

What do Boards Know?

The Value of CEO Selection and Monitoring*

Andrew Capron

University of North Carolina at Chapel Hill

Barton H. Hamilton

Olin Business School, Washington University in St. Louis

Andrés Hincapié

Federal Reserve Bank of Atlanta

Noah Lyman

Warwick Business School, University of Warwick

October 17, 2026

Abstract

How much do boards know about their CEOs, and what is that information worth? We estimate a dynamic model of CEO succession in which boards screen internal and external candidates and monitor the incumbent's evolving quality. Boards' information is valuable but limited: screening resolves about sixty percent of the uncertainty about internal candidates but only six percent about external ones. Relative to uninformed hiring and firing, this information raises firm value by 10.6 and 7.7 percent, respectively. Full information would add another 23 percent. Insiders' performance advantage at appointment reflects what boards know, not better candidate pools.

*Capron: Department of Economics, University of North Carolina at Chapel Hill (email: acapron@unc.edu); Hamilton: Olin Business School, Washington University in St. Louis (email: hamiltonb@wustl.edu); Hincapié: Federal Reserve Bank of Atlanta (email: andres.hincapie@atl.frb.org); Lyman: Warwick Business School, University of Warwick (email: Noah.Lyman@wbs.ac.uk). We thank participants at seminars at the University of North Carolina at Chapel Hill and at the Southern Economic Association for their valuable feedback and comments, and Simone Hanna for her expert research assistance.

1 Introduction

When to replace the firm’s chief executive and who to replace him with are two of the most important decisions boards make, but they have received uneven attention in the literature. Beginning with Weisbach (1988), the sensitivity of termination to firm performance has served as a standard measure of whether a board is doing its job, and various studies have refined this measure: how it has changed over time (Huson et al., 2001; Kaplan and Minton, 2012), whether boards attribute industry shocks to the CEO (Jenter and Kanaan, 2015), and how much sensitivity is obscured by the classification of departures (Jenter and Lewellen, 2020). There is no comparable statistic for the appointment decision. This asymmetry in our understanding of termination and appointment is exacerbated by the fact that the impact these decisions have in firm performance crucially depends on what boards know about the quality of their chief executive (the extent of *monitoring*), and what they know about the quality of the internal and external prospects who can replace him (the extent of *selection*). And we do not have a clear understanding of what boards know at either decision margin. In this paper we address this asymmetry by providing a quantitative measure of what boards know at termination and appointment, and weighing the relative importance of monitoring and selection.

Weighing what boards know at both margins has proven elusive. At appointment, what a board learns about a candidate before choosing is never observed. Cziraki and Jenter (2021) document that the market for chief executives is dominated by internal promotion without evidence that insiders are intrinsically superior, and conjecture that boards simply know more about the internal candidates they employ. He and Schroth (2024) ask directly why external successions are rare, and Cvijanović et al. (2023) study the planning through which boards prepare for successions. However, the existing literature does not quantify how much boards know about the candidates they appoint. At termination, the existing measure (the sensitivity of termination to firm performance) is the product of two components: how much a year of performance moves the board’s assessment of its chief executive, and how readily the board acts on a given assessment.¹ Hence, a board that rarely dismisses a poor performer may be a board that tolerates one, or a board that cannot yet tell.

This paper separates what boards know from their willingness to act on it at both margins: at *selection*, what a board can tell about a candidate before appointing them, and at *monitoring*, what it learns about the executive it already has. We specify and estimate a dynamic model of chief executive retention and succession in which the precision of the board’s

¹Engel et al. (2003) show that the turnover-performance relation varies with the informativeness of the performance measure, and Bushman et al. (2010) document a related dependence on noise.

information is parsimoniously captured by a small number of parameters. A board employs one chief executive whose productivity is persistent but evolving, and which the board never observes directly. Each period the board decides whether to retain its CEO after learning about him from profitability, which confounds the executive’s contribution with industry conditions, and from a noisy benchmark of those conditions, so that the degree to which industry luck is credited to the chief executive is itself an object to be estimated. Upon turnover, the board chooses between an insider, who inherits part of the predecessor’s trajectory, and an outsider drawn from the market, and observes a signal about each prospect before appointing, with a precision that may differ between internal and external candidates. Boards need not act on firm value alone; we capture non-monetary motivations for board behavior by allowing two non-pecuniary costs for dismissal and external appointment. We estimate the model by Simulated Method of Moments on a panel of 6,307 chief executive spells from 1995 to 2019.

We find that boards do have information about candidates and incumbents, and that this information is limited in different ways at the two decision margins. Before an appointment, boards resolve roughly sixty percent of the uncertainty about an internal candidate’s productivity and about six percent of the uncertainty about an external one; the confidence interval for the internal signal rejects both extremes: the signal is neither uninformative nor fully revealing. In comparison, learning about the sitting chief executive is slow because a year of firm performance is a fairly noisy signal of his contribution. In terms of precision, the pre-appointment signal about an internal candidate is worth roughly nine years of CEO performance signals, and the signal about an external candidate only a few months, consistent with the gradual resolution of post-hire uncertainty documented by Taylor (2013) and Pan et al. (2015). When considering industry trends, boards do a reasonable filtering job: they filter roughly three-quarters of industry performance out of their assessments, placing them between the polar benchmarks of absolute and relative performance evaluation. Boards decisions are also affected by non-monetary motivations that create material wedges between firm value maximization and board utility maximization: the non-pecuniary cost of dismissal amounts to about ninety percent of a year of average profitability, and the cost of external appointment to about half a year.

Our central results price the board’s information at each decision margin. At appointment, replacing the board’s optimal informed rule with random appointment, holding fixed the empirical share of internal appointments, reduces firm value by 9.6%. At dismissal, replacing the board’s optimal informed rule with random dismissal, holding fixed the empirical rate of dismissal, reduces firm value by 7.2%. Replacing both optimal rules costs the firm 19.3% of value. Boards’ optimal use of what they know is therefore worth more when they

select a prospect as chief executive than when they monitor an incumbent CEO. However, only the dismissal decision has a standard measure of appropriate board responses to information. These information prices reflect frictions as well as knowledge: a board willing to fire freely, disregarding the dismissal non-pecuniary wedge, would get more than twice as much out of what it learns from monitoring.

Given the substantial value of information, we next ask how far boards stand from being fully informed. Endowing boards with perfect knowledge of chief executive productivity would raise firm value by 22.9%. Strikingly, either margin can deliver most of that gain on its own: perfect selection alone is worth 21.7% and perfect monitoring alone 14.8%. The two are therefore close substitutes, because a board that knows its candidates has little left to learn about the executive it installed, and a board that can identify a failing executive at once has less to gain from screening prospects beforehand. Overall, selection is the more valuable margin: boards' existing information is worth more at appointment, and so is the value they could still gain if they were fully informed at this margin.

Industry attribution, by contrast, matters little for firm value. While boards do respond to industry conditions in their dismissal decisions, in our data as elsewhere (Kaplan and Minton, 2012; Jenter and Kanaan, 2015), the impact on firm value is minimal.² We find that firms' responses to industry conditions are consistent with the rational use of an imperfect benchmark rather than a failure to benchmark, and the value at stake is small either way: perfect filtering of industry conditions would add 0.05 percentage points to firm value and abandoning the filtering boards already do would cost 0.26. Contrasted against the 21.7% gain in firm value from perfect selection, these magnitudes are minuscule.

Addressing the longstanding question of whether insiders make better chief executives (Shen and Cannella, 2002; Zhang and Rajagopalan, 2010; Quigley et al., 2019), our estimates reveal that the observed advantage of insiders is not intrinsic, and it arises at appointment from boards' information about internal talent. We find that the internal and external candidate pools are nearly identical in underlying quality: the average productivity of internal candidates is slightly below that of external ones, by 0.5% of assets, against a cross-sectional dispersion of executive productivity of 5.1%. The observed advantage of insiders arises from what boards know about internal prospects and from the successful regimes those candidates inherit. Removing the board's pre-appointment signal about internal candidates eliminates the observed gap in CEO performance; substituting the firm's optimal appointment rule with a coin toss (which ignores information) does the same. The insider advantage is therefore an information advantage rather than an endowment of the internal pipeline. This inter-

²This response contradicts the informativeness principle of Holmström (1979), a tension that has generated an active literature including Gibbons and Murphy (1990) and Bertrand and Mullainathan (2001).

pretation has theoretical precedent: Ferreira and Nikolowa (2023) show that when more is known about an employer’s own managers than about outside candidates, promoted insiders outperform external hires even though both are drawn from a common talent distribution. Our contribution to this debate is the quantification of what is known about internal talent, and the finding that for chief executives the informational channel accounts for the observed gap in CEO performance essentially on its own.

Our paper contributes to three more literatures. The first is the structural analysis of chief executive turnover. Taylor (2010) measures what boards’ reluctance to dismiss a chief executive costs shareholders; in his framework a departing executive is replaced by a random draw from the population, so there is no appointment decision to price. We measure the value of boards’ ability to make informed choices at the appointment margin, and find it larger than the value of their information at the dismissal margin. Taylor (2013) and Pan et al. (2015) infer the pace of board learning from compensation dynamics and from the decline in return volatility over tenure; we infer it from the personnel decisions themselves, and recover in addition the share of uncertainty resolved before the appointment, for both internal and external candidates. A related structural literature estimates models of executive careers in which the object of interest is the compensation contract (Margiotta and Miller, 2000; Gayle and Miller, 2009; Gayle et al., 2015). Our board writes no contract; what we recover is the precision of its information rather than the incentives it provides.

The second is the empirical literature on boards. Analyzing the minutes of board meetings, Schwartz-Ziv and Weisbach (2013) find that directors devote most of their attention to supervising management, and that the appointment or dismissal of an executive is the issue on which boards most often take an active hand. By separating what boards know from how they act on it, we show that the turnover-performance sensitivity understates board effectiveness to the extent that dismissal is limited by information rather than by will. For the other half of the board’s personnel problem, the appointment decision, we estimate how much boards know about internal and external candidates and what that knowledge is worth to shareholders.

Finally, we also contribute to the literature on internal labour markets and employer learning. That firms know more about the workers they already employ than about outside applicants is a long-standing idea (Greenwald, 1979; Lazear, 1998; DeVaro and Morita, 2013), and the informational advantage boards hold over internal candidates is the accumulated product of having observed those candidates in junior roles (Baker et al., 1994). Ferreira and Nikolowa (2023) study this idea in equilibrium: employers who have discovered talent internally retain it, because outside employers cannot assess it as well, so promotion from within prevails and appointments from outside are comparatively rare. Our estimates put

a value on that accumulated knowledge at the top of the organization: the option to select on fit revealed before the appointment, which is what distinguishes an internal promotion from an external hire (Chan, 1996; Bidwell, 2011), is among the most valuable features of the succession process. This literature has also asked how fast such knowledge accumulates (Altonji and Pierret, 2001). Employers’ initial expectation errors about a young worker fall by half within three years (Lange, 2007), whereas, according to our estimates, an error in a board’s assessment of its chief executive takes roughly nine years to fall by half, a difference consistent with how noisy firm performance is as a signal of any one executive’s contribution. This has a practical counterpart in succession planning, whose disclosure has itself been the subject of governance attention (McConnell and Qi, 2022; Cvijanović et al., 2023), and which a recent literature takes as its subject: Celentano and Mello (2024) ask why firms so often lack a succession plan at all, and Geelen and Hajda (2024) study how the anticipation of succession shapes firm policy. The precision of the board’s pre-appointment signal is, in light of our model, what such planning produces.

