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

LUCIAN A. TAYLOR

Summary:

The published article estimates a dynamic model of CEO turnover. The Internet Appendix provides additional details on the model's solution (Appendices A through C), the estimation method (Appendices D and E), and robustness of the results (Appendix F).

Contents:

* Citation format: Taylor, Lucian A., 20??, Internet Appendix to "Why are CEOs rarely fired? Evidence from structural estimation," *Journal of Finance* ???, ???–???, http://www.afajof.org/IA/???.asp. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the author of the article.

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

$$\delta_{z,t} \equiv z_t - \mu_t \tag{IA.1}$$

$$\delta_{y,t} \equiv \frac{1}{\phi}\left(y_t - y_{t-1}\right) + y_{t-1} - \mu_t = \alpha + \frac{1}{\phi}\epsilon_t - \mu_t. \tag{IA.2}$$

Standard results on Bayesian learning (e.g., Pástor and Veronesi (2009)) 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

$$\sigma^2(\tau) = \sigma_0^2\left[1 + \tau\left(\kappa_\epsilon^{-1} + \kappa_z^{-1}\right)\right]^{-1}. \tag{IA.3}$$

The posterior mean evolves according to

$$\mu_{t+1} = \mu_t + \delta_{y,t}\theta_y(\tau_t) + \delta_{z,t}\theta_z(\tau_t) \tag{IA.4}$$

$$\theta_y(\tau) \equiv \frac{\sigma^2(\tau)\phi^2}{\sigma_\epsilon^2}\left(1 + \sigma^2(\tau)\phi^2/\sigma_\epsilon^2 + \sigma^2(\tau)/\sigma_z^2\right)^{-1} \tag{IA.5}$$

$$= \kappa_\epsilon^{-1}\left(1 + (\tau+1)\left(\kappa_\epsilon^{-1} + \kappa_z^{-1}\right)\right)^{-1} \tag{IA.6}$$

$$\theta_z(\tau) = \kappa_z^{-1}\left(1 + (\tau+1)\left(\kappa_\epsilon^{-1} + \kappa_z^{-1}\right)\right)^{-1}. \tag{IA.7}$$

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 (IA.2)).

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 (IA.4)-(IA.7). 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,

$$\frac{\theta_z(\tau)\sigma_z}{\theta_X(\tau)\sigma_\epsilon/\phi} = \frac{\kappa_z^{-1}\sigma_z}{\kappa_X^{-1}\sigma_\epsilon/\phi} = \frac{\sigma_\epsilon}{\sigma_z\phi} \equiv P. \tag{IA.8}$$

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*

$$\frac{U_t}{\kappa B_t} = E_t\left[\sum_{s=0}^{\infty}\beta^s v_{t+s}\right] + \left(\frac{1-\phi}{1-\beta(1-\phi)}\right) y_{t-1} + \qquad \text{(IA.9)}$$

$$\left(\frac{\phi}{1-\beta(1-\phi)}\right)\left(\frac{1}{1-\beta}\right)\mu_0 + V\left(\eta_t^{inc}, \tau_t, b_t\right) \qquad \text{(IA.10)}$$

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

$$V(\eta, \tau, 0) = \max\{V_{fire}, V_{keep}(\eta, \tau)\} \qquad \text{(IA.11)}$$

$$V_{fire} = V(0,0,0) - c \qquad \text{(IA.12)}$$

$$c \equiv c^{(firm)} + c^{(pers)}/\kappa \qquad \text{(IA.13)}$$

$$V_{keep}(\eta, \tau) = \left(\frac{\phi}{1-\beta(1-\phi)}\right)\eta + \beta f(\tau) V(\eta, \tau, 1) + \qquad \text{(IA.14)}$$

$$\beta(1 - f(\tau)) E\left[V(\eta + \theta_X(\tau)\delta_X + \theta_z(\tau)\delta_z, \tau + 1, 0)\right] \qquad \text{(IA.15)}$$

