \theta_z\left(\tau_t\right)\delta_{z,t} \quad \text{(A45)}$$

$$\eta_{t+1}^{inc} = \eta_t^{inc} + \theta_y\left(\tau_t\right)\delta_{y,t} + \theta_z\left(\tau_t\right)\delta_{z,t} \quad \text{(A46)}$$

I write the Bellman in its final form by dropping time and incumbent subscripts and substituting in for V^{retire}. End of proof.

Equation (A9) shows that the board's objective function is the sum of an industry-specific component, a component due to persistence in profitability, and a component V which depends on the CEO's posterior mean skill and tenure in office. Each period the board makes a firing decision by comparing its utility from firing the CEO (V_{fire}) and not firing him (V_{keep}) (Equation (A11)). Expression (A12) shows that after firing the CEO, the board hires a new one and incurs the firing cost; the firing utility V_{fire} is constant over time. The board's decision depends on the total κ-adjusted turnover cost, defined in equation (A13), not on the firm and personal costs separately. In equation (A14), the utility V_{keep} from keeping the CEO depends on his expected contribution this period (the μ term) and the expected utility V next period, which in turn depends on whether the CEO quits (with

probability $f(\tau)$) at the end of the period. If the CEO does not quit, he enters next period with posterior mean (minus the prior) equal to $\eta' = \eta + \theta_X(\tau)\delta_X + \theta_z(\tau)\delta_z$ (from the learning rule), and one more year of tenure (hence $\tau + 1$). The boundary condition in equation (A16) shows that following a voluntary succession, the board hires a new CEO and pays cost c. The prior mean μ_0 drops out of the Bellman equation, which still depends on η, distance between the posterior mean and the prior mean.

Appendix C: Numerical Solution of Bellman Equation

This Appendix describes how I numerically solve the Bellman equation to find the board's optimal CEO firing rule. I obtain an approximate solution for $V(\mu, \tau, 0)$ by discretizing the state space and iterating on the Bellman equation.

I approximate the value function using the Jacobi Iteration method. I start by discretizing the state space. State variable τ_t takes values in set $\varsigma = \{0, 1, ..., \bar{\tau} - 1\}$, where $\bar{\tau} = \sup \tau$ is the maximum possible number of terms in office. I let μ takes values in finite set M, which contains 1,001 equally spaced points in the interval $[\mu_0 - c_{fire} - 2\sigma_0, \mu_0 + c_{fire} + 2\sigma_0]$; the length of the interval does not need to be extremely large, because the extrapolation used below ends up being quite accurate. To speed up the iteration, I start with a guess of V^0 over the grid $\varsigma \times M$:

$$V^0(\mu, \tau, 0) = \left(\frac{\phi}{1 - \beta(1 - \phi)}\right)\left[\frac{\mu_0}{(1 - \beta)} + \max(\mu - \mu_0, 0)\frac{1 - \beta^{\bar{\tau} - \tau}}{1 - \beta}\right]. \qquad (A47)$$

Then I update the value function according to

$$V^{t+1}(\mu, \tau, 0) = \max\{V^t(\mu_0, 0, 0) - c_{fire}, \left(\frac{\phi}{1 - \beta(1 - \phi)}\right)\mu + \qquad (A48)$$

$$\beta f(\tau)\left[V^t(\mu_0, 0, 0) - c_{retire}\right] + \qquad (A49)$$

$$\beta(1 - f(\tau))E\left[V^t(\mu + \theta_X(\tau)\delta_X + \theta_z(\tau)\delta_z, \tau + 1, 0)\right]\}. \qquad (A50)$$

$$\begin{pmatrix} \delta_X \\ \delta_z \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\epsilon^2/\phi^2 + \sigma^2(\tau) & 0 \\ 0 & \sigma_z^2 + \sigma^2(\tau) \end{bmatrix} \right) \tag{A51}$$