The remainder of the paper proceeds as follows. Section 2 describes the data and the facts that motivate and discipline the model. Section 3 presents the model. Section 4 discusses identification and estimation. Section 5 reports the estimates. Section 6 prices the board’s information through a series of counterfactuals. Section 7 concludes.

2 Data and Stylized Facts

Data from S&P 1500 firms confirms the observation in Cziraki and Jenter (2021) that internal promotion dominates the market for CEOs: 66.2% of CEO successions in our sample are internal. Moreover, our data suggest that the information boards possess about their CEO prospects leaves traces in what boards do and in the ensuing performance of the firm: rather than a persistent advantage over outsider executives, the insider CEO edge in firm performance is concentrated at appointment and fades over tenure. Consistent with Jenter and Lewellen (2020), we find that boards weigh recent performance above more distant results when dismissing a chief executive. In addition, we find that this recency phenomenon extends to the hiring margin when boards choose between internal and external candidates. In the rest of this section we describe our sample and document the features of CEO succession that motivate and discipline our model.

Sample construction. We construct a panel of publicly traded firms spanning 1995 to 2019 from four sources: ExecuComp provides CEO-level information for S&P 1500 firms, including executives’ prior experience at the firm, tenure, and turnover; Compustat provides

firm fundamentals, including assets and operating income; CRSP provides stock returns; and the CEO dismissal database of Gentry et al. (2021) allows us to classify CEO departures as either forced or voluntary.³ We employ two measures of performance. The first is operating ROA (operating income before depreciation scaled by total assets), which we employ both raw and industry-adjusted; the adjustment subtracts the industry-year mean computed over the universe of Compustat firms.⁴ The second is the market-adjusted stock return. Following Parrino (1997), we classify CEOs as *insiders* or *outsiders* using their prior tenure with the firm at the time of appointment: insiders are CEOs who are founders or who have more than one year of prior experience at the firm, and all other CEOs are classified as outsiders.

Our final sample consists of 38,213 firm-year observations covering 3,090 distinct firms and 6,307 CEO spells. Roughly two-thirds of firm-years are insider-led. The sample contains 3,410 CEO successions, of which 66.2% result in an internal appointment. The predominance of internal promotion is a stable feature of the market for CEOs: comparable insider shares have been documented by Zhang and Rajagopalan (2010), Cremers and Grinstein (2014), and Cziraki and Jenter (2021), with external hiring concentrated among smaller firms and declining with firm size.

Table 1 reports summary statistics separately for insider- and outsider-led firm-years. Three patterns are worth noting. First, insider-led firm-years display higher profitability: operating ROA averages 12.7% against 10.7% for outsider-led firm-years, and a gap of half a percentage point survives industry adjustment. Second, insider-led firms are substantially larger, with roughly twice the assets and revenue of outsider-led firms. Third, the dispersion of profitability is markedly higher under outsider CEOs with a standard deviation of industry-adjusted ROA of 13.6% against 11.3% under insider CEOs. The higher mean and lower variance of performance (as measured by ROA) under insider CEOs are consistent with long-standing conjectures that outsider appointments are draws from a riskier pool (Parrino, 1997; Hermalin, 2005; Chen and Hambrick, 2012). Notably, the insider advantage does not appear in returns, as one would expect if an anticipated difference in performance is already reflected in the price.

Stylized facts. Our main stylized facts confirm and expand regularities previously observed in the literature. We find that boards place more weight on recent performance over distant results when it comes to firing, and that this same recency phenomenon shows up in the hiring decision to appoint internally or externally. We also find that the insider

³The database codes the circumstances of each departure of an S&P 1500 chief executive from contemporaneous press reports; Appendix A provides details.

⁴Industries are GICS industry groups; Appendix A provides details on the construction of the industry benchmark.

Table 1: Summary Statistics by CEO Type

	Outsider-led		Insider-led		Difference
	Mean	SD	Mean	SD	
Operating ROA	0.107	(0.133)	0.127	(0.105)	+0.019***
Industry-adjusted operating ROA	0.064	(0.136)	0.069	(0.113)	+0.005***
Market-adjusted return	0.066	(0.535)	0.054	(0.449)	−0.012**
Assets (\$ billions)	10.4	(65.9)	21.5	(112.6)	+11.1***
Revenue (\$ billions)	3.5	(10.0)	7.4	(22.7)	+3.9***
Observations (firm-years)	13,122		25,091		

Notes: Firm-year observations, split by whether the current CEO was appointed from within the firm. Difference is the insider mean minus the outsider mean; significance is assessed with unequal-variance t -tests. Assets and revenue are in billions of nominal U.S. dollars. $*p < .1$, $**p < .05$, $***p < .01$.

performance edge appears at appointment but disappears quickly after, suggesting that the underlying mechanism driving the gap is not permanent. Throughout this section we decompose performance into a firm-specific and an industry component: the industry component is the industry-year mean and the firm-specific component is the deviation from it. Three regressors summarize a CEO’s record: *recent* performance (the one-year lag of the firm-specific component), *distant* performance (the mean of the two- and three-year lags), and the *industry* component (its one-year lag). We measure performance both with market-adjusted stock returns and with industry-adjusted ROA, and we standardize all covariates, so that coefficients measure the change in probability per standard deviation of the regressor.

Table 2 reports linear probability models of forced turnover. Two patterns stand out. First, consistent with Jenter and Kanaan (2015), termination is considerably more sensitive to recent than to distant performance: column (6) shows that a one-standard-deviation improvement in recent returns lowers the probability of termination by roughly one percentage point, a third of the baseline rate, with distant performance carrying about half that weight. Second, consistent with Jenter and Kanaan (2015), termination also responds to the industry component of performance (returns) at 0.4 percentage points per standard deviation, below the firm-specific sensitivity but away from zero. Column (3) shows that insider CEOs are 1.5 percentage points less likely to be fired than outsiders, though the gap disappears within firm in column (6); the gap reflects the kinds of firms that promote from within rather than boards holding promoted CEOs to a more lenient standard than outside hires.

Table 3 explores the recency phenomenon in the hiring margin, applying the same specification to the choice of successor. The regressions are conditional on CEO turnover and the

Table 2: Forced Turnover and the CEO's Record

	(1)	(2)	(3)	(4)	(5)	(6)
Recent performance, ROA	−0.017*** (0.003)		−0.012*** (0.003)	−0.019*** (0.004)		−0.014*** (0.004)
Distant performance, ROA	0.010*** (0.003)		0.006** (0.003)	0.007** (0.003)		0.005* (0.003)
Industry, ROA	−0.003** (0.001)		−0.003** (0.001)	−0.006 (0.005)		−0.004 (0.005)
Recent performance, returns		−0.012*** (0.001)	−0.011*** (0.001)		−0.011*** (0.002)	−0.010*** (0.002)
Distant performance, returns		−0.003** (0.002)	−0.002 (0.002)		−0.007*** (0.002)	−0.005*** (0.002)
Industry, returns		−0.002 (0.001)	−0.002 (0.001)		−0.004** (0.001)	−0.004** (0.001)
Insider	−0.014*** (0.003)	−0.015*** (0.003)	−0.015*** (0.003)	0.011 (0.008)	0.011 (0.008)	0.011 (0.008)
Firm + year FE	No	No	No	Yes	Yes	Yes
Observations				20,813		

Notes: Linear probability models of forced CEO turnover on standardized components of the CEO's performance record. Recent and distant performance are the firm-specific components of one-year-lagged and two-to-three-year-lagged performance; the industry regressor is the lagged industry-year mean. Standard errors clustered by firm in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

dependent variable indicates whether the successor was promoted from within. The probability of internal promotion rises with the *departing* CEO's performance, again more strongly with recent than with distant performance: column (6) indicates that a one-standard-deviation improvement in recent returns raises the probability of an internal appointment by roughly five percentage points. Column (3) shows that the probability of internal promotion rises with the industry component of ROA, by about five percentage points per standard deviation, though the correlation loses significance within firm in column (6). The hiring and firing margins thus respond symmetrically to the same information: a record that would have kept the incumbent also promotes her deputy. The pattern is consistent with boards treating a successful record as evidence that the incumbent regime works, and promoting from within to preserve it, while a poor record argues for the fresh draw an outsider represents. The symmetry also carries a subtler implication: whatever shapes the board's assessment of the departing CEO, including any component of performance the board fails to attribute correctly, propagates into the choice of who runs the firm next.

Given that boards appear to promote insiders after strong records, do the insiders they promote perform better? Figure 1 plots the insider performance premium by tenure, estimated from a regression of industry-adjusted ROA on CEO type interacted with tenure

Table 3: Internal Promotion and the Departing CEO's Record

	(1)	(2)	(3)	(4)	(5)	(6)
Recent performance, ROA	0.062*** (0.019)		0.025 (0.019)	0.059* (0.033)		0.019 (0.033)
Distant performance, ROA	-0.022 (0.019)		0.000 (0.019)	-0.004 (0.030)		0.009 (0.031)
Industry, ROA	0.052*** (0.011)		0.049*** (0.011)	0.039 (0.037)		0.045 (0.037)
Recent performance, returns		0.069*** (0.008)	0.065*** (0.008)		0.054*** (0.013)	0.052*** (0.013)
Distant performance, returns		0.035*** (0.009)	0.032*** (0.009)		0.033** (0.013)	0.030** (0.014)
Industry, returns		0.013 (0.008)	0.009 (0.008)		-0.004 (0.014)	-0.008 (0.014)
Firm + year FE	No	No	No	Yes	Yes	Yes
Observations			2,799			

Notes: Linear probability models of an internal appointment, estimated on succession events. The dependent variable equals one if the incoming CEO was promoted from within the firm; regressors are the standardized components of the *departing* CEO's performance record, defined as in Table 2. Standard errors clustered by firm in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

bins, with firm and year fixed effects. Newly appointed insiders outperform newly appointed outsiders by 2.1 percentage points of ROA in their first year, but the premium fades quickly, and is statistically indistinguishable from zero from the fourth year onward. The disappearance of the insider edge over tenure provides clues as to why insiders outperform. If insiders were simply better managers, the gap should be persistent. If instead their advantage derives from the circumstances of their appointment, e.g. information boards may hold about them and the successful regimes they inherit, then the gap should decay as those circumstances lose value with tenure. The profile favors the second story: the insider advantage is likely a feature of how and when insiders are chosen, rather than a persistent quality difference between candidate pools. For instance, it is reasonable to expect that firms possess more information about insider prospects, which would result on higher homogeneity across appointed insiders. Consistent with this conjecture, among CEO spells within the same firm, the between-CEO dispersion of average performance is two-thirds larger for outsider than for insider CEOs (0.0040 versus 0.0024).