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

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

$$V(\eta, \tau, 1) = V(0,0,0) - c. \qquad \text{(IA.16)}$$

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 (IA.13) into (4), and then substituting the result into (3), the board's optimization problem is

$$\max_{\{d_{t+s}\}_{s=0}^{\infty}} U_t = \max_{\{d_{t+s}\}_{s=0}^{\infty}} \kappa E_t\left[\sum_{s=0}^{\infty}\beta^s B_{t+s}\left(v_{t+s} + y_{t+s} - d_{t+s}c_{fire} - b_{t+s}c_{retire}\right)\right], \qquad \text{(IA.17)}$$

where d_t and b_t are indicator variables equal to one 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

$$\max_{\{d_{t+s}\}_{s=0}^{\infty}} \frac{U_t}{\kappa B_t} = \max_{\{d_{t+s}\}_{s=0}^{\infty}} E_t \left[\sum_{s=0}^{\infty} \beta^s \left(v_{t+s} + y_{t+s} - d_{t+s} c_{fire} - b_{t+s} c_{retire} \right) \right] \quad \text{(IA.18)}$$

$$= E_t \left[\sum_{s=0}^{\infty} \beta^s v_{t+s} \right] + VF_t, \quad \text{(IA.19)}$$

$$VF_t = \max_{\{d_{t+s}\}_{s=0}^{\infty}} E_t \left[\sum_{s=0}^{\infty} \beta^s \left(y_{t+s} - d_{t+s} c_{fire} - b_{t+s} c_{retire} \right) \right]. \quad \text{(IA.20)}$$

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

$$y_t = y_{t-1}(1-\phi) + \phi \mu_t + \phi \delta_{y,t} \quad \text{(IA.21)}$$

$$y_{t+1} = \left[y_{t-1}(1-\phi) + \phi \mu_t + \phi \delta_{y,t} \right](1-\phi) + \phi \mu_{t+1} + \phi \delta_{y,t+1} \quad \text{(IA.22)}$$

$$\vdots \quad \text{(IA.23)}$$

$$y_{t+s} = y_{t-1}(1-\phi)^{s+1} + \phi \sum_{\tau=0}^{s} \mu_{t+\tau}(1-\phi)^{s-\tau} + \phi \sum_{\tau=0}^{s} \delta_{y,t+\tau}(1-\phi)^{s-\tau} \quad \text{(IA.24)}$$

$$E_t[y_{t+s}] = y_{t-1}(1-\phi)^{s+1} + E_t \left[\phi \sum_{\tau=0}^{s} \mu_{t+\tau}(1-\phi)^{s-\tau} \right], \quad \text{(IA.25)}$$

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

$$E_t \left[\sum_{s=0}^{\infty} \beta^s y_{t+s} \right] = \sum_{s=0}^{\infty} \beta^s E_t[y_{t+s}] \quad \text{(IA.26)}$$

$$= \sum_{s=0}^{\infty} \beta^s \left[y_{t-1}(1-\phi)^{s+1} + E_t \left[\phi \sum_{\tau=0}^{s} \mu_{t+\tau}(1-\phi)^{s-\tau} \right] \right] \quad \text{(IA.27)}$$

$$= y_{t-1}(1-\phi) \sum_{s=0}^{\infty} \beta^s (1-\phi)^s + \quad \text{(IA.28)}$$

$$\phi \sum_{s=0}^{\infty} \sum_{\tau=0}^{s} \beta^s (1-\phi)^{s-\tau} E_t[\mu_{t+\tau}] \quad \text{(IA.29)}$$

$$= \left(\frac{1-\phi}{1-\beta(1-\phi)} \right) y_{t-1} + \left(\frac{\phi}{1-\beta(1-\phi)} \right) \sum_{s=0}^{\infty} \beta^s E_t[\mu_{t+s}] \quad \text{(IA.30)}$$

$$= \left(\frac{1-\phi}{1-\beta(1-\phi)} \right) y_{t-1} + \quad \text{(IA.31)}$$

$$\left(\frac{\phi}{1-\beta(1-\phi)} \right) \sum_{s=0}^{\infty} \beta^s \left(\mu_0 + E_t[\eta_{t+s}] \right), \quad \text{(IA.32)}$$

where I have used the relation

$$\mu_t = \mu_0 + \eta_t. \tag{IA.33}$$

In sum, we have

$$E_t\left[\sum_{s=0}^{\infty}\beta^s y_{t+s}\right] = \left(\frac{1-\phi}{1-\beta(1-\phi)}\right)y_{t-1} + \left(\frac{\phi}{1-\beta(1-\phi)}\right)\left(\frac{1}{1-\beta}\right)\mu_0 \tag{IA.34}$$

$$+\left(\frac{\phi}{1-\beta(1-\phi)}\right)\sum_{s=0}^{\infty}\beta^s E_t\left[\eta_{t+s}\right]. \tag{IA.35}$$

