Internet Appendix for “CEO wage dynamics: Estimates from a learning model”

Lucian A. Taylor

July 19, 2012

Contents

1. Model Solution

1.1. Learning and first steps

First I solve the learning problem, which is a Kalman filtering problem. Since prior beliefs and signals are normally distributed, Bayes' rule tells us that agents' posterior beliefs about CEO ability will also be normally distributed. At the end of year t, agents' beliefs are distributed as

$$\eta_i \sim N\left(m_{it}, \sigma^2_{\tau_{it}}\right), \tag{IA.1}$$

where τ_{it} as the number of years completed by CEO of firm i as of the end of year t. For simplicity I drop the subscripts on τ. Agents update their beliefs about CEO ability by observing the mean-zero surprises in profitability and the additional signal:

$$\widetilde{Y}_{it} = Y_{it} - a_i - v_t - m_{t-1} = \eta_i + \varepsilon_{it} - m_{it-1} \tag{IA.2}$$

$$\widetilde{z}_{it} = z_{it} - m_{it-1}. \tag{IA.3}$$

Applying Bayes' rule, the posterior variance follows

$$\sigma^2_\tau = \sigma^2_0\left(1 + \tau\left(\frac{\sigma^2_0}{\sigma^2_\varepsilon} + \frac{\sigma^2_0}{\sigma^2_z}\right)\right)^{-1}, \tag{IA.4}$$

which goes to zero in the limit where tenure τ becomes infinite. The posterior mean belief m_{it} follows a martingale:

$$m_{it} = m_{it-1} + \frac{\sigma^2_\tau}{\sigma^2_\varepsilon}\widetilde{Y}_{it} + \frac{\sigma^2_\tau}{\sigma^2_z}\widetilde{z}_{it}. \tag{IA.5}$$

Next I solve for the changes in expected pay. From assumption 6, we have

$$\Delta E_{t-1}\left[w_{it}\right] = \theta_t B_{it}\left(m_{it-1} - m_{it-2}\right). \tag{IA.6}$$

Substituting in Eq. (IA.5) yields

$$\Delta E_{t-1}\left[w_{it}\right] = \theta_t B_{it}\left(\frac{\sigma^2_{\tau-1}}{\sigma^2_\varepsilon}\widetilde{Y}_{it-1} + \frac{\sigma^2_{\tau-1}}{\sigma^2_z}\widetilde{z}_{it-1}\right). \tag{IA.7}$$

This equation relates changes in expected compensation to the previous year's earnings surprise $\widetilde{Y}_{it-1}$ and additional signal surprise $\widetilde{z}_{it-1}$.

The dividend at the end of year t equals profits minus CEO pay

$$D_{it} = B_{it}Y_{it} - w_{it}, \tag{IA.8}$$

and the firm's value at the beginning of year t equals

$$M_{it} = E_{t-1}\left[\sum_{s=0}^{\infty}\beta^{s+1}D_{it+s}\right]. \tag{IA.9}$$

From this equation I derive the firm's stock return, the average industry return (which equals a constant plus v_t), and the stock return in excess of the industry.

1.2. Excess returns

Prediction 1 (excess returns): The excess stock return (firm minus industry) in year t equals

$$r_{it} \approx \frac{B_{it}}{M_{it}}\widetilde{Y}_{it} + \frac{B_{it}}{M_{it}}\left(1-\theta_{t+1}\right)\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)\left(m_{it}-m_{it-1}\right) - median\left(r_{it}\right). \tag{IA.10}$$

This equation uses the approximation that the pay-performance sensitivity b_{it} is much less than the firm's market value, which I confirm empirically.

Proof:

First I show that the industry shock to profitability, v_t, is observable. I adjust profitability and average across the N_i firms k in firm $i's$ industry:

$$\lim_{N_i\to\infty}\frac{1}{N_i}\sum_{k=1}^{N_i}\left(Y_{kt}-a_k-m_{kt-1}\right) = v_t + \lim_{N_i\to\infty}\frac{1}{N_i}\sum_{k=1}^{N_i}\left(\eta_k - m_{kt-1}+\varepsilon_{kt}\right) \tag{IA.11}$$

$$= v_t. \tag{IA.12}$$

The model assumes agents know or can observe all quantities on the left-hand side, so it follows that they can also observe the right-hand side, which converges to the industry shock v_t since $\eta_k - m_{kt-1} + \varepsilon_{kt}$ has mean zero.

In the remainder of this section I drop the firm subscript i, for convenience. Also, since assets B_{it} are constant over time, I denote them B.

The unexpected stock return is

$$R_t - E_{t-1}\left[R_t\right] = M_t^{-1}\left(D_t - E_{t-1}\left[D_t\right] + M_{t+1} - E_{t-1}\left[M_{t+1}\right]\right) \tag{IA.13}$$

The unexpected dividend is

$$D_t - E_{t-1}\left[D_t\right] = B\left(\eta - m_{t-1} + v_t + \varepsilon_t\right) - \left(w_t - E_{t-1}\left[w_t\right]\right) \tag{IA.14}$$

$$= B\left(\widetilde{Y}_t + v_t\right) - b_t r_t, \tag{IA.15}$$

since (as I show later) the expected excess return r_t equals zero, implying $w_t - E_{t-1}[w_t] = b_t r_t$. Recalling from the learning results that

$$m_t = m_{t-1} + \frac{\sigma_\tau^2}{\sigma_\varepsilon^2}\widetilde{Y}_t + \frac{\sigma_\tau^2}{\sigma_z^2}\widetilde{z}_t \tag{IA.16}$$

we have

$$D_t - E_{t-1}[D_t] = B\left(\frac{\sigma_\varepsilon^2}{\sigma_\tau^2}(m_t - m_{t-1}) - \frac{\sigma_\varepsilon^2}{\sigma_z^2}\widetilde{z}_t + v_t\right) - b_t r_t. \tag{IA.17}$$

The surprise in future market value is

$$\begin{aligned} M_{t+1} - E_{t-1}[M_{t+1}] &= E_t - E_{t-1}\left[\sum_{s=0}^{\infty}\beta^{s+1}D_{t+1+s}\right] && \text{(IA.18)} \\ &= E_t - E_{t-1}\left[\sum_{s=0}^{\infty}\beta^{s+1}(B\eta - w_{it+1+s})\right], && \text{(IA.19)} \end{aligned}$$

because firm fixed effect a_i is known (hence no change in its expected value), and shocks ε and v have conditional mean zero. Combining the results above yields

$$\begin{aligned} R_t - E_{t-1}[R_t] &= M_t^{-1}\left(B\left(\widetilde{Y}_t + v_t\right) - b_t r_t\right) && \text{(IA.20)} \\ &\quad + M_t^{-1}\left(E_t - E_{t-1}\left[\sum_{s=1}^{\infty}\beta^s(B\eta - w_{it+s})\right]\right). && \text{(IA.21)} \end{aligned}$$

The CEO's last period is T, so there are $T - \tau_{t+1}$ periods left at the beginning of period $t+1$. In periods $T+1$ and later, a new CEO is in office. Before period T, agents learn nothing about this new CEO's ability or his expected pay. Therefore we have

$$M_{t+1} - E_{t-1}[M_{t+1}] = E_t - E_{t-1}\left[\sum_{s=0}^{T-\tau_t-1}\beta^{s+1}(B\eta - w_{t+1+s})\right]. \tag{IA.22}$$