Taken together, the facts presented in this section admit more than one reading. That boards promote internally after strong records, and that the insiders they promote outperform outsiders at appointment, may reflect what boards know about internal candidates before appointing them, or genuine differences between the internal and external talent pools. The prevalence of internal promotion may also reflect performance continuity, or a taste

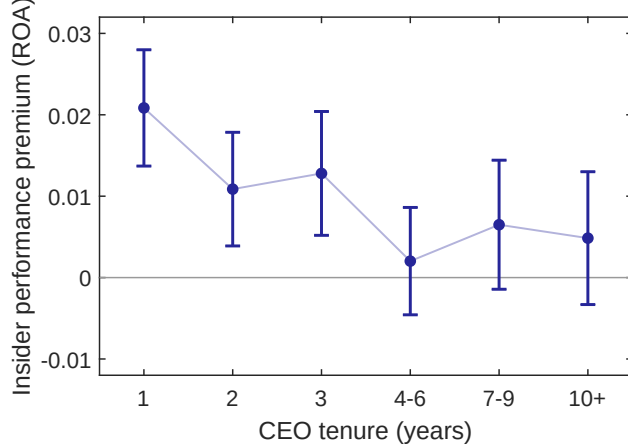


Figure 1: Insider Performance Premium over Tenure

Notes: Point estimates and 95% confidence intervals for the coefficients on insider $\times$ tenure-bin indicators from a regression of industry-adjusted ROA on CEO type interacted with tenure, with firm and year effects; standard errors are clustered by firm. The regression also includes an indicator for the years surrounding a succession; its coefficient is -0.0061 (0.0029).

for in-house management unrelated to firm value. Separating these explanations requires a model that accounts for selection and that can reproduce the patterns documented here: the weight boards place on recent and distant performance, the symmetry of these weights in the hiring and firing margins, an insider advantage concentrated at appointment and fading thereafter, and a promoted pool markedly less dispersed than one hired externally. Section 3 presents such a model.

3 A Dynamic Model of CEO Turnover

A board employs one CEO whose productivity, evolving but partially persistent, it never observes directly. Instead, it learns about the productivity of its CEO from firm profitability, which confounds the executive's contribution with industry conditions, and from a separate, noisy signal of those conditions (such as analyst commentary on industry trends). At succession the board chooses between an insider and an outsider, receiving a pre-hire signal about each; the insider partially inherits the predecessor's productivity trajectory, while the outsider's trajectory is independent of prior firm management. Additional incentives shape turnover decisions beyond information: the board faces pecuniary and non-pecuniary costs of firing and a non-pecuniary cost of external appointment, which captures insider favoritism.

At each time t , firm j in industry n employs a single CEO i . We let $\tau_{it} \in \mathbb{Z}_+$ denote the executive's tenure, $\theta_{it} \in \mathbb{R}$ denote their productivity, and $h_i \in \{I, O\}$ denote whether their appointment was internal or external. CEOs can depart voluntarily, $v_{jt} \in \{0, 1\}$, or they

can be fired, $d_{jt}^f \in \{0, 1\}$. At the beginning of each period, voluntary separation is realized. If the CEO retires ($v_{jt} = 1$) the board must hire a replacement. If the CEO does not retire ($v_{jt} = 0$) the board has the option to retain ($d_{jt}^f = 0$) or fire ($d_{jt}^f = 1$) their CEO. Upon turnover, the board chooses to appoint a new executive from the pool of internal or external talent, and it always has one candidate available from each pool.

Firm profitability. The profitability of firm j is given by:

$$Y_{jt} = \iota_j + y_{jt} - c(v_{jt} + d_{jt}^f) \quad (1)$$

where $\iota_j \sim N(\bar{\iota}, \sigma_\iota^2)$ is a permanent firm component, y_{jt} is the time-varying component of profitability, and c is the monetary cost of CEO turnover, capturing the financial cost of executive search, severance payouts, and organizational disruptions to the firm that occur during managerial transitions. The time-varying component of profitability is persistent and depends on CEO ability and an industry factor according to the following mean-reverting process:

$$y_{jt} = \rho y_{jt-1} + (1 - \rho) \theta_{it} + g_{nt} + \varepsilon_{jt} \quad (2)$$

where ρ captures persistence in firm performance, g_{nt} is the profitability factor common to all firms in industry n , and $\varepsilon_{jt} \sim N(0, \sigma_\varepsilon^2)$ is an idiosyncratic profitability shock. The industry factor is itself persistent such that:

$$g_{nt} = \phi_g g_{nt-1} + \zeta_{nt} \quad (3)$$

where ϕ_g captures persistence in industry-level profitability and where $\zeta_{nt} \sim N(0, \sigma_\zeta^2)$ are industry innovations. Through its dependence on prior profitability, current profitability y_{jt} is affected by CEO productivity and shocks from prior periods. Thus, holding the current executive's productivity constant, firms where the new CEO succeeds a high-productivity one will enjoy higher performance on average.

CEO productivity and succession. The executive's impact on firm profitability is not fixed and it evolves according to:

$$\theta_{it} = \mu_{h_i} + \phi(\theta_{it-1} - \mu_{h_i}) + \eta_{it} \quad (4)$$

where μ_h is a type-specific long-run mean capturing population differences between internal and external talent, ϕ governs the persistence of executive productivity over their tenure at the firm, and $\eta_{it} \sim N(0, \sigma_\eta^2)$ is an innovation to CEO productivity. We normalize the

outsider mean to zero, $\mu_O \equiv 0$, and leave μ_I as a free parameter capturing the effect of firm-specific human capital accumulated at lower ranks within the firm on the productivity of insider appointees (Murphy and Zabojnik, 2006; Custódio et al., 2013; Frydman, 2019). While we refer to θ_{it} as CEO productivity, under our specification θ_{it} can be understood more broadly as the quality of the firm’s leadership under CEO i : it evolves as implemented strategies succeed, adapt, and decay, rather than remaining a fixed trait of any executive. The specification in (4) spans the two modeling approaches of managerial quality in the turnover literature: as $\phi \rightarrow 1$, productivity becomes a fixed attribute that the board learns over time, as in Jovanovic (1979) and Taylor (2010), while $\phi < 1$ allows leadership quality to drift over an executive’s tenure, consistent with evidence that both persistent managerial styles (Bertrand and Schoar, 2003) and evolving executive–firm match quality (Allgood and Farrell, 2003; Pan, 2017) shape performance.

The executive’s productivity at the beginning of their tenure, in other words the initial condition of their productivity trajectory at the firm, depends on whether they are hired externally or internally. Each candidate is drawn from the stationary productivity distribution of its own type, with mean μ_k and variance $\sigma_\theta^2 = \sigma_\eta^2/(1 - \phi^2)$ implied by (4), $k \in \{I, O\}$. The internal candidate, unlike the external one, inherits part of the organizational trajectory of the predecessor. We capture this organizational continuity through the joint distribution of productivity for the two executives involved in an internal succession: the departing CEO and the internal candidate are drawn from the stationary distribution of their own type, and the two draws are jointly normal with correlation λ . Hence, conditioning on the predecessor’s productivity θ_{it} , the productivity θ_t^k of the type- k candidate is given by:⁵

$$\theta_t^k = \mu_k + \lambda_k(\theta_{it} - \mu_{h_i}) + \chi_t^k \quad (5)$$

where $\lambda_I = \lambda$ measures *continuity*, capturing the extent to which internal successors carry forward the strategy of the regime that produced them, and $\lambda_O = 0$, as the external candidate has no predecessor at the firm to inherit from. The term $\chi_t^k \sim N(0, \sigma_{\chi,k}^2)$ captures individual heterogeneity in executive ability not inherited from the predecessor, where $\sigma_{\chi,k}^2 = (1 - \lambda_k^2) \sigma_\theta^2$. We parameterize λ as a fraction of ϕ in (4), $\lambda = \kappa\phi$ with $\kappa \in [0, 1]$, so that a successor inherits at most as much predecessor productivity as the predecessor herself would carry into the next period.⁶

⁵Formally, we take $(\theta_{it}, \theta_t^I)$ to be jointly normal with $\theta_{it} \sim N(\mu_{h_i}, \sigma_\theta^2)$, $\theta_t^I \sim N(\mu_I, \sigma_\theta^2)$, and $\text{Corr}(\theta_{it}, \theta_t^I) = \lambda$. Conditioning on θ_{it} delivers a normal distribution with mean $\mu_I + \lambda(\theta_{it} - \mu_{h_i})$ and variance $(1 - \lambda^2)\sigma_\theta^2$. The external candidate is independent of the predecessor.

⁶The restriction $\lambda \leq \phi$ rules out the pathological region in which the board would prefer to fire a well-performing incumbent in order to “harvest” a superior copy of her.

Information structure. Boards observe profitability and know its permanent component ι_j in (1) and the population distributions of CEO productivity for internal and external talent. However, they do not observe separately the idiosyncratic shocks to profitability ε_{jt} or the contributions to profitability coming from its CEO’s productivity θ_{it} and from the industry factor g_{nt} . Instead, the board forms beliefs about the evolving latent quantities θ_{it} and g_{nt} using two signals. First, realizations of time-varying profitability, as governed by (2), generate noisy signals u_{jt} of CEO productivity:

$$u_{jt} \equiv y_{jt} - \rho y_{jt-1} = (1 - \rho) \theta_{it} + g_{nt} + \varepsilon_{jt} \quad (6)$$

which confound the CEO’s contribution with industry conditions and idiosyncratic shocks. Second, the board observes a noisy signal s_{nt}^g of the industry factor itself:

$$s_{nt}^g = g_{nt} + \omega_{jt}, \quad \omega_{jt} \sim N(0, \sigma_\omega^2) \quad (7)$$

reflecting the information about industry conditions available to the board from peer performance, analyst commentary, and its own reading of the environment. The precision of this signal determines how well the board can separate industry luck from CEO productivity. This information structure with latent CEO ability and industry factor is motivated by a substantial body of evidence showing that industry conditions leak into decisions which, under the informativeness principle of Holmström (1979), should be purged of them.⁷ Rather than imposing either complete or absent industry benchmarking, we let the benchmark precision σ_ω determine how much industry luck contaminates the board’s assessment of its CEO.

Given the two signals, boards use Bayes’ rule to update their beliefs. The pair (θ_{it}, g_{nt}) and the signals (u_{jt}, s_{nt}^g) form a linear-Gaussian filtering problem in two latent states, and beliefs are jointly normal throughout. We restrict attention to steady-state learning, so that the posterior covariance is constant at its stationary value and the board’s beliefs about its CEO are summarized by the posterior mean $\tilde{\theta}_{it}$. Under this restriction the board’s uncertainty about its CEO is at its stationary level at every tenure, and the updating gain characterized in Proposition 1 is constant over a spell.⁸ Proposition 1 characterizes the

⁷CEOs are more likely to be dismissed after industry downturns (Jenter and Kanaan, 2015; Kaplan and Minton, 2012) and are rewarded for sector luck in pay (Bertrand and Mullainathan, 2001).

⁸This restriction is standard in models of learning about a time-varying latent state; see Demarzo and Sannikov (2017) for a prominent example. It concerns the board’s beliefs about the executive it currently employs: its uncertainty about a *candidate* at the moment of appointment is a separate object and can differ by type, since the precision of the pre-hire signal may differ across types and inheritance further reduces it for an internal candidate. In Appendix F we relax this restriction and re-estimate the model under non-stationary learning: no parameter moves by more than two standard errors of the difference, and the fit is worse. We therefore retain stationary learning as our benchmark.

resulting belief dynamics; all of the learning mechanics used in the remainder of the paper follow from it.