I approximate the expectation above using Gauss-Hermite quadrature, as follows. Recall $V^t(\mu, \tau)$ is defined only for μ in the finite set M. First, I create a function $\widehat{V}^t(\mu, \tau)$ which is defined for all $\mu \in \mathbb{R}$ by performing piecewise cubic spline interpolation and extrapolation of the function $V^t(\mu, \tau)$. Second, I apply two-dimensional Gauss-Hermite quadrature with 7 nodes as follows: For each $\mu \in M$ and $\tau = 0, 1, ..., \bar{\tau} - 1$,

$$E\left[V^t(\mu + \theta_X(\tau)\delta_X + \theta_z(\tau)\delta_z, \tau + 1, 0) \right] \tag{A52}$$

$$\approx \pi^{-1} \sum_{i=1}^{7} \sum_{j=1}^{7} \omega_i \omega_j \widehat{V}^t(\mu + \theta_X(\tau) \left[\sqrt{2(\sigma_\epsilon^2/\phi^2 + \sigma^2(\tau))} x_i \right] + \tag{A53}$$

$$\theta_z(\tau) \left[\sqrt{2(\sigma_z^2 + \sigma^2(\tau))} x_j \right], \tau + 1, 0) \tag{A54}$$

where $\{x_i\}$ and $\{\omega_i\}$ are the Gauss-Hermite quadrature nodes and weights, respectively. I stop iterating as soon as

$$\max_{(\tau, \mu) \in \varsigma \times M} \left| V^{t+s} - V^t \right| < 10^{-5}. \tag{A55}$$

Appendix D: Simulation Method

I define a CEO spell as all the periods a CEO serves in office. To simulate a single spell, I draw the CEO's true skill α from the prior distribution, I generate firm-specific profitability y_t and additional signals z_t using the CEO's true skill α, and I update the board's beliefs according to the learning rule in equation (AA4). Simulated CEOs are fired according to the optimal rule from the Bellman equation, and they leave office voluntarily with probability $f(\tau)$.

Appendix E: Additional Details on SMM Estimation

I use the optimal weighting matrix

$$W = \left[N \mathrm{var} \left(\widehat{M}_N \right) \right]^{-1}. \tag{A56}$$

I compute the 14x14 covariance matrix $\widehat{M}_N$ using the seemingly unrelated regressions approach. The moments can be expressed as the coefficients from the following system of regression equations:

$$
\begin{aligned}
y_{it}^* &= \lambda_0 + \lambda_1 y_{it-1}^* + & (A57) \\
&\quad \Delta^{(-2)} + \Delta^{(-1)} + \Delta^{(0)} + \Delta^{(1)} + \Delta^{(2)} + \delta_{it} & (A58) \\
\delta_{it}^2 &= Var\left(\delta\right) + w_{it} & (A59) \\
d_{it} &= h^{(1-2)} + h^{(2-3)} + h^{(4-6)} + h^{(7+)} + \eta_{it} & (A60) \\
Var_i\left(X_{it}\right) &= E\left[Var\left(X\right)\right] + e_i & (A61) \\
\left(E_i\left[X_{it}\right] - E\left[E_i\left[X_{it}\right]\right]\right)^2 &= Var\left(E\left[X\right]\right) + \iota_i & (A62)
\end{aligned}
$$

The coefficients $h^{(j)}$ are fixed effects for tenure (j). Var_i denotes variance within CEO spell i, and E_i denotes average within CEO spell i. I estimate each regression separately using ordinary least squares, which provides consistent estimates for each moment as well as regression disturbances. Each regression above has the form

$$Y_i = X_i \beta_i + \varepsilon_i, \tag{A63}$$

where Y_i is $N_i \times 1$ and β_i is $k_i \times 1$. The covariance between moments estimators β_i and β_j is the $k_i \times k_j$ matrix

$$Cov\left(\widehat{\beta}_i, \widehat{\beta}_j\right) = \left(X_i'X_i\right)^{-1} X_i' \Omega_{ij} X_j \left(X_j'X_j\right)^{-1}, \tag{A64}$$

where $\Omega_{ij} = Cov\left(\varepsilon_i, \varepsilon_j\right)$ is the $N_i \times N_j$ matrix whose element t, s is $Cov\left(\varepsilon_{it}, \varepsilon_{js}\right)$. I estimated the covariance matrix Ω_{ij} for each pair of moments ij, allowing for time series autocorrelation and also correlation across regressions.