Decomposing into the two pieces and using fact that firm size and η are constant over time,

$$\begin{aligned} M_{t+1} - E_{t-1}[M_{t+1}] &= B(E_t - E_{t-1}[\eta])\sum_{s=0}^{T-\tau-1}\beta^{s+1} - \sum_{s=0}^{T-\tau-1}\beta^{s+1}(E_t - E_{t-1}[w_{t+1+s}]) \\ &= B(m_{it} - m_{it-1})\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right) - \sum_{s=0}^{T-\tau-1}\beta^{s+1}(E_t - E_{t-1}[w_{it+1+s}]). \end{aligned}$$

Starting with $s = 0$, we want to know

$$E_t - E_{t-1}[w_{t+1}] = E_t[w_{t+1}] - E_{t-1}[E_t[w_{t+1}]]. \tag{IA.23}$$

Recall that θ_{t+1} is known at the beginning of period $t+1$ but not at the beginning of t. We therefore need to treat θ_{t+1} as a random variable at time t. It is possible to show that

$$\begin{aligned} E_t - E_{t-1}[w_{t+1}] &= \theta_{t+1} B (m_t - m_{t-1}) + \\ & BE_{t-1}[m_t - m_{t-1} | m_t - m_{t-1} > 0] \left(\frac{\theta^{down} - \theta^{up}}{2} \right). \end{aligned} \tag{IA.24}$$

Using results for the truncated normal distribution, and denoting $\phi(0)$ the pdf of the standard normal distribution evaluated at zero, we have

$$E_t - E_{t-1}[w_{t+1}] = \theta_{t+1} B (m_t - m_{t-1}) + B\kappa(\tau). \tag{IA.25}$$

$$\kappa(\tau) \equiv (Var_{t-1}(m_{ijt} - m_{ijt-1}))^{1/2} \phi(0) \left(\theta^{down} - \theta^{up}\right).$$

Using backwards induction, it follows that

$$E_t - E_{t-1}[w_{t+1+s}] = E_t - E_{t-1}[w_{t+1}].$$

Plugging this result in, we have

$$\begin{aligned} M_{t+1} - E_{t-1}[M_{t+1}] &= B(m_t - m_{t-1}) \beta \left(\frac{1-\beta^{T-\tau}}{1-\beta} \right) - \\ & (E_t - E_{t-1}[w_{t+1}]) \sum_{s=0}^{T-\tau-1} \beta^{s+1} \\ &= \beta \left(\frac{1-\beta^{T-\tau}}{1-\beta} \right) B \left[(1-\theta_{t+1})(m_t - m_{t-1}) - \kappa(\tau) \right], \end{aligned}$$

and the firm's stock return is

$$\begin{aligned} R_t &= E[R_t] + \frac{B}{M_t} v_t - \frac{b_t}{M_t} r_t \\ & + \frac{B}{M_t}(m_t - m_{t-1}) \left[\frac{\sigma_\varepsilon^2}{\sigma_\tau^2} + \beta \left(\frac{1-\beta^{T-\tau}}{1-\beta} \right)(1-\theta_{t+1}) \right] \\ & + \frac{B}{M_t} \left(-\frac{\sigma_\varepsilon^2}{\sigma_z^2}(z_t - m_{t-1}) - \beta \left(\frac{1-\beta^{T-\tau}}{1-\beta} \right) \kappa(\tau) \right) \end{aligned} \tag{IA.26}$$

It is possible to show that the average of excess returns r_t across industry firms goes to zero in the limit as the number of industry firms becomes infinite. Since all firms in the industry have the same assumed expected return $E[R]$, then the average realized industry return $\overline{R}_t$ equals

$$\overline{R}_t = E[R_t] + \overline{\left(\frac{B}{M_t} \right) v_t}$$

and the return in excess of the industry return equals[1]

$$r_t \approx \frac{B}{M_t}\left[\frac{\sigma_\varepsilon^2}{\sigma_\tau^2}+\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)(1-\theta_{t+1})\right](m_t-m_{t-1}) \qquad \text{(IA.27)}$$
$$-\frac{B}{M_t}\left(\frac{\sigma_\varepsilon^2}{\sigma_z^2}(\widetilde{z}_t+\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)\kappa(\tau)\right).$$

We can also write the excess return as

$$r_t \approx \frac{B}{M_t}\left(\left[\frac{\sigma_\varepsilon^2}{\sigma_\tau^2}+\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)(1-\theta_{t+1})\right]\frac{\sigma_\tau^2}{\sigma_\varepsilon^2}\widetilde{Y}_t\right) \qquad \text{(IA.28)}$$
$$+\frac{B}{M_{it}}\left(\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)(1-\theta_{t+1})\frac{\sigma_\tau^2}{\sigma_z^2}\widetilde{z}_t-\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)\kappa(\tau)\right)$$
$$= \frac{B}{M_t}\left(\left[1+\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)\frac{\sigma_\tau^2}{\sigma_\varepsilon^2}(1-\theta_{t+1})\right]\widetilde{Y}_t\right) \qquad \text{(IA.29)}$$
$$+\frac{B}{M_t}\left(\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)(1-\theta_{t+1})\frac{\sigma_\tau^2}{\sigma_z^2}\widetilde{z}_t-\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)\kappa(\tau)\right) \qquad \text{(IA.30)}$$
$$= \frac{B}{M_t}\left(\widetilde{Y}_t+\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)(1-\theta_{t+1})(m_t-m_{t-1})-\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)\kappa(\tau)\right) \qquad \text{(IA.31)}$$

The various forms of this equation will be useful in various places later in the Appendix.

While $\widetilde{Y}_t$ and $\widetilde{z}_t$ are normally distributed with mean zero, the excess return r_t is not normally distributed, because θ_{t+1} is a binary discrete random variable perfectly correlated with the sign of $\widetilde{Y}_t\frac{\sigma_\tau^2}{\sigma_\varepsilon^2}+\widetilde{z}_t\frac{\sigma_\tau^2}{\sigma_z^2}$. The expected excess return is zero, by construction. The median of $\widetilde{Y}_t\frac{\sigma_\tau^2}{\sigma_\varepsilon^2}+\widetilde{z}_t\frac{\sigma_\tau^2}{\sigma_z^2}$. is zero, and so is the median of $\theta_{t+1}\left(\widetilde{Y}_t\frac{\sigma_\tau^2}{\sigma_\varepsilon^2}+\widetilde{z}_t\frac{\sigma_\tau^2}{\sigma_z^2}\right)$. The median return is therefore

$$median\left(r_{it}|\tau,T,\frac{B}{M_t};\beta,\sigma_\varepsilon,\sigma_0,\sigma_z,\theta^{down},\theta^{up}\right)=-\frac{B}{M_t}\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)\kappa(\tau), \qquad \text{(IA.32)}$$

which has the same sign as $-\left(\theta^{down}-\theta^{up}\right)$. Substituting this expression into the equations above yields Prediction 1. End of proof.

1.3. Return volatility

Prediction 2 (excess return volatility):

[1]Here we use the approximation that the firm's time-t book to market ratio, B/M_t, approximately equals the industry average ratio, $\overline{(B/M_t)}$. We also use the approximation $b_{ijt} << M_{it}$.