Proposition 1 (Belief dynamics). *Under the information structure (6)-(7) and steady-state learning, the board's belief about its incumbent CEO evolves according to*

$$\tilde{\theta}_{it+1} = \mu_h + \phi(\tilde{\theta}_{it} - \mu_h) + \phi K_u \left[\underbrace{(u_{jt} - \hat{u}_{jt})}_{\text{performance surprise}} - \psi_g \underbrace{(s_{nt}^g - \tilde{g}_{nt})}_{\text{benchmark surprise}} \right] \quad (8)$$

where $\hat{u}_{jt} = (1 - \rho)\tilde{\theta}_{it} + \tilde{g}_{nt}$ is the predicted profitability innovation, $\tilde{g}_{nt}$ is the board's belief about the industry factor, and the updating gain K_u and the benchmark-adjustment share ψ_g are constants determined by the stationary filtering problem. Moreover, if industry persistence is lower than executive productivity persistence ($\phi_g \leq \phi$):

(i) $K_u > 0$ and $\psi_g > 0$: beliefs rise with unexpectedly good firm performance, and fall when the industry benchmark is unexpectedly favorable;

(ii) $\psi_g \rightarrow 1$ as $\sigma_\omega \rightarrow 0$ and $\psi_g \rightarrow 0$ as $\sigma_\omega \rightarrow \infty$.

PROOF: See Appendix B.

The updating rule in Proposition 1 indicates that the board evaluates its CEO on *benchmark-adjusted* performance. It nets out a fraction ψ_g of the industry surprise before crediting the CEO, and K_u governs the overall speed at which beliefs respond to news. Part (i) formalizes how the two signals jointly identify the CEO's contribution: holding performance fixed, good news about the industry is bad news about the CEO. Part (ii) pins the degree of industry-adjustment as a function of benchmark noise σ_ω . When $\psi_g = 1$, the board nets out the entire industry surprise: relative performance evaluation (RPE). When $\psi_g = 0$, industry luck is fully credited to the CEO: absolute performance evaluation (APE). Intermediate values capture partial adjustment: as the benchmark ranges from perfectly revealing to uninformative, ψ_g falls from one to zero.⁹ Given the remaining parameters, every value of $\psi_g \in (0, 1)$ maps one-to-one from benchmark noise σ_ω ; we therefore parametrize the attribution problem and directly estimate ψ_g , which is more economically interpretable than σ_ω .

⁹The interpretation of ψ_g as an informational share can be made exact in a one-shot version of the learning problem. With a single performance observation $u = (1 - \rho)\theta + g + \varepsilon$ and benchmark signal $s^g = g + \omega$, the posterior mean of θ loads on u and s^g with coefficients whose ratio implies $\psi_g = p_g/(p_g + \sigma_\omega^2)$ exactly, where p_g denotes the prior variance of the industry factor: the share of industry uncertainty that the benchmark resolves, the precise analogue of ψ_I and ψ_O in the screening problem (Equation (10)). In the fully dynamic problem, ψ_g retains this essence, with prior variances replaced by their steady-state counterparts and a more involved functional form.

Screening. Boards have prior information about both candidates before hiring, though not necessarily of equal quality. Concretely, before choosing between its two candidates, the board observes a signal about each:

$$s_t^k = \chi_t^k + \nu_t^k, \quad \nu_t^k \sim N(0, \sigma_{\nu,k}^2), \quad k \in \{I, O\} \quad (9)$$

centered on χ_t^k , the candidate's heterogeneity in productivity in (5) that is not already predictable from the predecessor. Using Bayes' rule, the signal shifts the board's belief about the type- k candidate by:

$$\xi_t^k = \mathbb{E}[\chi_t^k | s_t^k] = \psi_k s_t^k, \quad \psi_k = \frac{\sigma_{\chi,k}^2}{\sigma_{\chi,k}^2 + \sigma_{\nu,k}^2}, \quad \xi_t^k \sim N(0, \psi_k \sigma_{\chi,k}^2). \quad (10)$$

In parallel with our treatment of the industry attribution parameter ψ_g in (8), we parametrize the precision of the board's pre-hire information directly by $\psi_k \in [0, 1]$: the share of the candidate-specific uncertainty $\sigma_{\chi,k}^2$ that the board resolves before hiring, with the signal noise $\sigma_{\nu,k}$ implied. Screening operates through selection: because the board acts on ξ_t^k , a sharper signal about a type (smaller $\sigma_{\nu,k}$) raises the average productivity and reduces the dispersion of the appointees of that type, relative to its pool of candidates. The realized signal s_t^k is what the board learns about a particular candidate, while ψ_k measures how well it knows candidates of that type: the share of their candidate-specific uncertainty that its information resolves. For internal candidates, prior information accumulates through direct assessment and through the promotion decisions that produced the candidate. For external prospects, market research focusing on performance with prior employers illuminates initial outsider beliefs and ultimately hiring choices (Fee and Hadlock, 2003); for these external candidates ψ_O measures how much is revealed via market research. The gap $\psi_I - \psi_O$ therefore measures any difference in how well boards know the two types of candidates; while a positive gap seems intuitive, as in-house information is likely more abundant (e.g through experience in lower ranks within the firm), we do not impose the sign of the gap.

Because the board does not observe the productivity of the departing CEO, only having a posterior belief $\tilde{\theta}_{it}$ about it, the board's beliefs about the incoming candidates apply the inheritance mapping to the posterior mean $\tilde{\theta}_{it}$. Hence, since the board has rational expectations, its belief about the type- k candidate upon turnover is:

$$\tilde{\theta}_0^k(\xi^k) = \mu_k + \lambda_k(\tilde{\theta}_{it} - \mu_{h_i}) + \xi_t^k \quad (11)$$

which for the external candidate reduces to the signal alone, $\tilde{\theta}_0^O = \xi_t^O$, since $\mu_O = \lambda_O = 0$.

A predecessor perceived as strong therefore raises the board’s assessment of the internal bench, and the realized signals move its assessment of each candidate at hand. These entry conditions are where any information advantage about insiders would reside: the board’s belief about an insider can draw on two sources, the predecessor’s assessed trajectory and its own signal, while its belief about an outsider draws on the signal alone. A long-standing notion in personnel economics suggests boards know insiders better: firms know more about the workers they already employ than about outside applicants (Lazear, 1998; DeVaro and Morita, 2013), and boards assess executive talent by observing it in lower ranks (Hermalin and Weisbach, 2017; Ferreira and Nikolowa, 2023). However, we let our estimates determine the direction and magnitude any information advantage between internal and external prospects.

3.1 The Board’s Problem

The equity value of the firm is given by:

$$P_{jt} = \mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s b_{jt+s} Y_{jt+s} \right]. \quad (12)$$

The flow term $b_{jt}Y_{jt}$ represents firm profits, obtained by multiplying firm assets b_{jt} by profitability Y_{jt} , and $\beta \in (0, 1)$ is the discount factor. We assume firm assets are constant for simplicity, thus $b_{jt} \equiv b_j$ for all t and assets drop from boards’ optimization problem. Were the board to perfectly advance the interests of shareholders, they would choose personnel management policies to maximize (12). However, we allow for the possibility that board decision-making is swayed by their private interests. In particular, we assume the board’s flow utility is given by:

$$U_{jt} = Y_{jt} - \pi d_{jt}^f - \gamma (d_{jt}^f + v_{jt}) \mathbf{1}[h_{it} = O] \quad (13)$$

where the two non-pecuniary terms reflect the wedge separating the preferences of shareholders and the board. The board incurs a non-pecuniary cost π from firing its CEO and a non-pecuniary cost γ from hiring an external prospect.¹⁰ These features of the model allow us to assess the relative importance of non-pecuniary motives for promoting from within the firm versus profit-driven incentives. The two wedges correspond to distinct governance frictions: π reflects CEO entrenchment that depresses turnover below the shareholder optimum (Hermalin and Weisbach, 1998; Taylor, 2010), while γ nests the view that boards favor

¹⁰We assume the non-pecuniary cost of firing a CEO remain stable over time. This is consistent with evidence showing that board composition is highly persistent (Graham et al., 2020).

insiders, or handicap outsiders, for reasons unrelated to value (Agrawal et al., 2006; Mobbs and Raheja, 2012) and can also capture the average additional cost of acquiring information about external candidates.

The board chooses firing and hiring policies to maximize the expected present value of its flow utility:

$$\max \mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s U_{jt+s} \right] \quad (14)$$

which differs from shareholder value (12) only through the non-pecuniary terms: when $\pi = \gamma = 0$ the two objectives coincide, up to the constant scale b_j , and the board's policy maximizes firm value. The remainder of this section represents (14) recursively.

Timing. Events within period t unfold in five steps. First, the retirement shock is realized: the incumbent departs voluntarily with exogenous probability $\delta(\tau_{it})$, a function of tenure that we estimate in a first stage.¹¹ Second, if the incumbent does not retire, the board has the option to retain or fire them (incurring the non-pecuniary cost π). Third, if the position becomes vacant, whether by retirement or dismissal, the board draws one candidate from each talent pool, internal and external, observes the signals about each candidate (s_t^I and s_t^O), and appoints either the insider or the outsider (incurring the non-pecuniary cost γ), depending on its initial beliefs. Fourth, profitability Y_{jt} is realized under the incumbent or the newly appointed CEO, net of the turnover cost c if a succession occurred. Lastly, the board observes the profitability innovation u_{jt} and the benchmark signal s_{nt}^g and updates its beliefs according to Proposition 1, delivering the state on which next period's decisions are based. The decision to retain or fire is therefore made under beliefs that reflect performance through $t - 1$.

The value of retention and turnover. Given state $x_{jt} = (\tilde{\theta}_{it}, \tilde{g}_{nt}, y_{jt-1}, \tau_{it}, h_i)$, which includes the board's beliefs about the incumbent CEO and the industry factor, lagged performance, and the tenure and type of the incumbent CEO, the board's value conditional on retaining the CEO is given by (suppressing the firm subscript):

$$V_0(x_t; h) = \mathbb{E} \left[Y_t + \beta \left(v_{t+1} R(x_{t+1}) + (1 - v_{t+1}) V(x_{t+1}; h) \right) \right] \quad (15)$$

where v_{t+1} is the indicator of whether the current CEO will depart voluntarily at the beginning of next period. The replacement value $R(x_t)$ and the value of having the option to fire

¹¹In practice $\delta(\tau)$ is logistic in a cubic in tenure; Appendix D gives the specification, the estimates, and the implied hazard profile.

$V(x_t; h)$ are respectively given by:

$$R(x_t) = \mathbb{E}_{\xi^I, \xi^O} \left[\max \{ V_0(x_t^I(\xi^I); I), V_0(x_t^O(\xi^O); O) - \gamma \} \right] \quad (16)$$

$$V(x_t; h) = \max \{ V_0(x_t; h), -\pi + R(x_t) \} \quad (17)$$

where $x_t^I(\xi^I)$ and $x_t^O(\xi^O)$ denote the entry states of the insider and outsider candidates, with beliefs given by (11) and tenure reset to one. The replacement value $R(x_t)$ reflects the board's optimal hiring choice between internal and external talent following a turnover event, the function $V(x_t; h)$ is the value of having the option to trigger such turnover. Note that the monetary impact of forced turnover (the cost c) is implicit in the conditional value functions in (16) and (17), as they incorporate the impact of CEO departure on profits. The optimal firing and hiring decisions characterized in this section capture the effect of prior performance on current beliefs, including the extent to which industry luck is attributed to the CEO and the expectations boards have regarding future firm performance. These expectations are affected by the persistence of CEO and industry productivity, by the persistence in profitability itself, by management continuity across successions, and by what boards learn about internal and external candidates before hiring.