I define

$$G_N = M_N - \frac{1}{S}\sum_{s=1}^{S} m_n^s(\theta). \tag{A65}$$

Applying the result of Pakes and Pollard (1989) with the efficient weighting matrix, we obtain

$$\sqrt{N}\left(\widehat{\theta} - \theta_0\right) \rightarrow \ ^d\mathcal{N}\left(0, \Omega\right) \tag{A66}$$

$$\Omega = \left(1 + \frac{1}{S}\right)\left(\Gamma'\Lambda^{-1}\Gamma\right)^{-1}, \tag{A67}$$

where S is the number of simulated data sets (I choose $S = 10$), $\Gamma = \text{plim}_{N\to\infty}\, \partial\widehat{G}\left(\theta_0\right)/\partial\theta'$ and $\Lambda = N\text{avar}(\widehat{M}\left(\theta_0\right)) = N\text{avar}(\widehat{m}\left(\theta_0\right))$. I estimate Γ by numerically differentiating $\widehat{G}\left(\widehat{\theta}\right)$ with respect to θ, and using $\widehat{\Lambda} = N\widehat{\text{var}}\left(\widehat{M}\right)$, as described above.

We have

$$\sqrt{N}\widehat{G}\left(\theta_0\right) \rightarrow^d \mathcal{N}\left(0, \left(1 + \frac{1}{S}\right)\Lambda\right), \tag{A68}$$

so SMM provides the following test of the model's over-identifying restrictions:

$$\frac{NS}{1+S}\widehat{G}\left(\theta_0\right)'\Lambda^{-1}\widehat{G}\left(\theta_0\right)' \rightarrow^d \chi^2\left(\#\text{moments - }\#\text{parameters}\right). \tag{A69}$$

Appendix F: Firing threshold when entrenchment increases with tenure

With constant CEO turnover costs, the firing threshold rises with tenure. In the robustness exercise in Section IV.C of Taylor (20??), I estimate the model forcing the threshold to be flat. I do so by getting the model's predicted threshold (which rises with tenure), taking this threshold's value in CEOs' first year in office, and forcing the threshold to be constant and equal to this first-year value at all tenure levels. I argue in that section that a flat threshold may obtain from turnover costs that increase with tenure.

In this Appendix I show how turnover costs must change with tenure in order to obtain a firing threshold that is flat and equal to its value in year one. The main result is that turnover costs must increase with tenure (as expected), and (less expected) the average turnover cost across tenures is higher than the cost estimates I report in Section IV.C of Taylor (20??). This result supports the claim I make in Section IV.C of Taylor (20??): "These conclusions are even stronger in the story where entrenchment increases with tenure, since flattening the firing threshold requires the total turnover cost to start at 5.99% and then increase with tenure, as I show in the Internet Appendix."

I solve an extension of the model in which the total turnover cost $c(\tau)$ is an arbitrary function of tenure τ. The extension is straightforward, since tenure is already a state variable in the board's dynamic optimization problem. To keep this exercise simple, I assume CEOs retire after and only after 10 years in office, and I use the same parameter values I used to create Figure 1 in Taylor (20??). The top panel of Figure 1 plots the firing threshold that obtains when the total turnover cost $c(\tau)$ is constant at 3% of assets at all tenures τ. As expected, the firing threshold increases with tenure. Next I numerically search for values of $c(\tau)$, $t = 1, 2, ..., 9$, so that the firing threshold becomes flat and equal the first-year value of the rising threshold produced by the constant cost $c(\tau) = 3\%$. I do not constrain the

numerical search in any way. The line with O's in the top panel of Figure 1 shows the resulting flat threshold, and the line with O's in the bottom panel shows the time-varying turnover costs that produce this flat threshold. For comparison, the bottom panel also plots the 3% turnover cost that produced the rising threshold in the top panel (line with X's).

The main result is that, to produce the flat threshold, the turnover costs rise from 3.2% at $\tau = 1$ to 4.8% at $\tau = 9$. Since the turnover costs start rising at approximately 3% (the level that produced the rising threshold), the average of these turnover costs across tenures is well above 3%. There are two important caveats. I have not proven that this result obtains for parameter values besides the ones I use in this exercise, nor have I proven that the turnover costs shown in Figure 1 are the unique values that produce the flat threshold in the Figure. However, if these results hold more generally then they have implications for the results in Section IV.C of Taylor (20??). In that robustness section I report that the model with a flat threshold requires an estimated total turnover cost of 5.99% to fit the data. A constant turnover cost of 5.99% will produce a firing threshold that increases with tenure. In that exercise, I take the increasing threshold and make it flat and equal to the first-year value of the increasing threshold. The results of this Appendix suggest that flattening the threshold in this way would require the turnover cost to start near 5.99% in CEOs' first year in office, and then increase with tenure.