Plugging this into the expression for VF,

$$VF_t \equiv \left(\frac{1-\phi}{1-\beta(1-\phi)}\right)y_{t-1} + \left(\frac{\phi}{1-\beta(1-\phi)}\right)\left(\frac{1}{1-\beta}\right)\mu_0 + V_t^*, \tag{IA.36}$$

$$V_t^* = \max_{d_t}\left\{\left(\frac{\phi}{1-\beta(1-\phi)}\right)\eta_t - d_t c_{fire} - b_t c_{retire} + \beta E_t\left[V_{t+1}^*\right]\right\}, \tag{IA.37}$$

and so

$$V\left(\eta_t^{inc}, \tau_t, b_t\right) = \max_{d_t}\left\{\left(\frac{\phi}{1-\beta(1-\phi)}\right)\eta_t - d_t c_{fire} - b_t c_{retire} + \beta E_t\left[V\left(\eta_{t+1}^{inc}, \tau_{t+1}, b_{t+1}\right)\right]\right\}. \tag{IA.38}$$

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

$$V_{retire} = V\left(\eta_t^{inc}, \tau_t, 1\right) = V(0,0,0) - c_{retire}. \tag{IA.39}$$

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, yielding

$$V_{fire}\left(\eta_t^{inc}, \tau_t, 0\right) = V(0,0,0) - c_{fire}. \tag{IA.40}$$

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

$$V_{keep}\left(\eta_t^{inc}, \tau_t, 0\right) = \left(\frac{\phi}{1-\beta(1-\phi)}\right)\eta_t^{inc} + \beta E_t\left[V\left(\eta_{t+1}^{inc}, \tau_{t+1}, b_{t+1}\right)\right] \tag{IA.41}$$

$$= \left(\frac{\phi}{1-\beta(1-\phi)}\right)\eta_t^{inc} + \beta f(\tau_t) V^{retire} + \tag{IA.42}$$

$$\beta\left(1 - f(\tau_t)\right) E_t\left[V\left(\eta_{t+1}^{inc}, \tau_{t+1}, 0\right)\right]. \tag{IA.43}$$

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

$$V\left(\eta_t^{inc}, \tau_t, 0\right) = \max\left\{V^{fire}\left(\eta_t^{inc}, \tau_t, 0\right), V^{keep}\left(\eta_t^{inc}, \tau_t, 0\right)\right\}. \tag{IA.44}$$

Recalling from equation (IA.4) that

$$\mu_{t+1}^{inc} = \mu_t^{inc} + \theta_y(\tau_t)\,\delta_{y,t} + \theta_z(\tau_t)\,\delta_{z,t} \tag{IA.45}$$

$$\mu_0 + \eta_{t+1}^{inc} = \mu_0 + \eta_t^{inc} + \theta_y(\tau_t)\,\delta_{y,t} + \theta_z(\tau_t)\,\delta_{z,t} \tag{IA.46}$$

$$\eta_{t+1}^{inc} = \eta_t^{inc} + \theta_y(\tau_t)\,\delta_{y,t} + \theta_z(\tau_t)\,\delta_{z,t}. \tag{IA.47}$$

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

Equation (IA.9) 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 that 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 (IA.11)). Expression (IA.12) 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 (IA.13), not on the firm and personal costs separately. In equation (IA.14), 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 (IA.16) 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, the 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 the set $\varsigma = \{0, 1, ..., \overline{\tau} - 1\}$, where $\overline{\tau} = \sup \tau$ 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 $\varsigma \times M$:

$$V^0(\mu,\tau,0) = \left(\frac{\phi}{1-\beta(1-\phi)}\right)\left[\frac{\mu_0}{(1-\beta)} + \max(\mu-\mu_0,0)\frac{1-\beta^{\overline{\tau}-\tau}}{1-\beta}\right]. \tag{IA.48}$$

I the update the value function according to

$$V^{t+1}(\mu,\tau,0) = \max\{V^t(\mu_0,0,0) - c_{fire}, \left(\frac{\phi}{1-\beta(1-\phi)}\right)\mu + \tag{IA.49}$$

$$\beta f(\tau)\left[V^t(\mu_0,0,0) - c_{retire}\right] + \tag{IA.50}$$

$$\beta(1-f(\tau))E\left[V^t(\mu+\theta_X(\tau)\delta_X+\theta_z(\tau)\delta_z,\tau+1,0)\right]\}, \tag{IA.51}$$

$$\begin{pmatrix}\delta_X\\ \delta_z\end{pmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\sigma_\epsilon^2/\phi^2+\sigma^2(\tau) & 0\\ 0 & \sigma_z^2+\sigma^2(\tau).\end{bmatrix}\right) \tag{IA.52}$$