Finally, equity value is the expected present value (12) of shareholder cash flows evaluated under the board's policy: it incorporates the monetary cost c of succession but excludes the non-pecuniary components π and γ , which enter board utility rather than cash flow. Boards and shareholders share the same information. The performance record, the industry benchmark and the pre-appointment signals are public, so the belief that enters the price is the common one, and the difference between ψ_I and ψ_O is a difference in how much can be known about the two types of candidate, not in who knows it.¹² Abnormal returns in the model are constructed from P_t exactly as in the data, as returns net of the industry benchmark. Because the price capitalizes the board's beliefs, abnormal returns move with the *revisions* of those beliefs: news moves the price exactly insofar as it changes the board's assessment of the firm's leadership. Abnormal returns are therefore the model's natural summary of the arrival of decision-relevant information and we accordingly use returns as the performance measure for the hiring and firing margins in the estimation process described below.¹³

¹²The pre-appointment signals summarise what the public record says about a candidate: an internal candidate's performance in the senior roles and promotions that preceded the appointment, an external candidate's record at a previous employer. The same record is available to shareholders, so the beliefs that guide the board's decisions and the beliefs that price the firm coincide.

¹³This link between belief revisions and prices is the mechanism documented empirically by Pan et al. (2015), who show that uncertainty about CEO ability is capitalized into stock prices and resolves gradually over tenure, and it is consistent with the long-standing evidence that prices respond to top-management

4 Identification and Estimation

We recover parameter estimates using a two-step estimation procedure. In a first stage, we estimate the parameters of the voluntary separation process, identified off of variation in voluntary turnover rates across CEO tenure. These estimates are taken as an input in the second stage, where the remaining 16 parameters are estimated using the Simulated Method of Moments. The discount factor is calibrated throughout to $\beta = .95$.¹⁴

We first discuss how we disentangle the parameters governing firm profitability in (1): the permanent firm component $(\bar{\iota}, \sigma_{\iota})$, the idiosyncratic shock σ_{ϵ} , firm persistence ρ , and the industry factor (ϕ_g, σ_{ζ}) . The mean and the across-firm variance of profitability pin down $\bar{\iota}$ and σ_{ι} . We discipline the industry factor directly by targeting the variance and the first- and second-order autocorrelations of industry-level profitability, computed over the universe of Compustat firms within each industry. Identification of ρ and σ_{ϵ} comes from the dynamics of performance within a CEO’s spell. Performance moves within a spell both because of idiosyncratic shocks and because the quality of the firm’s leadership evolves, and the two are distinguished by how long they last: an idiosyncratic shock feeds into subsequent profitability at rate ρ and then dies out, while a movement in quality persists at rate ϕ and continues to affect performance for the remainder of the spell. We target the within-spell variance of industry-adjusted profitability together with the variances of its one-year and two-year changes. Differencing removes the permanent firm component, and comparing the change variances across the two horizons measures how quickly the transitory component decays: change variances that grow little between the one- and two-year horizons indicate rapid mean reversion, while variances that continue to grow indicate longer-lived innovations. Together with the pooled autocorrelations of profitability, these moments are primarily informative about ρ and σ_{ϵ} .

The tenure profile of performance differences across CEO types carries information about the succession channels: what boards know about each type at the time of hire, inheritance, and the type mean. To separate these channels, we target the coefficients of the following regression of profitability on tenure-binned type indicators:

$$y_{ijt} = \sum_b \delta_b^y \mathbf{1}[\tau_{ijt} \in b] \cdot \text{insider}_{ijt} + \sum_{b \geq 3} \delta_b^0 \mathbf{1}[\tau_{ijt} \in b] + \delta_0^y \mathbf{1}[\text{turnover}]_{jt} + \alpha_j + \alpha_t + \epsilon_{ijt}^y \quad (18)$$

where the tenure bins b are $\{1, 2, 3, 4-6, 7-9, 10+\}$ and α_j and α_t are firm and year effects. The tenure-bin main effects, entered for years three and later, control for the common tenure

changes (Warner et al., 1988).

¹⁴Appendix C lists the targeted moments in full and gives details on the weight matrix, the numerical routine, the standard-error formulas, and model fit.

profile of performance; the turnover indicator absorbs the first year, so the profile is measured relative to the second year, and each δ_b^y measures the insider-outsider performance difference within tenure bin b .¹⁵ The tenure profile of the coefficients δ_b^y traces the life cycle of the insider premium, and its shape carries information about distinct channels. The entry level of the premium is the difference in entering productivity between the two types and is shaped by the board’s pre-hire information about each type, inheritance, and the type mean; its long-tenure level carries information about μ_I , mediated by survival selection. The change in the premium over tenure reflects the information catch-up of whichever type entered less selected, as boards learn and replace poor performers of both types, and so disciplines the speed of learning together with the firing margin. The coefficient on the turnover indicator, capturing the dip in performance around CEO transitions, is most informative about the turnover cost c .

The dispersion of CEO quality is identified from the between-CEO variation in average within-firm performance: the differences in spell-average profitability across the executives who have led the same firm. We target the variance of these spell means across all spells at firms with at least two CEOs, net of the sampling variance that averaging a short spell adds to its mean.¹⁶ Since $\sigma_\theta^2 = \sigma_\eta^2/(1 - \phi^2)$, this moment together with the persistence ϕ determines the innovation variance σ_η . The persistence of quality ϕ , the attribution parameter ψ_g , and the two costs the board bears in personnel decisions, the firing cost π and the outsider hiring cost γ , are identified off of the hiring and firing margins. We target the coefficients of regressions of the firing and hiring decisions on the information available to the board at the time of the decision:

$$d_{ijt}^f = \delta_r^d R_{ijt}^{\text{rec}} + \delta_o^d R_{ijt}^{\text{old}} + \delta_g^d R_{jt}^{\text{ind}} + \alpha_j + \alpha_t + \varepsilon_{ijt}^d \quad (19)$$

$$\text{insider}_{i'jt}^h = \delta_r^h R_{ijt}^{\text{rec}} + \delta_o^h R_{ijt}^{\text{old}} + \delta_g^h R_{jt}^{\text{ind}} + \alpha_j + \alpha_t + \varepsilon_{ijt}^h \quad (20)$$

where R_{ijt}^{rec} and R_{ijt}^{old} are the firm’s recent and longer-lagged market-adjusted returns under the current (in (19)) or departing (in (20)) CEO and R_{jt}^{ind} is the contemporaneous industry return.¹⁷ The firing regression is estimated on all CEO-years, while the hiring regression is estimated on succession years only: i' indexes the incoming CEO, and the coefficients

¹⁵The first year of a spell is the succession year, so the turnover indicator and a first-year dummy coincide.

¹⁶Averaging n years of profitability leaves sampling variance of order σ^2/n in the spell mean. The correction subtracts each spell’s own within-spell variance divided by its length, adjusted for the part that firm-demeaning already absorbs; Appendix C gives the formula. The same estimator is applied to the data and to every simulated panel.

¹⁷The three performance measures are defined as in Section 2: recent performance is the one-year lag of the firm-specific component of returns, older performance the mean of its two- and three-year lags, and the industry component the one-year lag of the industry-year mean, all standardized.

measure how the departing CEO’s record shifts the probability that the successor is an insider. The performance coefficients discipline the mapping from information to decisions: the relative weight on recent versus longer-lagged performance reflects the speed at which boards update, which the learning parameters must reproduce. The same recency structure is what identifies the persistence of quality: the more persistent is quality, the more weight older performance carries relative to recent performance in the board’s firing decisions (Jenter and Lewellen, 2020). Returns enter the regressions decomposed into their firm-specific and industry components, so the industry coefficients measure the extent to which industry conditions enter personnel decisions holding the firm’s own performance fixed. Industry conditions reach the board’s decisions only through its beliefs about the CEO, so a board that fully filters industry performance from its assessments ($\psi_g = 1$) fires and hires without regard to the industry component of returns, while a board that credits industry performance to the CEO responds to it as it does to firm-specific news; the response of turnover to industry returns relative to firm-specific returns therefore pins down ψ_g . Conditional on the parameters governing profitability, belief formation, and pre-hire information, π is identified off of the empirical rate of forced CEO termination, and γ off of the level of the insider share.¹⁸

What remains are the parameters that govern what the board knows about a candidate at the time of appointment: the precision of its signal about internal candidates, ψ_I , the precision of its signal about external candidates, ψ_O , and continuity λ , the rate at which the predecessor’s assessed productivity passes to an internal successor. The three shape the distribution of the board’s beliefs at entry in (11): an outsider enters at the belief its signal produces, and an insider at its own signal plus the internal mean plus the inherited part of the predecessor’s assessment. A sharper signal about a type does two things to that type’s entry beliefs. It raises their mean among appointees, because the board declines candidates whose signals are poor, and it disperses them, because the board’s belief moves with the signal it observes. Our moments measure both features. The premium regression in (18) measures the difference in entry means between the two types, and the timing of dismissal within a spell, which we introduce below, measures the dispersion of entry beliefs within each type.

Consider the means first. As discussed above, the long-tenure level of the insider premium pins down μ_I , mediated by survival selection. The entry level then measures how much better appointed insiders are at entry than appointed outsiders, net of the difference in the

¹⁸Jenter and Lewellen (2020) caution that some nominally voluntary exits are in fact performance-induced. To the extent such departures remain in our voluntary margin, the forced-turnover rate understates the true incidence of belief-driven dismissal and the estimated firing margin understates the responsiveness of turnover to beliefs, making our estimates of what boards know conservative.

pools they are drawn from. Screening raises the entry productivity of appointed insiders relative to μ_I , and information about outsiders raises that of appointed outsiders relative to the population mean. Continuity raises the entry productivity of insiders through the predecessor’s trajectory rather than through the candidate’s own draw, and the hiring regression in (20) measures this channel directly: continuity ties the internal candidate’s entering productivity, and the board’s belief about it, to the departing CEO, so a larger λ steepens the response of the insider appointment probability to the predecessor’s performance, while sharper screening flattens it, because the more the board learns from the candidate’s own signal, the less its choice depends on the prior it brings from the predecessor.

The premium’s entry level is one moment for three contributions. Telling them apart comes from the dispersion of entry beliefs, which has its sharpest implication for when, within a spell, a CEO is dismissed. The board dismisses a CEO when its belief about the CEO’s productivity falls below the value of replacement, and its belief moves only with news. A board with no pre-hire information about a type appoints every CEO of that type at the same belief, at a common distance above that threshold. Crossing it requires an accumulation of bad news, so few are dismissed in their first years, and dismissal becomes likely only once the record of a poor appointment has built up. A board with pre-hire information appoints at dispersed beliefs, and its hiring rule admits a candidate at a low belief whenever the alternative is worse. A candidate appointed against a weak alternative enters close to the threshold, and a single bad year suffices for dismissal; one appointed on a strong signal starts far from it and is rarely dismissed at all. Pre-hire information about a type therefore moves that type’s dismissals toward the start of the spell, and the more precise the information, the further they move. For insiders, inheritance is a second source of dispersion: the predecessor’s assessed trajectory differs across successions and passes to the successor’s entry belief at rate λ . Information about one type also changes who of the other type is appointed. With an informative signal about the outsider, an internal candidate is sometimes appointed because the outsider under consideration was poor rather than because the insider was strong, so appointed insiders include marginal ones and their early-tenure dismissals rise; information about the insider, by making the internal alternative sometimes known to be weak, does the same for outsiders.