1. In the special case with no learning, i.e., $\sigma_0^2 = 0$, or in the limit when tenure goes to infinity, then the variance of excess stock returns equals

$$var_{t-1}(r_{it}) = \left(\frac{B_{it}}{M_{it}}\right)^2 \sigma_\varepsilon^2. \tag{IA.33}$$

2. In the special case in which $\theta^{up} = \theta^{down} = 1$, meaning the CEO captures the entire surplus from, then the variance equals

$$var_{t-1}(r_{it}) = \left(\frac{B_{it}}{M_{it}}\right)^2 \left(\sigma_{\tau-1}^2 + \sigma_\varepsilon^2\right), \tag{IA.34}$$

   where $\sigma_{\tau-1}^2$ is the uncertainty about CEO ability at the beginning of year t, given in Eq. (IA.4).

3. If prior uncertainty $\sigma_0^2 > 0$, the CEO's share $\theta^{up} = \theta^{down} = \theta$, and $0 < \theta < 1$, then the variance of excess stock returns decreases with CEO tenure, increases with prior uncertainty σ_0, and decreases with the CEO's share of the surplus θ.

Proof: The proof uses the following result on asymmetric random normal variables. If X is distributed as $N(0, \sigma)$ and

$$\begin{aligned} Y &= \theta_+ X \text{ if } X \geq 0 && \text{(IA.35)} \\ &= \theta_- X \text{ if } X < 0, && \text{(IA.36)} \end{aligned}$$

then the variance of Y equals

$$Var(Y) = \sigma^2 \left[\frac{\theta_+^2 + \theta_-^2}{2} - \phi(0)^2 (\theta_+ - \theta_-)^2\right]. \tag{IA.37}$$

I use Eq. (IA.31) to compute return volatility. Random variable θ_{t+1} depends on the sign of $(m_t - m_{t-1})$. I introduce notation that will come in handy soon:

$$\begin{aligned} \theta_{+/-} &= \beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)(1-\theta_{t+1}) && \text{(IA.38)} \\ \theta_+ &= \beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)(1-\theta^{up}) && \text{(IA.39)} \\ \theta_- &= \beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)(1-\theta^{down}), && \text{(IA.40)} \end{aligned}$$

$$\lambda\left(T-\tau; \theta^{up}, \theta^{down}, \beta\right) \equiv \frac{\theta_+^2 + \theta_-^2}{2} - \phi(0)^2 (\theta_+ - \theta_-)^2 \tag{IA.41}$$

The variance of the second term in Eq. (IA.31) is therefore

$$\left(\frac{B}{M_t}\right)^2 \left(\frac{\sigma_\tau^4}{\sigma_\epsilon^4}(\sigma_{\tau-1}^2+\sigma_\epsilon^2)+\frac{\sigma_\tau^4}{\sigma_z^4}\left(\sigma_{\tau-1}^2+\sigma_z^2\right)+2\frac{\sigma_\tau^4}{\sigma_\epsilon^2\sigma_z^2}\sigma_{\tau-1}^2\right)\lambda\left(T-\tau;\theta^{up},\theta^{down},\beta\right) \quad \text{(IA.42)}$$

The variance of the first term in Eq. (IA.31) above is

$$\left(\frac{B_{it}}{M_{it}}\right)^2 \left(\sigma_{\tau-1}^2+\sigma_\epsilon^2\right).$$

The variance of returns therefore equals

$$\begin{aligned} var_t\left(r_{it}\right) &= \left(\frac{B}{M_t}\right)^2 \left(\sigma_{\tau-1}^2+\sigma_\epsilon^2\right)+ \qquad \text{(IA.43)}\\ & \left(\frac{B}{M_t}\right)^2 \left(\frac{\sigma_\tau^4}{\sigma_\epsilon^4}(\sigma_{\tau-1}^2+\sigma_\epsilon^2)\right)\lambda\left(T-\tau;\theta^{up},\theta^{down},\beta\right)+\\ & \left(\frac{B}{M_t}\right)^2 \left(\frac{\sigma_\tau^4}{\sigma_z^4}\left(\sigma_{\tau-1}^2+\sigma_z^2\right)+2\frac{\sigma_\tau^4}{\sigma_\epsilon^2\sigma_z^2}\sigma_{\tau-1}^2\right)\lambda\left(T-\tau;\theta^{up},\theta^{down},\beta\right)\\ & 2\frac{B}{M_t}\beta\left(\frac{1-\beta^{T-\tau}}{1-\beta}\right)cov_t\left(\tilde{Y}_t,\left(1-\theta_{t+1}\right)\left(m_t-m_{t-1}\right)\right). \end{aligned}$$

It is possible to show that the covariance term above equals

$$cov_{t-1}\left(\tilde{Y}_t,\theta_{+/-}\left(Z_t\right)\right)=\frac{\theta_++\theta_-}{2}\left[\frac{\sigma_\tau^2}{\sigma_\epsilon^2}\left(\sigma_\varepsilon^2+\sigma_{\tau-1}^2\right)+\frac{\sigma_\tau^2}{\sigma_z^2}\sigma_{\tau-1}^2\right],$$

so return volatility equals

$$\begin{aligned} var_{t-1}\left(r_{it}\right) &= \left(\frac{B}{M_t}\right)^2 \left(\sigma_{\tau-1}^2+\sigma_\epsilon^2\right)+ \qquad \text{(IA.44)}\\ & 2\left(\frac{B}{M_t}\right)^2 \left(\frac{\theta_++\theta_-}{2}\left[\frac{\sigma_\tau^2}{\sigma_\epsilon^2}\left(\sigma_\varepsilon^2+\sigma_{\tau-1}^2\right)+\frac{\sigma_\tau^2}{\sigma_z^2}\sigma_{\tau-1}^2\right]\right)+\\ & \left(\frac{B}{M_t}\right)^2 \left(\frac{\sigma_\tau^4}{\sigma_\epsilon^4}(\sigma_{\tau-1}^2+\sigma_\epsilon^2)+\frac{\sigma_\tau^4}{\sigma_z^4}\left(\sigma_{\tau-1}^2+\sigma_z^2\right)+2\frac{\sigma_\tau^4}{\sigma_\epsilon^2\sigma_z^2}\sigma_{\tau-1}^2\right)\lambda\left(T-\tau;\theta^{up},\theta^{down},\beta\right). \end{aligned}$$

Claim: Holding constant $T-\tau$ (number of years left in office) or setting $T-\tau$ to infinity, then we have

$$\lim_{\tau\to\infty} var_{t-1}\left(r_{it}\right)=\left(\frac{B_{it}}{M_t}\right)^2\sigma_\varepsilon^2. \quad \text{(IA.45)}$$

This results follows by noting

$$\lim_{\tau\to\infty}\sigma_{\tau-1}^2=0. \quad \text{(IA.46)}$$

Claim: If $\theta^{up} = \theta^{down} = \theta$, $\sigma_0^2 > 0$, and $0 \leq \theta \leq 1$, then $var_{it}(r_{it})$ decreases in tenure.

Proof: Posterior variance σ_τ^2 is strictly positive and decreasing in τ if $\sigma_0^2 > 0$. The quantity $\left(1 - \beta^{T-\tau}\right)$ is also decreasing in τ. The result follows by inspecting the expression for return variance, noting that quantities decreasing in τ are multiplied by weakly positive quantities. Even if $\theta = 1$, the first term in Eq. (IA.44) still decreases in τ.