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 $\widehat{V}^t(\mu,\tau)$ that 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 seven nodes as follows: For each $\mu\in M$ and $\tau=0,1,...,\overline{\tau}-1$,

$$E\left[V^t(\mu+\theta_X(\tau)\delta_X+\theta_z(\tau)\delta_z,\tau+1,0)\right] \tag{IA.53}$$

$$\approx \pi^{-1}\sum_{i=1}^{7}\sum_{j=1}^{7}\omega_i\omega_j\widehat{V}^t(\mu+\theta_X(\tau)\left[\sqrt{2(\sigma_\epsilon^2/\phi^2+\sigma^2(\tau))}x_i\right] + \tag{IA.54}$$

$$\theta_z(\tau)\left[\sqrt{2(\sigma_z^2+\sigma^2(\tau))}x_j\right],\tau+1,0), \tag{IA.55}$$

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

$$\max_{(\tau,\mu)\in\varsigma\times M}\left|V^{t+s}-V^t\right| < 10^{-5}. \tag{IA.56}$$

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 (IA.4). 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

$$W = \left[N\text{var}\left(\widehat{M_N}\right)\right]^{-1}. \tag{IA.57}$$

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:

$$\begin{aligned}
y_{it}^* &= \lambda_0 + \lambda_1 y_{it-1}^* + && \text{(IA.58)}\\
& \quad \Delta^{(-2)} + \Delta^{(-1)} + \Delta^{(0)} + \Delta^{(1)} + \Delta^{(2)} + \delta_{it} && \text{(IA.59)}\\
\delta_{it}^2 &= Var(\delta) + w_{it} && \text{(IA.60)}\\
d_{it} &= h^{(1-2)} + h^{(2-3)} + h^{(4-6)} + h^{(7+)} + \eta_{it} && \text{(IA.61)}\\
Var_i(X_{it}) &= E[Var(X)] + e_i && \text{(IA.62)}\\
(E_i[X_{it}] - E[E_i[X_{it}]])^2 &= Var(E[X]) + \iota_i. && \text{(IA.63)}
\end{aligned}$$

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

$$Y_i = X_i\beta_i + \varepsilon_i, \tag{IA.64}$$

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

$$Cov\left(\widehat{\beta}_i, \widehat{\beta}_j\right) = \left(X_i'X_i\right)^{-1} X_i'\Omega_{ij}X_j\left(X_j'X_j\right)^{-1}, \tag{IA.65}$$

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

I define

$$G_N = M_N - \frac{1}{S}\sum_{s=1}^{S} m_n^s(\theta). \tag{IA.66}$$

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

$$\begin{aligned}
\sqrt{N}\left(\widehat{\theta} - \theta_0\right) &\rightarrow {}^d\mathcal{N}(0, \Omega) && \text{(IA.67)}\\
\Omega &= \left(1 + \frac{1}{S}\right)\left(\Gamma'\Lambda^{-1}\Gamma\right)^{-1}, && \text{(IA.68)}
\end{aligned}$$

where S is the number of simulated data sets (I choose $S = 10$), $\Gamma = \text{plim}_{N\to\infty}\,\partial\widehat{G}(\theta_0)/\partial\theta'$, and $\Lambda = N\text{avar}(\widehat{M}(\theta_0)) = N\text{avar}(\widehat{m}(\theta_0))$. 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

$$\sqrt{N}\widehat{G}(\theta_0) \to^d \mathcal{N}\left(0, \left(1+\frac{1}{S}\right)\Lambda\right), \tag{IA.69}$$

so SMM provides the following test of the model's overidentifying restrictions:

$$\frac{NS}{1+S}\widehat{G}(\theta_0)'\,\Lambda^{-1}\widehat{G}(\theta_0)' \to^d \chi^2\left(\text{\#moments - \#parameters}\right). \tag{IA.70}$$

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 the published paper, I estimate the model forcing the threshold to be flat. I do so by calculating the model's predicted threshold (which rises with tenure), taking this threshold's value in a CEO's 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 the published paper. This result supports the claim I make in Section IV.C of the published paper: "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 the published paper. The top panel of Figure IA.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 to 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 IA.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 IA.1 are the unique values that produce the flat threshold in the figure. However, to the extent that these results hold more generally, they have implications for the results in Section IV.C of the published paper. 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 a CEO's first year in office, and then increase with tenure.