To capture these profiles we target the coefficients of the following regression of the dismissal indicator on tenure bins and their insider interactions, with spell and year fixed effects:

$$d_{ijt}^f = \sum_{b \geq 2} \delta_b^f \mathbf{1}[\tau_{ijt} \in b] + \sum_{b \geq 2} \delta_b^{fI} \mathbf{1}[\tau_{ijt} \in b] \cdot \text{insider}_{ijt} + \alpha_{ij} + \alpha_t + \varepsilon_{ijt}^f \quad (21)$$

with the tenure bins of (18), estimated on all CEO-years. The spell effects α_{ij} absorb every fixed characteristic of the firm and of the executive, including the executive’s type, so the coefficients measure when within a spell dismissal occurs, not how often; the overall incidence is a separate target. Since the indicator is one only in a spell’s final year, each coefficient is the frequency with which a year in its bin ends a spell, relative to the first year. The profiles rise with tenure as fewer years remain; what the moments carry is how steeply, and how differently across types. The outsider profile δ_b^f measures the dispersion of outsider entry beliefs, which ψ_O produces. The insider difference δ_b^{fI} measures the dispersion of insider entry beliefs relative to outsiders, which ψ_I and λ produce together. The profiles also carry information about the persistence of quality from a different angle than the recency structure of (19): how quickly a poor appointment is recognized depends on how durable productivity is. With λ pinned down by the hiring gradient, the insider timing profile pins down ψ_I , the outsider profile pins down ψ_O , and the entry level of the premium, each of whose components is now pinned elsewhere, adds a further restriction on the three together.

The choice margins provide additional discipline. Information about either type raises the forced-turnover rate, because the board can more often identify a successor as clearly better than the incumbent. A cost of hiring from outside raises the insider share, and the firing cost lowers the forced rate. The two levels pin down π and γ once the profiles have pinned down the information parameters, and the premium tells the two sources of a high insider share apart: a cost of hiring from outside leads the board to promote internal candidates it would otherwise pass over, which lowers their entry productivity, whereas screening raises it.

We target a total of 40 moments to recover 16 model parameters. A summary of model fit is presented in Table 4. The estimated model fits the data well: 39 of the 40 targeted moments are within two standard errors of their empirical values. In Appendix E, we formalize the mapping between moments and parameters using the sensitivity matrix outlined in Andrews et al. (2017), which confirms that each estimate is driven primarily by the moments highlighted above.

5 Structural Estimates

Despite the statistically significant difference in the performance of newly appointed insiders and outsiders shown in Figure 1, a pattern at the center of an open and longstanding debate over the firm-performance consequences of internal appointments (Shen and Cannella, 2002; Zhang and Rajagopalan, 2010; Quigley et al., 2019), our estimates in Table 5 reveal no long-run productivity advantage for insiders: if anything, the average internal candidate is

Table 4: Model Fit

Moment	Model	Data	t	Moment	Model	Data	t
<i>Profitability dynamics and variances</i>				<i>Hiring</i>			
ROA AR(1)	0.692	0.650	1.47	recent performance	0.036	0.054	-1.30
ROA AR(2)	0.162	0.162	0.01	older performance	0.030	0.033	-0.22
Mean ROA	0.120	0.120	0.09	industry performance	0.006	-0.004	0.71
Var. firm effect	0.0105	0.0109	-0.45	insider share	0.659	0.662	-0.36
Within-spell var.	0.0035	0.0036	-0.44	<i>Firing</i>			
Var. 1-yr change	0.0051	0.0047	1.24	recent performance	-0.012	-0.011	-0.29
Var. 2-yr change	0.0072	0.0071	0.27	older performance	-0.011	-0.007	-2.55
Btw-CEO var	0.0018	0.0022	-1.83	industry performance	-0.003	-0.004	0.11
<i>Insider premium by tenure</i>				forced rate	0.028	0.028	0.27
Year 1	0.016	0.020	-1.16	<i>Industry factor</i>			
Year 2	0.014	0.013	0.21	g AR(1)	0.717	0.710	0.15
Year 3	0.011	0.013	-0.71	g AR(2)	0.418	0.401	0.24
Years 4–6	0.007	0.003	1.13	Var. industry mean	0.0011	0.0011	0.19
Years 7–9	0.002	0.006	-0.97	<i>Baseline tenure profile (outsiders)</i>			
Years 10+	-0.001	0.006	-1.82	Year 3	0.005	-0.001	1.80
Succession dip	-0.009	-0.008	-0.63	Years 4–6	0.010	0.009	0.19
<i>Dismissal timing within spell, outsiders</i>				Years 7–9	0.014	0.008	1.35
Year 2	0.043	0.046	-0.49	Years 10+	0.017	0.012	1.30
Year 3	0.066	0.064	0.22	<i>Dismissal timing, insider difference</i>			
Years 4–6	0.086	0.093	-0.90	Year 2	-0.017	-0.016	-0.13
Years 7–9	0.096	0.108	-1.36	Year 3	-0.034	-0.020	-1.62
Years 10+	0.097	0.110	-1.22	Years 4–6	-0.049	-0.038	-1.32
				Years 7–9	-0.054	-0.047	-0.71
				Years 10+	-0.053	-0.056	0.31

Notes: Model moments are computed at the estimated parameter vector; t -statistics are the difference between model and data moments divided by the empirical standard error of the moment.

slightly less productive than their external counterpart, with $\mu_I = -0.47\%$ of firm assets.¹⁹ This is a small gap relative to the estimated cross-sectional stationary dispersion of CEO productivity, $\sigma_\theta = \sigma_\eta / \sqrt{(1 - \phi^2)} \approx 5.1\%$ of assets, which echoes the descriptive finding of Cziraki and Jenter (2021) that internal promotion dominates the market for CEOs even as insiders are not intrinsically superior. The discrepancy is driven in good part by how well boards *know* their candidates, rather than by what the candidates *are* on average.

Three main mechanisms explain the discrepancy, and they underscore the importance of accounting for the endogeneity of hiring and retention. First, internal successors inherit part of the trajectory of the regime that produced them ($\lambda = 0.48$), rendering them more

¹⁹Since our measure of profitability is the industry-adjusted ROA, the parameters in Table 5 are expressed in terms of a percentage of firm assets. We relegate estimates of the voluntary turnover model to Appendix D.

attractive candidates during good times, and resulting in genuinely better fundamentals from the start. This is consistent with prior descriptive observations showing that insiders are disproportionately likely to be hired by firms that have been performing well, while boards replacing failing CEOs tend to look outside (Parrino, 1997; Huson et al., 2001), an empirical regularity our model reproduces. Second, the inheritance mechanism is exacerbated by firm persistence. Given the estimated operational persistence, $\rho = 0.42$, firm performance following the appointment of an insider, who likely inherited better fundamentals, will on average be persistently higher. Third, and most importantly, we find that boards know their internal candidates far better than their external ones ($\psi_I = 0.60$ against $\psi_O = 0.06$), which likely reflects boards' prior information on internal personnel accumulated during their tenure in junior roles before CEO appointment (Baker et al., 1994).

Our estimates describe how well boards are informed at each of their two margins of action: what they learn about a candidate before an appointment, which we refer to as *selection*, and what they learn about an incumbent executive as they decide whether to retain them, which we refer to as *monitoring*. Consider first the selection margin. Boards can draw on a pre-hire signal about either candidate and, for an internal candidate, also on what they have already learned about the outgoing CEO whose quality the candidate partially inherits. Boards know little about external candidates before appointing them: the estimated precision of the signal about outsiders is $\psi_O = 0.06$, so boards resolve only about 6% of their uncertainty about an external candidate before the appointment. They know far more about internal candidates. Inheritance and the pre-hire signal together resolve about 63% of the board's uncertainty about an internal candidate's productivity, 17 percent through inheritance and 46 percent through the signal.²⁰ The estimated signal precision, $\psi_I = 0.60$, implies that boards resolve roughly 60% of the uncertainty that remains after inheritance (about χ_t^I in (5)) before making the hire, and the confidence interval rejects both no screening and perfect screening.²¹ Screening individual candidates, rather than knowledge about their inherited path, is thus the more important source of discriminating information.

Figure 2 shows what this difference in precision does to the CEOs boards appoint. In panel (a), the productivity distribution of appointed insiders lies 0.37 standard deviations

²⁰Both are variance reductions relative to σ_θ^2 , the board's uncertainty about a candidate before any pre-hire information. A board holding posterior variance Σ about the departing CEO faces prior variance $\lambda^2\Sigma + (1 - \lambda^2)\sigma_\theta^2$ about an internal candidate. Inheritance is therefore worth $\lambda^2(\sigma_\theta^2 - \Sigma)$, or 16.8 percent of σ_θ^2 at our estimates with Σ at its steady-state value. The signal resolves a further $\psi_I(1 - \lambda^2)\sigma_\theta^2$, or 45.7 percent, against $\psi_O\sigma_\theta^2$, or 6.1 percent, for an external candidate.

²¹The implied signal noise is $\sigma_{\nu,I} = 0.036$ while the standard deviation of an internal candidate's productivity innovation is $\sigma_{\chi,I} = 0.044$. The two combine into the screening precision as $\psi_I = \sigma_{\chi,I}^2 / (\sigma_{\chi,I}^2 + \sigma_{\nu,I}^2)$. Hence, the noise in what a board observes about an internal candidate ($\sigma_{\nu,I}^2$) is about two-thirds the size of the quality differences it is trying to detect ($\sigma_{\chi,I}^2$).

Table 5: Structural Estimates

Parameter	Description	Estimate
<i>CEO productivity and succession</i>		
ϕ	Productivity persistence	0.987 (0.007)
σ_η	Productivity-innovation s.d.	0.008 (0.001)
σ_θ	Implied stationary dispersion	0.051 (0.009)
μ_I	Insider productivity premium	-0.005 (0.002)
κ	Continuity (as fraction of ϕ)	0.491 (0.082)
$\lambda = \kappa\phi$	Implied inheritance	0.484 (0.082)
<i>Board information</i>		
ψ_I	Insider signal precision (share of candidate uncertainty resolved)	0.597 (0.129)
$\sigma_{\nu,I}$	Implied signal noise	0.036
ψ_O	Outsider signal precision	0.061 (0.015)
$\sigma_{\nu,O}$	Implied signal noise	0.199
$\psi_I - \psi_O$	Information advantage at hiring	0.536 (0.127)
ψ_g	Industry-performance filtering	0.756 (0.126)
σ_ω	Implied benchmark noise	0.018
<i>Firm profitability</i>		
$\bar{\iota}$	Firm-effect mean	0.095 (0.004)
σ_ι	Firm-effect s.d.	0.098 (0.004)
ρ	Operational persistence	0.423 (0.041)
σ_ε	Transitory shock s.d.	0.061 (0.002)
ϕ_g	Industry-factor persistence	0.509 (0.094)
σ_ζ	Industry-factor innovation s.d.	0.022 (0.002)
c	Monetary turnover cost	0.0023 (0.0026)
<i>Board preferences</i>		
π	Non-pecuniary firing cost	0.108 (0.011)
γ	Outsider hiring cost	0.060 (0.008)

Notes: Standard errors are clustered by firm and account for the first-stage estimation of the retirement hazard; we report them in parentheses. Since our measure of profitability is the industry-adjusted ROA, the parameters are expressed in terms of a percentage of firm assets. The standard errors for σ_θ , λ , ψ_I and $\psi_I - \psi_O$ are computed by the delta method.

of CEO productivity to the right of the pool of internal candidates, while appointed outsiders are nearly indistinguishable from the pool they are drawn from, shifting by only 0.12 standard deviations. Panels (b) and (c) show the two channels behind the insider shift: prior information and inheritance. Panel (b) shows that the appointment probability of an internal candidate rises from 0.11 for executives at the bottom of the productivity distribution to 0.99 for those at the top, while that of an external candidate rises only from 0.25 to 0.43: boards select sharply among internal candidates and only weakly among external ones. Panel (c) shows that inheritance has a similar but weaker impact: the probability of an internal appointment rises from 0.50 to 0.85 as the productivity of the departing CEO goes from the bottom to the top of the support. The gradient is flatter in panel (c) because quality transfers only partially across regimes, and because the board observes the departing executive's contribution with error.