Claim: If $\theta^{up} = \theta^{down} = \theta$, and $0 \leq \theta \leq 1$, $var_t(r_{it})$ is increasing in σ_0^2.

Proof: Posterior variance σ_τ^2 is increasing in σ_0^2. Terms multiplying σ_τ in the expression above for return variance are all weakly positive, so the conclusion follows.

Claim: If $\theta^{up} = \theta^{down} = \theta$, $\sigma_0 > 0$, $0 < \theta < 1$, then return variance is strictly decreasing in θ.

Proof: Inspecting the expression for return variance, the term multiplying $(1-\theta)$ is positive, so the entire term is strictly decreasing in θ.

1.4. CEO pay

Prediction 3 (CEO pay): The change in expected CEO compensation, scaled by the firm's lagged market value, equals

$$\frac{\Delta E_{t-1}[w_{it}]}{M_{it-1}} \approx \gamma r_{it-1} + \gamma \frac{B_{it}}{M_{it-1}} \left(\frac{\sigma_\varepsilon^2}{\sigma_z^2}\right) \widetilde{z}_{it-1} + g(\cdot) \tag{IA.47}$$

$$\gamma(\tau, T; \beta, \sigma_\varepsilon, \sigma_0, \theta_t) = \frac{\sigma_{\tau-1}^2 \theta_t}{\sigma_\varepsilon^2 + \sigma_{\tau-1}^2 \beta \left(\frac{1-\beta^{T-\tau+1}}{1-\beta}\right)(1-\theta_t)}, \tag{IA.48}$$

where g is a deterministic function given in the Appendix.

Proof:

Using the last model assumption, we have

$$\frac{\Delta E_{t-1}[w_t]}{M_{t-1}} = \theta_t \frac{B}{M_{t-1}} (m_{t-1} - m_{t-2})$$

Rearranging Eq. (IA.27) yields

$$(m_{t-1} - m_{t-2}) \approx \frac{r_{it-1} \frac{M_{t-1}}{B} + \beta \left(\frac{1-\beta^{T-\tau+1}}{1-\beta}\right) \kappa(\tau - 1) + \frac{\sigma_\varepsilon^2}{\sigma_z^2} \widetilde{z}_{t-1}}{\frac{\sigma_\varepsilon^2}{\sigma_{\tau-1}^2} + \beta \left(\frac{1-\beta^{T-\tau+1}}{1-\beta}\right)(1-\theta_t)},$$

so

$$\begin{aligned}
\Delta E_{t-1}[w_t] &\approx \theta_t B \frac{r_{t-1}\frac{M_{it-1}}{B_{it-1}} + \beta\left(\frac{1-\beta^{T-\tau+1}}{1-\beta}\right)\kappa(\tau-1) + \frac{\sigma_\varepsilon^2}{\sigma_z^2}\widetilde{z}_{ijt-1}}{\frac{\sigma_\varepsilon^2}{\sigma_{\tau-1}^2} + \beta\left(\frac{1-\beta^{T-\tau+1}}{1-\beta}\right)(1-\theta_t)} \\
\frac{\Delta E_{t-1}[w_t]}{M_{t-1}} &\approx \gamma r_{t-1} + \gamma\frac{B}{M_{t-1}}\left(\frac{\sigma_\varepsilon^2}{\sigma_z^2}\right)\widetilde{z}_{t-1} + g(\cdot) \\
\gamma(\tau, T; \beta, \sigma_\varepsilon, \sigma_0, \theta_t) &= \frac{\sigma_{\tau-1}^2\theta_t}{\sigma_\varepsilon^2 + \sigma_{\tau-1}^2\beta\left(\frac{1-\beta^{T-\tau+1}}{1-\beta}\right)(1-\theta_t)} \\
g\left(\tau, T; \beta, \sigma_\varepsilon, \sigma_0, \theta_t, \frac{B}{M_{t-1}}\right) &= \frac{B}{M_{t-1}}\kappa(\tau-1)\beta\left(\frac{1-\beta^{T-\tau+1}}{1-\beta}\right)\gamma(.)
\end{aligned}$$

Comparative statics for γ : If there is no learning, i.e., $\sigma_0 = 0$, then $\gamma = 0$ since $\sigma_{\tau-1}^2 = 0$. Also, in limit where τ goes to infinity then we have $\sigma_{\tau-1}^2 = 0$ and hence $\gamma = 0$. By inspection, for $\sigma_0 > 0$, slope γ is increasing in θ_t, independent of firm size M_{t-1} or B, decreasing in signal noise σ_ε^2, increasing in initial uncertainty σ_0^2, and independent of the additional signal's precision $1/\sigma_z$. It is straightforward to show that γ is decreasing in tenure.

2. Estimating the value of vesting options and restricted stock

This Appendix explains how I estimate $vovest_{jt}$, the value of CEO j's options that vest during year t, and $vsvest_t$, the value of a CEO shares that vest during year t. The value of options vesting equals the number of options vesting ($novest_t$) times the price of each option vesting ($pvest_t$):

$$vovest_t = novest_t \times povest_t. \tag{IA.49}$$

A similar formula applies to shares vesting:

$$vsvest_t = nsvest_t \times psvest_t. \tag{IA.50}$$

The number of options vesting during the year is

$$novest_t = opt\_unex\_exer\_num_t \frac{ajex_t}{ajex_{t-1}} - opt\_unex\_exer\_num_{t-1} + opt\_exer\_num_t. \tag{IA.51}$$

$opt\_unex\_exer\_num_t$ is Execucomp's number of unexercised exercisable options held by the CEO at the end of fiscal year t. The ratio $ajex_t/ajex_{t-1}$ (also Execucomp variables) adjusts for stock splits during year t. $opt\_exer\_num_t$ is Execucomp's number of shares obtained upon exercising options during year t. The explanation for the formula above is that the

CEO starts year t with a supply of options $opt\_unex\_exer\_num_{t-1}$ that have already vested but have not yet been exercised. An amount $novest_{jt}$ of new options vests, then the CEO gets rid of some of these options by exercising them ($opt\_exer\_num_t$), so the CEO is left with a supply $opt\_unex\_exer\_num_t$ of vested but unexercised options at the end of year t. The formula assumes that options are exercised before any stock splits occur. I set $novest$ equal to zero for fewer than 5% of observations that are negative.

The number of shares vesting during the year is given by

$$nsvest_t = stock\_unvest\_num_{t-1} - stock\_unvest\_num_t \frac{ajex_t}{ajex_{t-1}} + new\_granted\_num_t \quad \text{(IA.52)}$$

$stock\_unvest\_num_t$ is Execucomp's number of shares of restricted stock held by the executive that had not yet vested by the end of year t. $new\_granted\_num_t$ is the number of new shares of restricted stock granted during the year, which I estimate by dividing the dollar value of newly granted options (Execucomp variable $rstkgrnt_t$ before 2006, $stock\_awards\_fv_t$ in 2006 and later) by $\overline{S}_t$, the midpoint of the starting and ending share price for the year. To understand the formula for $nsvest_t$, the CEO starts with a supply of unvested shares at the beginning of the year ($stock\_unvest\_num_{t-1}$), then he or she receives some new shares ($new\_granted\_num_t$), then $nsvest_t$ shares vest, so the CEO is left with a supply $stock\_unvest\_num_t$ of unvested shares at the end of the year. I set $nsvest_t$ to zero if it takes a negative value. Since I do not know the exact date when the shares vest, I assume they vest at a share price $psvest_t$ midway between the starting and ending price for the year.