REFERENCES

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

Pástor, Ľuboš, and Pietro Veronesi, 2009, Learning in financial markets, *Annual Review of Financial Economics* 1, 361–381.

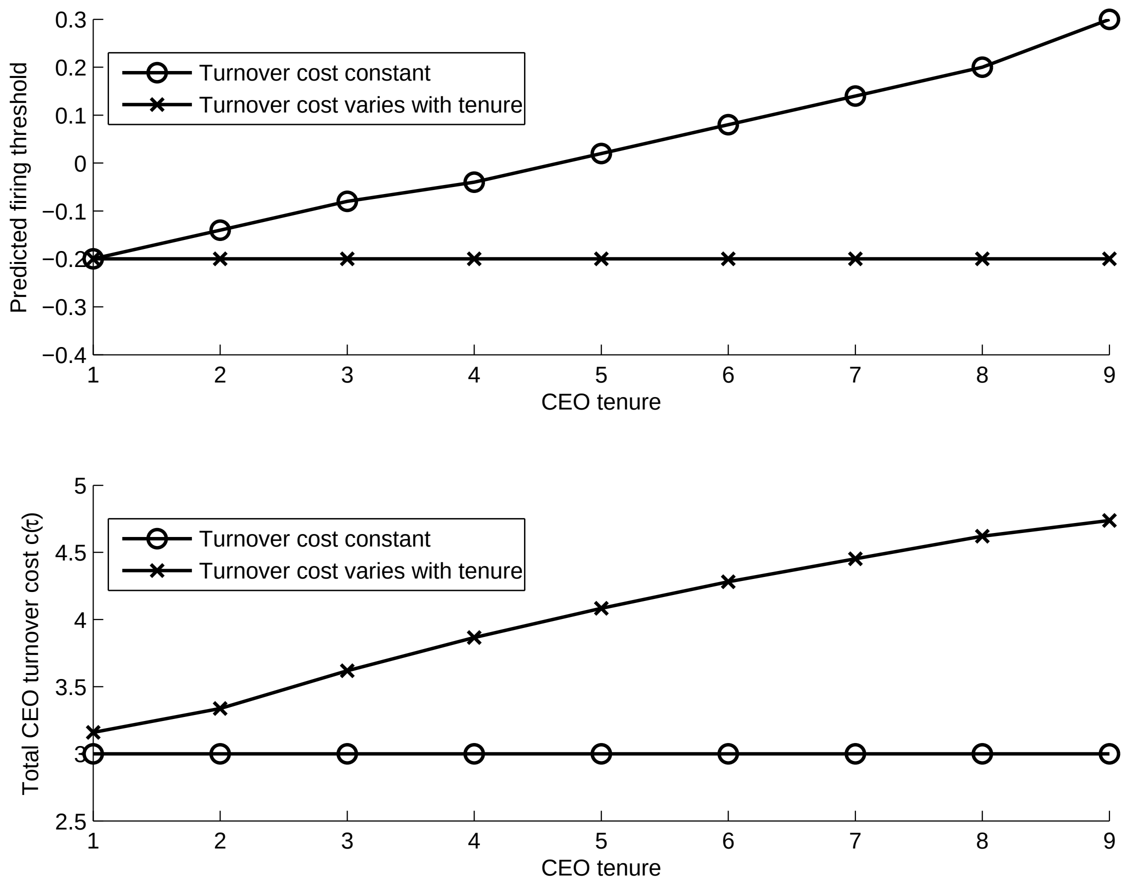

Figure IA.1. Firing threshold with turnover costs that vary with tenure. The top panel shows two predicted firing thresholds as a function of CEO tenure. The line with X's shows the firing threshold produced by constant CEO turnover costs; parameter values are $\beta = 0.9$, $\mu_0 = 1\%$, $\sigma_0 = 2\%$, $\sigma_\epsilon = 3\%$, $c = 3\%$, $\phi = 0.12$, and $\sigma_z = 7\%$, and voluntary turnover occurs after (and only after) completing 10 periods in office. The line with O's shows the predicted firing threshold using these 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.