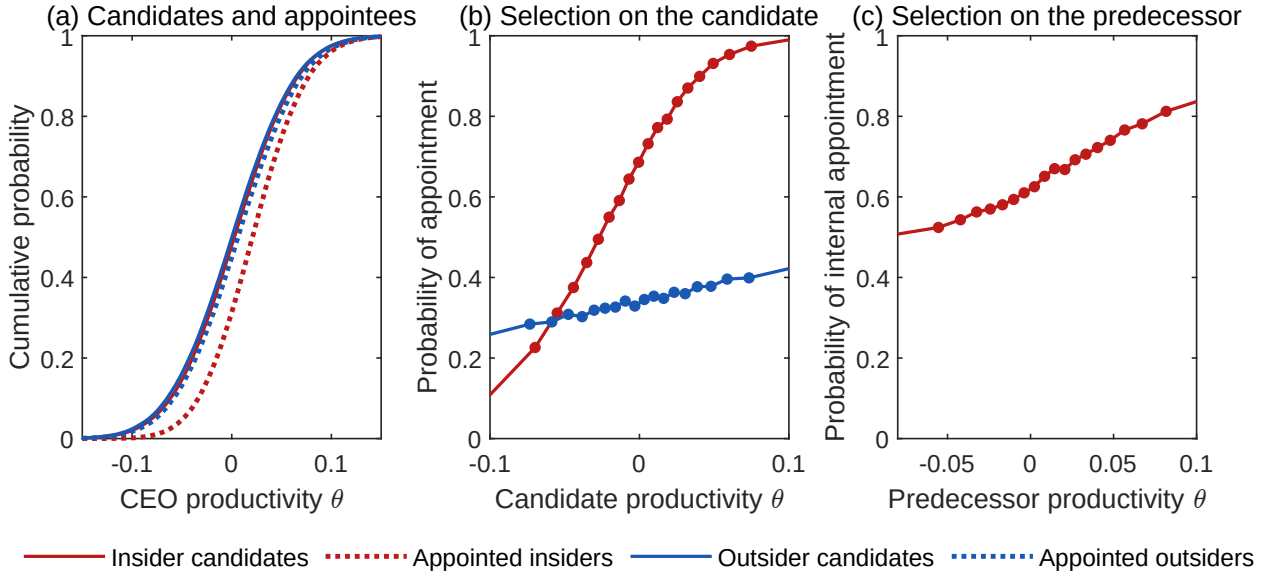


Figure 2: Selection at the Appointment Margin

Notes: Panel (a) plots CDFs of productivity for candidates and appointees of each type; the candidate distribution contains the candidate of that type on offer at each succession, whether or not that candidate was appointed. Panel (b) plots the probability of appointment against the candidate's own productivity, separately for the internal and the external candidate. Panel (c) plots the probability that the internal candidate is appointed against the productivity of the departing CEO; the probability of an external appointment is its complement, and the unconditional probability of an internal appointment is 0.66. Insiders in red, outsiders in blue. Simulated at the estimated parameters.

Consider now the monitoring margin, which begins after selection, as boards attempt to learn about the evolving productivity of their CEOs through the record they generate. Our estimates indicate that the evolving CEO productivity in equation (4) is highly persistent ($\phi = 0.99$) with modest volatility in year-to-year innovations ($\sigma_\eta = 0.8\%$ of assets). Hence,

productivity differences across CEOs ($\sigma_\theta \approx 5.1\%$ of assets) are large and durable, and the environment resembles the fixed-type benchmark of Jovanovic (1979) and Taylor (2010). The board’s problem is therefore dominated by learning about a durable attribute of considerable value, and the record it learns from is a poor signal of that attribute. CEO productivity shows up in a year’s profitability mixed with transitory shocks and industry conditions that the benchmark removes only in part. Expressed in units of productivity, the noise in a year of performance has a standard deviation of 10.8% of assets, about twice the dispersion of productivity across CEOs.²² A year of performance is thus three times as noisy as the pre-hire signal about an internal candidate ($\sigma_{\nu,I} = 3.6\%$) and about half as noisy as the signal about an external candidate ($\sigma_{\nu,O} = 19.9\%$). The record is slow to reveal a CEO’s productivity, consistent with the gradual learning that Taylor (2013) infers from CEO pay and with the evidence of Pan et al. (2015) that uncertainty about CEO ability remains substantial years into tenure. Set against this record, the two pre-hire signals differ sharply in what they are worth. In terms of precision, the signal about an internal candidate is equivalent to roughly nine years of performance, longer than most CEOs in our sample hold the job, while the signal about an external candidate is equivalent to less than four months. A board that promotes from within acts on information that the record would take most of a decade to supply, whereas a board that hires from outside must wait for the record.

Our estimated attribution parameter ($\psi_g = 0.76$) indicates that boards filter most, though not all, industry performance from their assessments of CEO productivity: an industry-driven surprise in profitability moves the board’s assessment of the CEO about a quarter as much as a firm-specific surprise of the same size. Boards are thus considerably more sophisticated in their treatment of industry conditions than an absolute-performance benchmark ($\psi_g = 0$) would suggest. Prior reduced-form evidence in the literature has established that industry conditions reach personnel decisions (Kaplan and Minton, 2012; Jenter and Kanaan, 2015) and pay (Bertrand and Mullainathan, 2001). Our results complement this literature by establishing how much industry performance boards credit to the CEO. Given our estimated magnitude of attribution, partial responsiveness to industry returns emerges as the rational use of an imperfect benchmark rather than a failure to benchmark at all.

Finally, we turn to the two non-pecuniary wedges in the board’s objective. At the hiring margin, we estimate a material cost of appointing an outsider ($\gamma = 0.060$), equivalent to about half a year of average operating profitability. A board therefore passes over an external candidate unless it expects that candidate to outperform the internal one by enough to

²²The profitability innovation in (6) loads on productivity with weight $1 - \rho$. Dividing it by $1 - \rho$ gives a signal of θ_{it} whose noise combines the transitory shock with the industry variation that the benchmark leaves unresolved. The standard deviation of this noise is 0.108 at our estimates, of which $\sigma_\varepsilon/(1 - \rho) = 0.105$ comes from the transitory shock alone.

cover the cost, consistent with the handicap outsiders face in CEO successions (Agrawal et al., 2006). At the firing margin, the estimated non-pecuniary firing cost ($\pi = 0.108$) is larger, amounting to roughly 90% of a year of average operating profitability, in line with the substantial entrenchment frictions found in the literature (Taylor, 2010; Lyman, 2024). These wedges govern how readily boards act on what they know, at the appointment and in monitoring, impacting the value of that information. We defer the welfare implications of these frictions, and the prices they imply for information at each margin, to the counterfactual analysis in Section 6.

6 The Value of Board Information

We now use the estimated model to quantify the value of board information at each of the two margins, selection and monitoring. We conduct three counterfactual exercises. First, we price the board’s information at each of the two margins separately, by replacing the board’s decision at that margin with an uninformed rule. Second, we measure how far boards stand from full information about their candidates and executives, and at which margin the unrealized value lies. Third, we focus on the selection margin and value separately each of the ways in which an internal candidate differs from an external one (average productivity, strategy inheritance, and how well the board knows the candidate), assessing the implications for the insider performance premium documented in Section 2. Each counterfactual solves the board’s problem under the new environment and simulates the economy. Our welfare metric is firm value, the expected present value of shareholder cash flows under the board’s policy, which incorporates the monetary cost of succession and excludes the non-pecuniary components of board utility, π and γ . We report percentage changes relative to the estimated baseline.

6.1 The Value of Information at Each Margin

On personnel decisions, boards are graded almost entirely on the monitoring margin, and largely through one statistic: how strongly dismissal responds to performance. A weak response is ambiguous. It may mean that a board knows too little to act on or that it is unwilling to act on what it knows, and reduced-form estimates cannot distinguish between the two (Engel et al., 2003). Our model separates the two because the precision of what boards observe and the wedge governing how readily they act on what they observe are distinct mechanisms in the model associated with different parameters. With this separation in hand, we price the board’s information at each margin by replacing the board’s optimal

policy with an uninformed rule. At the hiring margin, the successor’s type is drawn at random rather than conditioned on the board’s information, with sampling probabilities calibrated to the observed insider share; this way, the experiment destroys the information content of the choice without changing the composition of appointments. At the firing margin, dismissal occurs at random rather than on the board’s assessment, with the hazard calibrated to reproduce the observed forced-turnover rate, so that the experiment does not confound uninformed dismissal with more or less frequent dismissal. In each case, the other decision margin and the price of the firm re-optimize under the constraint.

Table 6: The Value of Board Information at Each Margin

	Δ Value (%)	Insider share	Forced rate
<i>Panel A: at the estimated parameters</i>			
Baseline	—	0.659	0.028
Uninformed selection	−9.61	0.660	0.035
Uninformed monitoring	−7.17	0.613	0.028
Uninformed at both margins	−19.34	0.660	0.028
<i>Panel B: no dismissal wedge ($\pi = 0$)</i>			
Baseline	—	0.799	0.235
Uninformed selection	−8.41	0.800	0.210
Uninformed monitoring	−15.95	0.697	0.235
Uninformed at both margins	−27.81	0.800	0.235
<i>Panel C: no outsider hiring cost ($\gamma = 0$)</i>			
Baseline	—	0.541	0.033
Uninformed selection	−10.11	0.540	0.036
Uninformed monitoring	−7.37	0.538	0.033
Uninformed at both margins	−19.46	0.542	0.033

Notes: Each row replaces the board’s decision at the indicated margin with an uninformed rule and re-solves the model, with the remaining margin and the price of the firm re-optimizing. Panels B and C first re-solve the model with the indicated preference parameter set to zero and then repeat both experiments inside that economy; percentage changes are always relative to the baseline of the panel in which they appear. In every panel the random successor is drawn at the insider share prevailing in that panel’s own baseline, and the random dismissal hazard is calibrated to reproduce that baseline’s forced-turnover rate. Holding the mix fixed at the hiring margin and the rate fixed at the firing margin is what makes the two prices comparable: each experiment removes the information content of a decision without changing how often, or in what proportions, that decision goes each way.