I estimate the price of the vesting options using the Black-Scholes formula, adjusted for dividends. I estimate the strike price K_{t-1} for vesting options using the method of Core and Guay (2002), as described in Edmans, Gabaix, and Landier (2009):

$$K_{t-1} = S_t - \frac{opt\_unex\_exer\_est\_val_{t-1}}{opt\_unex\_exer\_num_{t-1}}. \quad \text{(IA.53)}$$

$opt\_unex\_exer\_est\_val_{t-1}$ is the Execucomp estimated value of unexercised exercisable options at the end of fiscal year $t-1$. The dividend rate is Execucomp variable bs_yield measured at end of fiscal year t, divided by 100. I impute a zero if this variable is missing. I also winzorize this variable at the 95th percentile each year. Black-Scholes volatility is given by Execucomp variable $bs\_volatility$ at end of fiscal year t. If this variable is missing, I replace it with the year's median value. I winzorize volatility at the 5th and 95th percentile each year. The risk free rate is the continuously compounded risk-free rate, derived from the one-month Treasury rate in July of year t. Following the method of Core and Guay (2002) and Edmans, Gabaix, and Landier (2008), I set the average maturity of maturing options equal to the maturity of options granted during year t (computed using Execucomp option maturity date, *exdate*), minus four years. If there were no new grants in year t then I set $T_t = 5.5$ years. In the case of multiple new grants during year t, I take the longest maturity option. If maturity becomes negative then I set maturity equal to 1 day.

3. Forecasting Final CEO Tenure T_j

This section explains how I forecast T_j (the total number of years CEO j spends in office) for CEOs who have not left office by the end of the sample period. Forecasted T_j equals the CEO's tenure in his last record in Execucomp plus the forecasted number of years left in office, denoted $YearsLeft_{jt}$. The forecast is based on the following regression:

$$\log\left(1 + YearsLeft_{jt}\right) = \log a_0 + b_1 \log Age_{jt} + b_2 \log Tenure_{jt} + \varepsilon_{jt}. \tag{IA.54}$$

Age_{jt} is CEO j's age in year t (Execucomp variable AGE). I estimate the regression by taking CEOs whose last year in office is in the database, and then creating one regression observation for each year the CEO spent in office, potentially including years before 1992. The regression uses 14,111 observations and has an R^2 value of 0.23. Forecasted T_j is then

$$\begin{aligned} \widehat{T}_j &= Tenure_{jt^*} + \widehat{a}_0 Age_{jt^*}^{\widehat{b}_1} Tenure_{jt^*}^{\widehat{b}_2} - 1 && \text{(IA.55)} \\ &= Tenure_{jt^*} + e^{12.5} Age_{jt^*}^{-2.75} Tenure_{jt^*}^{0.114} - 1, && \text{(IA.56)} \end{aligned}$$

where t^* denotes CEO j's last year in the database. $\widehat{T}_j$ is missing if Age_{jt} or $Tenure_{jt}$ is missing.

4. Details on Cleaning the Data

I clean the data as follows. First I fill in missing CEO indicators in Execucomp. I label an individual to be CEO in a firm/year observation if (i) Execucomp lists no one as CEO in the given firm/year, and (ii) either (a) this individual was CEO of the firm in previous and following year; (b) this individual was CEO in previous year, and we don't know who was CEO in following year; or (c) this person was CEO in following year, and we don't know who was CEO in previous year. I assume the CEO's first fiscal year is the one when he completes at least 6 full months in office. I use Execucomp variable BECAMECEO as the date the CEO started in office. I exclude observations where BECAMECEO is missing. Next I exclude all observations for those CEOs whose start date (BECAMECEO) is more than one year after their first yearly record as CEO in Execucomp; I assume these are data mistakes in Execucomp. Next I exclude firm/year observations where the CEO's first fiscal year in office is less than 6 months long; I keep these CEOs' later years in office. I cannot compute the vesting compensation measure in the CEO's first year in Execucomp, because computing the value of shares and options vesting in year t requires Execucomp data from year $t-1$. Therefore, I cannot compute the change in this pay measure in a CEO's first two years in Execucomp. In years when change in pay is missing for mechanical reasons, I keep the years' stock return observation but treat the change in pay variable as missing. For other years, I delete firm/year records where change in pay is missing. I exclude firm/years where I cannot observe or forecast the CEO's total tenure T_j. Next I exclude firm/years where I cannot find the firm's lagged market cap in CRSP, and then I eliminate firm/years in which the variance of excess returns is missing.

5. Model extension: Endogenous CEO turnover

This robustness section extends the model to allow endogenous CEO turnover and then estimates the extended model. The extended model is similar to Taylor (2010). All the assumptions are the same as in the original model, but now we assume the board chooses whether or not to fire the CEO at the beginning of each year. Firing the CEO costs the firm a fraction c of its assets. The board's goal is to maximize firm value.

The firm's dividend is now

$$D_{it} = Y_{it} - \mathbf{1}\left(fire_{it}\right) B_i c - w_{it} \tag{IA.57}$$

where $\mathbf{1}\left(fire_{it}\right)$ is an indicator equal to 1 if the firm fires the CEO at the end of year t. The board makes CEO firing decisions that maximize firm value.

The board's optimization problem is

$$\max_{\{fire_{it+s}\}_{s=0}^{\infty}} M_t, \tag{IA.58}$$

where M_t is the firm's market value at the beginning of year t, before the firing decision has been made.We therefore have

$$\begin{aligned} M_t^* &= \max_{\{fire_{it+s}\}_{s=0}^{\infty}} E_{t-1}\left[\sum_{s=0}^{\infty} \beta^{s+1} D_{it+s}\right] \\ &= \max_{\{fire_{it+s}\}_{s=0}^{\infty}} E_{t-1}\left[\sum_{s=0}^{\infty} \beta^{s+1}\left(B_i Y_{it+s} - \mathbf{1}\left(fire_{it+s}\right) B_i c - w_{it+s}\right)\right] \\ &= \max_{\{fire_{it+s}\}_{s=0}^{\infty}} E_{t-1}\left[\sum_{s=0}^{\infty} \beta^{s+1}\left(B_i\left(a_i + \eta_i\right) - \mathbf{1}\left(fire_{it+s}\right) B_i c - w_{it+s}\right)\right], \end{aligned}$$

since shocks v_t and ε_{it} have mean zero. We therefore have

$$\begin{aligned}\frac{M_t^*}{B_t} &= a_i\frac{\beta}{1-\beta}+\max_{\{fire_{it+s}\}_{s=0}^{\infty}} E_{t-1}\left[\sum_{s=0}^{\infty}\beta^{s+1}\left(\eta_i-\mathbf{1}\left(fire_{it+s}\right)c-\frac{w_{it+s}}{B_i}\right)\right]\\ &= a_i\frac{\beta}{1-\beta}+\max_{\{fire_{it+s}\}_{s=0}^{\infty}} E_{t}\left[\sum_{s=0}^{\infty}\beta^{s+1}\left(m_{it-1+s}-\mathbf{1}\left(fire_{it+s}\right)c-\frac{E_{t+s}\left[w_{it+s}\right]}{B_i}\right)\right]\end{aligned}$$