REFERENCES

Taylor, Lucian A., 200?, Why are CEOs rarely fired? Evidence from structural estimation, *Journal of Finance* ?, ??–??.

Figure 1: **Firing Threshold with Turnover Costs that Vary with Tenure**
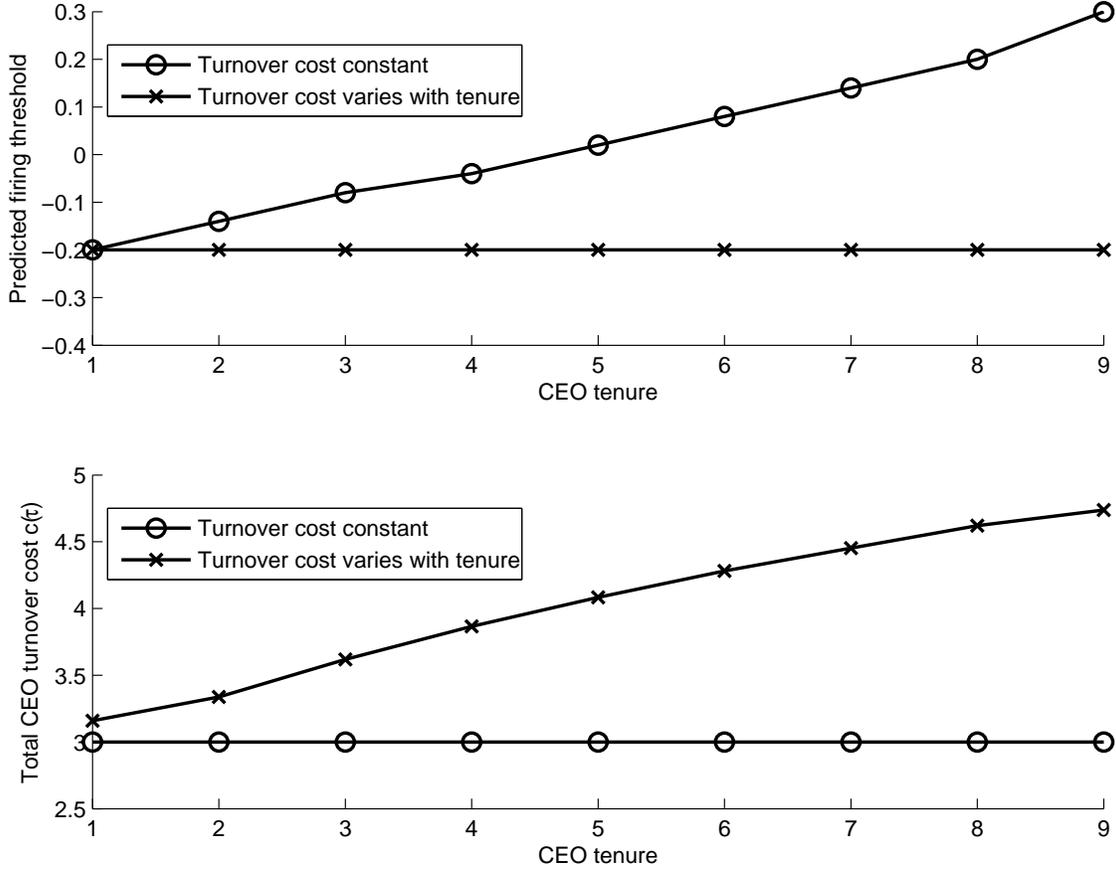
The top panel shows two predicted firing thresholds, as a function of CEO tenure. The line with X's shows the firing threshold produced by constant CEO turnover costs; parameter values are $\beta = 0.9$, $\mu_0 = 1\%$, $\sigma_0 = 2\%$, $\sigma_\epsilon = 3\%$, $c = 3\%$, $\phi = 0.12$, and $\sigma_z = 7\%$, and voluntary turnover occurs after (and only after) completing 10 periods in office, so $f(\tau) = 0$ for $\tau = 0, 1, ..., 8$, $f(9) = 1$. The line with O's shows the predicted firing threshold using those same parameter values, but using time-varying CEO turnover costs $c(\tau)$ as shown in the bottom panel. The line with X's in the bottom panel shows the constant turnover costs used to produce the firing threshold with X's in the top panel. The line with O's in the bottom panel shows the time-varying turnover costs $c(\tau)$ used to produce the flat firing threshold in the top panel.