I define Δ_t as the difference between posterior and the prior belief:

$$m_{it}=m_{i0}+\Delta_{it}. \tag{IA.59}$$

Also, I denote initial expected pay when the CEO enters office $E\left[w_0\right]$. Substituting in yields

$$\frac{M_t^*}{B_t} = \left(a_i+m_{i0}-E\left[w_{i0}\right]\right)\frac{\beta}{1-\beta}+ \tag{IA.60}$$

$$\max_{\{fire_{it+s}\}_{s=0}^{\infty}} E_{t}\left[\sum_{s=0}^{\infty}\beta^{s+1}\left(\Delta_{it-1+s}-\mathbf{1}\left(fire_{it+s}\right)c-\frac{E_{t+s}\left[w_{it+s}\right]-E\left[w_0\right]}{B_i}\right)\right]$$

$$\frac{M_t^*}{B_t} = \left(a_i+m_{i0}-E\left[w_{i0}\right]\right)\frac{\beta}{1-\beta}+V_t^* \tag{IA.61}$$

where V_t^* is the value function:

$$V_t^*=\beta\max_{fire_{it}}\left[\Delta_{it-1}-\frac{E_{t-1}\left[w_{it}\right]-E\left[w_0\right]}{B_i}-\mathbf{1}\left(fire_{it}\right)c\right]+\beta E_{t-1}\left[V_{t+1}^*\right]. \tag{IA.62}$$

The state variable is

$$x_{it}\equiv\Delta_{it-1}-\frac{E_{t-1}\left[w_{it}\right]-E\left[w_0\right]}{B_i}, \tag{IA.63}$$

which equals zero when the CEO first starts in office. State variable x equals the cumulative CEO surplus captured by the shareholders, plus a constant. We know the dynamics for Δ_{it} and $E_t\left[w_{it}\right]$:

$$\Delta_{it} = \Delta_{it-1}+\left(m_{it}-m_{it-1}\right) \tag{IA.64}$$

$$\frac{E_t\left[w_{it+1}\right]}{B_i} = \frac{E_{t-1}\left[w_{it-1}\right]}{B_i}+\theta_{t+1}\left(m_{it}-m_{it-1}\right) \tag{IA.65}$$

so

$$x_{it+1}=x_{it}+\left(m_{it}-m_{it-1}\right)\left(1-\theta_{t+1}\right), \tag{IA.66}$$

where

$$m_{it}-m_{it-1}\sim N\left(0,\sigma_{\tau-1}^2\sigma_\tau^2\left(\frac{1}{\sigma_\varepsilon^2}+\frac{1}{\sigma_z^2}\right)\right) \tag{IA.67}$$

by Bayes' Rule. Since the dynamics of x depend on tenure τ, τ is also a state variable. I therefore write V^* as a function: $V^*(x,\tau)$. If the board fires its CEO, it hires a new one so that the value function resets to

$$V_{fire} = V(x=0,\tau=1) - c. \tag{IA.68}$$

If the CEO retires voluntarily, then the value function resets to

$$V_{ret} = V(x=0,\tau=1). \tag{IA.69}$$

CEOs voluntarily retire after tenure year τ with probability $p_{ret}(\tau)$, estimated using data on voluntary successions as in Taylor (2010). If the firm chooses not to fire its CEO, then

$$\begin{aligned} V_{keep} &= \beta x_{it} + \beta E_{t-1}\left[V^*_{t+1}\right] && \text{(IA.70)} \\ &= \beta x_{it} + \beta\left(p_{ret}V_{ret} + (1-p_{ret})\,E_{t-1}\left[V\left(x_{it+1},\tau+1\right)\right]\right). && \text{(IA.71)} \end{aligned}$$

The firm chooses whether to fire the CEO at time t according to

$$\begin{aligned} V(x_{it},\tau_t)^* &= \max\left\{V_{fire}, V_{keep}\right\} && \text{(IA.72)} \\ &= \max\{V(0,1) - c, && \text{(IA.73)} \\ &\qquad \beta x_{it} + \beta\left(p_{ret}V_{ret} + (1-p_{ret})\,E_t\left[V\left(x_{it+1},\tau+1\right)\right]\right)\}. && \text{(IA.74)} \end{aligned}$$

Collecting results, the firm's market value equals

$$\begin{aligned} \frac{M^*_{it}}{B_i} &= k + V^*(x_{it},\tau_{it}) \\ k &\equiv (a_i + m_{i0} - E[w_{i0}])\frac{\beta}{1-\beta} \\ V^*(x_{it},\tau_{it}) &= \max\{V(0,1) - c, \\ &\qquad \beta x_{it} + \beta\left(p_{ret}(\tau)V(0,1) + (1-p_{ret}(\tau))\,E_{t-1}\left[V\left(x_{it+1},\tau+1\right)\right]\right)\} \\ V_{ret} &= V^*(0,1) \\ x_{it+1} &= x_{it} + (m_{it} - m_{it-1})(1-\theta_{t+1}) \\ m_{it} - m_{it-1} &\sim N\left(0, \sigma^2_{\tau-1}\sigma^2_\tau\left(\frac{1}{\sigma^2_\varepsilon} + \frac{1}{\sigma^2_z}\right)\right) \\ \theta_{t+1} &= \theta^{up} \text{ if } m_{it} - m_{it-1} > 0, \text{ otherwise } \theta_{t+1} = \theta^{down}. \end{aligned}$$

In simulations I choose k so that the simulated market-to-book ratio equals its empirical counterpart.

I simulate returns using the following equations:

$$
\begin{aligned}
R_{it} &= \frac{D_{it}+M_{it+1}}{M_{it}}-1 && \text{(IA.75)}\\
&= \frac{Y_{it}-E_t\left[w_{ijt}\right]/B-br_{it}/B+M_{it+1}/B}{M_{it}/B}-1 && \text{(IA.76)}\\
&\approx \frac{Y_{it}-E_t\left[w_{ijt}\right]/B+k+V(x_{t+1},\tau+1)}{k+V\left(x_t,\tau\right)}-1. && \text{(IA.77)}
\end{aligned}
$$

Unexpected returns equal

$$
\begin{aligned}
r_{it} &= R_{it}-E_t\left[R_{it}\right] && \text{(IA.78)}\\
&= \frac{\widetilde{Y}_{it}+V(x_{t+1},\tau+1)-E\left[V(x_{t+1},\tau+1)|x_t,\tau\right]}{k+V\left(x_t,\tau\right)}. && \text{(IA.79)}
\end{aligned}
$$

6. Model extension: Learning about firm quality

I make the following changes in notation. For convenience I drop subscripts on several variables. $\widehat{a}_{t|s}$ and $\widehat{\eta}_{|s}$ denotes the posterior mean of a_{it} and η_i, respectively, at the end of period s. Therefore, $\widehat{\eta}_{|s}=m_{is}$ from the original notation. $\Sigma_{a_t|s}$ and $\Sigma_{\eta|s}$ are the posterior variance of beliefs about a_{it} and η_i, respectively, at the end of period s. I drop firm subscripts i for convenience.

I write the problem in vector form to apply the multivariate version of Bayes' rule. State variable $x_t \equiv \left[\begin{array}{cc} a_t & \eta \end{array}\right]'$ follows (as long as CEO stays in office)

$$
\begin{aligned}
x_t &= \Phi x_{t-1}+(I-\Phi)\begin{pmatrix}\overline{a}\\0\end{pmatrix}+\begin{pmatrix}u_t\\0\end{pmatrix} && \text{(IA.80)}\\
\Phi &= \begin{pmatrix}\rho & 0\\0 & 1\end{pmatrix}. && \text{(IA.81)}
\end{aligned}
$$

Beliefs about x_t at the end of period $t-1$ are distributed as $N\left(\mu_{t|t-1},\Omega_{t|t-1}\right)$, and beliefs about x_t at end of period t are distributed as $N\left(\mu_{t|t},\Omega_{t|t}\right)$. From the law of motion for x we can immediately write

$$
\begin{aligned}
\mu_{t|t-1} &= \Phi\mu_{t-1|t-1}+(I-\Phi)\begin{pmatrix}\overline{a}\\0\end{pmatrix} && \text{(IA.82)}\\
\Omega_{t|t-1} &= \Phi'\Omega_{t-1|t-1}\Phi+\begin{pmatrix}\sigma_u^2 & 0\\0 & 0\end{pmatrix}. && \text{(IA.83)}
\end{aligned}
$$

When a new CEO takes office at the beginning of period t we set the off-diagonal elements of $\Omega_{t|t-1}$ to zero and the diagonal element corresponding to η to σ_0^2, meaning that uncertainty about the CEO resets to the prior uncertainty while uncertainty about firm quality keeps its current value. The signal observed each period is

$$X_t \equiv \begin{pmatrix} Y_t - v_t \\ z_t \\ \omega_t \end{pmatrix} = \Theta x_t + \varepsilon_t \tag{IA.84}$$

$$\Theta = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \varepsilon_t = \begin{pmatrix} \varepsilon_t \\ \delta_{zt} \\ \delta_{\omega t} \end{pmatrix} \sim N(0, \Sigma) \tag{IA.85}$$

$$\Sigma = \begin{pmatrix} \sigma_\varepsilon^2 & 0 & 0 \\ 0 & \sigma_z^2 & 0 \\ 0 & 0 & \sigma_\omega^2 \end{pmatrix}. \tag{IA.86}$$

Bayes rule states that

$$\mu_{t|t} = \Omega_{t|t}\left(\Omega_{t|t-1}^{-1}\mu_{t|t-1} + \Theta'\Sigma^{-1}X_t\right). \tag{IA.87}$$

$$\Omega_{t|t} = \left[\Omega_{t|t-1}^{-1} + \Theta'\Sigma^{-1}\Theta\right]^{-1}. \tag{IA.88}$$

Next I provide an expression for excess stock returns. In the original model, excess returns are given in Eq. (IA.27). The only difference in this model is that the firm's market value moves due to changes in beliefs about a_t, firm quality. The contribution of firm quality to market value at the beginning of period t is

$$B\beta \sum_{s=0}^{\infty} \beta^s E_{t-1}\left[a_{t+s}\right]. \tag{IA.89}$$

It is straightforward to show that

$$E_{t-1}\left[a_{t+s}\right] = \rho^{s+1}\widehat{a}_{t-1|t-1} + \overline{a}\left(1-\rho^{s+1}\right), \tag{IA.90}$$

hence

$$B\beta \sum_{s=0}^{\infty} \beta^s E_{t-1}\left[a_{t+s}\right] = B\beta \sum_{s=0}^{\infty} \beta^s \left(\rho^{s+1}\widehat{a}_{t-1|t-1} + \overline{a}\left(1-\rho^{s+1}\right)\right) \tag{IA.91}$$

$$= B\beta \sum_{s=0}^{\infty} \beta^s \left(\rho^{s+1}\left(\widehat{a}_{t-1|t-1} - \overline{a}\right) + \overline{a}\right) \tag{IA.92}$$

$$B\beta \sum_{s=0}^{\infty} \beta^s E_{t-1}\left[a_{t+s}\right] = B\beta\left[\frac{\overline{a}}{1-\beta} + \rho\frac{\widehat{a}_{t-1|t-1} - \overline{a}}{1-\beta\rho}\right]. \tag{IA.93}$$

The change in this contribution from the end of period t to the end of period $t-1$ is

$$B\beta\rho\frac{\widehat{a}_{t|t} - \widehat{a}_{t-1|t-1}}{1-\beta\rho}. \tag{IA.94}$$

The unexpected change in the contribution is

$$B\beta\rho\frac{\widehat{a}_{t|t}-E_{t-1}\left[\widehat{a}_{t|t}\right]}{1-\beta\rho}, \tag{IA.95}$$

where E_t denotes expectations conditioning on information known at the beginning of period t. Using equations (**??**)-(IA.88) one can show that the unexpected change in contribution to market value is

$$B\beta\rho\frac{\widehat{a}_{t|t}-\rho\widehat{a}_{t-1|t-1}-(1-\rho)\,\overline{a}}{1-\beta\rho}. \tag{IA.96}$$

Therefore, the excess stock return in year t is given by Eq. (**??**) plus the following term:

$$\frac{B_{it}}{M_{it}}\beta\rho\frac{\widehat{a}_{t|t}-\rho\widehat{a}_{t-1|t-1}-(1-\rho)\,\overline{a}}{1-\beta\rho} \tag{IA.97}$$

that comes from learning about firm quality.

I obtain predicted moments by first simulating values of state variable x_t, then simulating values of the signals X_t, updating beliefs according to the equations above, computing excess returns, and then taking the variance of simulated returns. I begin simulations with the variance of a_t at its long-run value.

7. Model extension: Persistent earnings shocks

This Appendix proves the following claim: Given the definitions of x, π, and Y in robustness Section ?, the extended model's predictions are identical to those in the main model.

The firm's unexpected return is

$$R_t - E_{t-1}\left[R_t\right] = M_t^{-1}\left(D_t - E_{t-1}\left[D_t\right] + M_{t+1} - E_{t-1}\left[M_{t+1}\right]\right) \tag{IA.98}$$

The unexpected dividend equals

$$\begin{aligned} D_t - E_{t-1}\left[D_t\right] &= B\left(\pi_{it}-E_{t-1}\left[\pi_{it}\right]\right)-\left(w_t-E_{t-1}\left[w_t\right]\right)\\ &= B\left(\sum_{s=-\infty}^{t} x_{is\to t} - E_{t-1}\left[\sum_{s=-\infty}^{t} x_{is\to t}\right]\right) - b_t r_t. \end{aligned}$$

Since contributions made at $s<t$ are known at time t, we have

$$D_t - E_{t-1}\left[D_t\right] = B\left(x_{it\to t} - E_{t-1}\left[x_{it\to t}\right]\right) - b_t r_t.$$

The surprise in future market value is

$$M_{t+1} - E_{t-1}[M_{t+1}] = E_t - E_{t-1}\left[\sum_{s=0}^{\infty} \beta^{s+1} D_{t+1+s}\right] \tag{IA.99}$$

$$= E_t - E_{t-1}\left[\sum_{s=0}^{\infty} \beta^{s+1}\left(B\pi_{it+1+s} - w_{it+1+s}\right)\right] \tag{IA.100}$$

$$= E_t - E_{t-1}\left[\sum_{s=0}^{\infty} \beta^{s+1}\left(B\sum_{\tau=-\infty}^{t+1+s} x_{i\tau\to t+1+s} - w_{it+1+s}\right)\right] \tag{IA.101}$$

Since beliefs about past contributions $x_{is\to t}$, $s < t$, do not change during period t, we have

$$M_{t+1} - E_{t-1}[M_{t+1}] = E_t - E_{t-1}\left[\sum_{s=0}^{\infty} \beta^{s+1}\left(B\sum_{\tau=t}^{t+1+s} x_{i\tau\to t+1+s} - w_{it+1+s}\right)\right]$$

Since

$$\begin{aligned}\sum_{s=0}^{\infty} \beta^{s+1}\left(\sum_{\tau=t}^{t+1+s} x_{i\tau\to t+1+s}\right) &= -x_{it\to t} + \sum_{s=0}^{\infty}\beta^{s}\sum_{\tau=0}^{\infty}\beta^{\tau} x_{it+s\to it+s+\tau} \\ &= -x_{it\to t} + \sum_{s=0}^{\infty}\beta^{s} Y_{it+s} = -x_{it\to t} + Y_{it} + \sum_{s=1}^{\infty}\beta^{s} Y_{it+s}\end{aligned}$$

we have

$$\begin{aligned}M_{t+1} - E_{t-1}[M_{t+1}] &= E_t - E_{t-1}\left[B\left(-x_{it\to t} + Y_{it} + \sum_{s=1}^{\infty}\beta^{s} Y_{it+s}\right) - \sum_{s=1}^{\infty}\beta^{s} w_{it+s}\right] \\ &= -B\left(x_{it\to t} - E_{t-1}[x_{it\to t}]\right) + Y_{it} - E_{t-1}[Y_{it}] \\ &\quad + E_t - E_{t-1}\left[\sum_{s=1}^{\infty}\beta^{s}\left(BY_{it+s} - w_{it+s}\right)\right]\end{aligned}$$

so

$$\begin{aligned}R_{t-1} - E_{t-1}[R_t] &= M_t^{-1}\left(B\left(Y_{it} - E_{t-1}[Y_{it}]\right) - b_t r_t\right) \\ &\quad + M_t^{-1}\left(E_t - E_{t-1}\left[\sum_{s=1}^{\infty}\beta^{s}\left(BY_{it+s} - w_{it+s}\right)\right]\right).\end{aligned}$$

Substituting in Eq. (1), the assumption that the firm fixed effect is known, and the industry shock is i.i.d. with mean zero, we have

$$\begin{aligned}R_t - E_{t-1}[R_t] &= M_t^{-1}\left(B\left(\widetilde{Y}_t + v_t\right) - b_t r_t\right) \\ &\quad + M_t^{-1}\left(E_t - E_{t-1}\left[\sum_{s=1}^{\infty}\beta^{s}\left(B\eta - w_{it+s}\right)\right]\right).\end{aligned}$$

This equation is identical to Eq. (IA.20) in Appendix 1.2, so the predictions about excess returns will be identical as in the main model. Since the equation for Y_{it} has not changed, the equations for learning dynamics will not change, and neither will the equations for wage dynamics.

8. CEO Tenure, Return Volatility, and the Variance of Profitability

Figure 2 in the main paper shows that excess stock return volatility declines after a new CEO takes office. The model attributes this decline to learning about CEO ability. In this section I test an alternate explanation, which is that earnings volatility declines with CEO tenure. First I estimate the shocks to profitability, then I check whether the volatility of these shocks changes with CEO tenure. For comparison, I confirm that return volatility declines with tenure even after including additional controls.

I compute annual return on assets (ROA) for every firm/year in the sample. I estimate earnings shocks ε_{it} using the following panel model:

$$ROA_{it} = \beta_0 + \beta_1 ROA_{it-1} + \beta_2 \log\left(Assets_{it-1}\right) + \beta_i + \beta_t + \beta_\tau + \varepsilon_{it}, \tag{IA.102}$$

where β_i is a firm fixed effect, β_t is a year fixed effect, and β_τ is a CEO tenure fixed effect for tenure categories $\tau = 1, ..., 10+$ years. The conditional mean of the squared residuals, $E\left[\varepsilon_{it}^2|\text{regressors}\right]$, equals the conditional variance of profitability. I estimate this conditional variance from the following regression:

$$\widehat{\varepsilon}_{it}^2 = \gamma_0 + \gamma_1 \log\left(Assets_{it-1}\right) + \gamma_i + \gamma_t + \gamma_\tau + u_{it}, \tag{IA.103}$$

where $\widehat{\varepsilon}_{it}^2$ is estimated from regression (IA.102), γ_i is a firm fixed effect, γ_t is a year fixed effect, and γ_τ is a CEO tenure fixed effect.

Table 1 shows the estimated tenure fixed effects γ_τ for the conditional variance of ROA. The fixed effect for tenure = 10+ years is normalized to zero. None of the tenure fixed effects is signficantly different from zero. The conditional variance of profitability shows no significant pattern with CEO tenure.

For comparison, I measure tenure fixed effects in excess return volatility. I regress RETVAR (the annualized variance of excess stock returns) on log lag assets, firm fixed effects, year fixed effects, tenure fixed effects, and (in one specification) the squared shocks to ROA ($\widehat{\varepsilon}_{it}^2$). Results are in Table 1. The fixed effect for tenure equal one (two) years is significantly positive at the one (ten) percent confidence level, and the remaining fixed effects are indistinguishable from zero, consistent with the result in Figure 3. In sum, return volatility declines significantly with tenure, but earnings volatility does not.

Table 1: **CEO Tenure, Return Volatility, and the Variance of Profitability**
This table shows the variance of firm profitability and excess stock returns, conditional on CEO tenure and other controls. The variance for CEOs with tenure = 10+ is normalized to zero. First I estimate shocks to return on assets (ROA) by regressing ROA on its lag, log(lag assets), firm fixed effects, year fixed effects, and CEO tenure fixed effects (results not shown). I then square the estimated residuals and regress these on log lag assets, firm fixed effects, year fixed effects, and tenure fixed effects; estimates are below. The table also shows the tenure fixed effects from a regression of RETVAR (annualized variance of excess stock returns) on log lag assets, firm fixed effects, year fixed effects, tenure fixed effects, and the squared shock to ROA. The sample contains is described in section IV.B. Standard errors are in parentheses.

	Dependent variable		
CEO tenure (yrs)	Squared ROA shock	Variance of excess returns	Variance of excess returns
1	0.0009	0.0264	0.0263
	(0.0031)	(0.0033)	(0.0033)
2	0.0025	0.0060	0.0050
	(0.0032)	(0.0034)	(0.0034)
3	0.0009	0.0034	0.0033
	(0.0033)	(0.0035)	(0.0035)
4	-0.0012	0.0027	0.0029
	(0.0034)	(0.0036)	(0.0036)
5	0.0001	-0.0001	-0.0004
	(0.0035)	(0.0037)	(0.0037)
6	-0.0015	0.0038	0.0038
	(0.0036)	(0.0038)	(0.0038)
7	-0.0024	-0.0003	-0.0007
	(0.0038)	(0.0040)	(0.0040)
8	-0.0027	-0.0017	-0.0014
	(0.0039)	(0.0041)	(0.0041)
9	-0.0028	-0.0025	-0.0022
	(0.0040)	(0.0043)	(0.0043)
10+	0	0	0
	N/A	N/A	N/A
log(lag assets)	0.0044	-0.0189	-0.0183
	(0.0018)	(0.0019)	(0.0019)
Squared ROA shock			0.0584
			(0.0079)
Year fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
N	20,400	20,482	20,400
R-squared	0.183	0.612	0.614