Table 6 shows that boards’ discretion is worth more at the hiring margin, 9.6% of firm value, than at the firing margin, 7.2%. However, information is worth only what boards do with it. The model contains two wedges between what a board does and what would maximize firm value: a non-pecuniary cost of dismissal, π , and a non-pecuniary cost of

external appointment, γ . A board that reads its CEO’s performance precisely but is reluctant to act on an unfavourable reading turns less of that information into firm value than a board that acts freely, and a long literature holds that boards are frequently of the first kind (Weisbach, 1988; Warner et al., 1988; Hermalin and Weisbach, 1998; Taylor, 2010).

Next, in Panels B and C of Table 6 we ask what information is worth when boards are not affected by these wedges. We repeat the experiments from Panel A while switching off each wedge separately. Without the dismissal wedge, boards replace their CEOs far more readily: the forced-turnover rate in the baseline of Panel B is 23.5% a year. In that economy the price of selection is little changed at 8.4%, while the price of monitoring more than doubles to 16.0%. Hence, the value of monitoring is suppressed by the wedge governing whether boards act on information at the firing margin, and selection gains relative strength in the presence of dismissal frictions (Panel A). The appointment wedge, by contrast, does little to the value of information. A cost of external appointment overrides the board’s information only when the two candidates are close in expected value, which is when the choice between them matters least. Removing it in Panel C raises the price of selection only from 9.6% to 10.1% and leaves the price of monitoring essentially unchanged.

6.2 The Distance to Full Information

After quantifying the value of available information at each decision margin, it is worth asking how much of the total value of information (available and not available) boards capture. Table 7 answers by endowing boards with perfect information at one or both margins. Perfect *selection* lets the board observe the productivity of both internal and external candidates at the moment of appointment. Perfect *monitoring* lets it observe the productivity of the incumbent CEO at all times.²³ Preference parameters are held at their estimates throughout, so these experiments vary what boards know while holding fixed how readily they act on it.

Panel A reports changes in firm value in the four resulting regimes. Full information raises firm value by 22.9%, measuring the total cost of the information frictions boards operate under. Notably, either margin delivers on its own a large part of the full information gain: perfect selection alone is worth 21.7% and perfect monitoring 14.8%. The two are therefore strong substitutes. Conditional on the other margin already being perfect, perfecting selection is worth 8.0 points and perfecting monitoring only 1.2. Hence, a board that

²³Perfect monitoring also removes any scope for misattributing industry conditions, since a board that observes productivity directly has no inference problem left to solve. Attribution carries little value on its own in any case: perfecting it (setting $\psi_g = 1$) while holding monitoring at its estimate raises firm value by 0.05%, and removing the board’s filtering of industry conditions altogether (setting $\psi_g = 0$) reduces firm value by 0.26%.

Table 7: The Distance to Full Information
Panel A: Firm value (% change from the estimated model)

	Baseline monitoring	Perfect monitoring	Gain from monitoring
Baseline selection	—	+14.83	+14.83
Perfect selection	+21.70	+22.86	+1.16
Gain from selection	+21.70	+8.03	

Panel B: Composition of turnover

	Insider share		Forced turnover rate	
	Baseline monitoring	Perfect monitoring	Baseline monitoring	Perfect monitoring
Baseline selection	0.659	0.551	0.028	0.058
Perfect selection	0.667	0.667	0.061	0.063

Notes: Perfect selection endows the board with direct observation of both candidates' productivity at the moment of appointment. Perfect monitoring endows it with direct observation of the incumbent's productivity, which also eliminates misattribution of industry conditions, since a board that observes productivity has no inference problem to solve. Preference parameters are held at their estimates in every regime. In Panel A the final column and the final row report the marginal effect of perfecting one margin while holding the other fixed. Panel B reports the insider share of new appointments and the forced-turnover rate in the same four regimes.

knows its candidates has little left to learn about the CEO it installed, and a board that can identify a failing CEO at once has less need to have screened the candidate beforehand.

Panel B reports the impact of full information on turnover. Better information makes boards dismiss about twice as often: the forced-turnover rate is 2.8% in the estimated economy and lies between 5.8% and 6.3% in every informed regime. Perfect selection alone raises the forced-turnover rate as much as perfect monitoring does. The reason is that the value of dismissal depends on the quality of the replacement, so a board that knows it can appoint well is more willing to remove the CEO it has. The substitution in Panel A reappears even more sharply at this margin: perfecting monitoring raises the forced-turnover rate by 3.0 percentage points when selection is at its estimated precision, but by only 0.2 points when selection is already perfect.

Taken together, the two exercises in Sections 6.1 and 6.2 place boards between two benchmarks: a board that makes personnel decisions at random, and a board that observes CEO productivity perfectly. Boards have valuable information about their candidates and executives: relative to uninformed decisions at both margins, their use of information is worth

19.3% of firm value. But they are far from being fully informed: complete information about CEO productivity would add another 22.9%. Selection is the more valuable margin on both counts: boards' existing information is worth more at selection than at monitoring (9.6% against 7.2%, Table 6), and so is the value still unrealized (21.7% against 14.8%, Table 7). Boards' information about CEO quality is, by these estimates, a first-order determinant of firm value, especially through selection.

6.3 Information and the Insider Premium

The previous counterfactual exercises price what boards know about candidates while holding fixed what candidates are. We now price the distinct features of an internal candidate and assess the impact of these features on the insider performance premium documented in Section 2. Internal candidates differ from external ones in three ways: their average productivity may differ (μ_I); they carry forward the strategy of the regime that produced them (λ) becoming attractive prospects for successful organizations seeking continuity; and boards know them better (ψ_I against ψ_O) thanks to their prior junior roles inside the firm. Relative to outsiders, the first two features make insiders genuinely different executives in their fundamentals, while the third one makes insiders differently known by boards. Table 8 prices each unique feature of an insider candidate.

The largest single lever is screening. Removing the board's pre-hire signal about internal candidates reduces firm value by 10.0%: unable to select among internal candidates, boards appoint insiders whose average productivity is 0.7% of assets, against 2.2% in the estimated model. Boards also dismiss half as often (the forced-turnover rate falls from 2.8% to 1.4%), because dismissal is worth less when the replacement cannot be chosen well. Conversely, endowing boards with a perfect pre-hire signal about internal candidates raises firm value by only 4.3%, since boards already resolve most of their uncertainty about these candidates, through the signal and through the record of the regime the candidate inherits. This is far short of the 21.7% available under perfect selection in Table 7: the remaining 17 points come from observing the external candidate as well, about whom boards know little.

The remaining two features matter far less. The difference in type means works against insiders: the average internal candidate is slightly less productive than the average external one ($\mu_I = -0.47\%$ of assets), and removing the difference raises firm value by 2.3%. Removing continuity costs only 0.9% of firm value, even though it lowers the average productivity of appointed insiders from 2.2% to 1.3% of assets. Inheritance cuts both ways: an internal candidate carries forward a successful regime when the predecessor retires and an unsuccessful one when the predecessor is dismissed. And because it lowers the quality of the

replacement precisely when a CEO has failed, it makes dismissal less attractive: without it, boards dismiss more readily (the forced-turnover rate rises from 2.8% to 3.7%), which offsets much of the loss in the quality of appointments. Eliminating all three differences at once, so that internal candidates are drawn from the outsider distribution and observed with the outsider’s precision, reduces firm value by 9.6%. This is almost exactly the cost of removing screening alone, so the value of the internal succession pipeline rests on the information it produces: the option to select on fit revealed before the appointment, the margin that distinguishes an internal promotion from an external hire (Chan, 1996; Bidwell, 2011).

Table 8: Counterfactuals: The Value of the Features of Insider Candidates

	Δ Value (%)	Insider share	Forced rate	CEO quality (%)	
				Insiders	Outsiders
Baseline	—	0.659	0.028	2.2	0.6
No screening ($\psi_I = 0$)	−10.02	0.644	0.014	0.7	0.9
Perfect screening ($\psi_I = 1$)	+4.34	0.671	0.038	2.8	0.5
Equal type means ($\mu_I = 0$)	+2.29	0.692	0.030	2.4	0.6
No continuity ($\lambda = 0$)	−0.93	0.665	0.037	1.3	0.8
Insiders equal to outsiders	−9.56	0.708	0.019	0.4	1.1

Notes: Each row re-solves the board’s problem and re-simulates the economy under the indicated change, holding all remaining parameters at their estimates. Δ Value is the percentage change in firm value relative to the baseline, and the next two columns report the insider share of successions and the annual forced-turnover rate under the counterfactual policy. CEO quality is the average realized productivity θ of appointed executives *measured in the year of appointment*, in percent of assets, reported separately by type; a dash indicates that no executive of that type is appointed. In the last row insider candidates are drawn from the outsider distribution and observed with the outsider’s precision ($\psi_I = \psi_O$), so the two candidates differ only in the cost of hiring from outside.

Given how valuable internal prospects are, and that their value stems from the information boards have about them, we now ask what would happen if firms’ appointment policies could not respond to this information. Concretely, we ask, what would boards lose if the composition of appointments were dictated rather than chosen, leaving the firing margin as the only decision node? Table 9 dictates the successor’s type, leaving every parameter of the model at its estimate. Every dictated composition is costly: forcing all hires to be internal costs 6.8% of firm value, an even split drawn at random 10.7%, and forcing all hires to be external 14.1%.²⁴ A dictated rule appoints whichever internal candidate is on offer, and

²⁴The uninformed-selection row of Table 6 is this same forced rule at the observed insider share, and costs 9.6%. A board deprived of its choice would therefore rather commit to promoting internally than reproduce its own observed mix, because the outsiders it appoints in equilibrium are appointed precisely when the internal candidate is weak; appointing that same fraction at random forfeits the reason for appointing them.

that candidate inherits from the predecessor: internal appointees average -0.8% of assets under all-internal promotion, where dismissals are more frequent (5.2% against 2.9%) and every failed regime passes to an insider, and -0.6% under the even split, which breaks such runs with external draws. External appointment is nonetheless the costlier rule, although external appointees are no less productive than internal ones (0.0% against -0.8%), because it gives up what the board knows about internal candidates and the inheritance of successful regimes. That knowledge retains its value when the choice is dictated: a board that must promote internally still knows which of its appointees are weak and dismisses them early, at a forced-turnover rate of 5.2% against 1.4% when every appointment is external.

Table 9: Counterfactuals: Forced Composition of Appointments

	Δ Value (%)	Insider share	Forced rate	CEO quality (%)	
				Insiders	Outsiders
Optimal hiring (baseline)	—	0.659	0.028	2.2	0.6
All insiders	-6.81	1.000	0.052	-0.8	—
Fifty-fifty	-10.74	0.500	0.029	-0.6	0.0
All outsiders	-14.05	0.000	0.014	—	0.0

Notes: Each row imposes the indicated composition of appointments; the firing policy re-optimizes and the information structure is unchanged. These experiments change the mix of hires rather than the information the board brings to the choice, and are therefore distinct from the uninformed-selection row of Table 6, which holds the mix at its observed value. CEO quality is the average realized productivity θ of appointed executives, measured in the year of appointment and defined as in Table 8; a dash indicates that no executive of that type is appointed. Under a dictated rule the appointed insider is the candidate on offer, so the insider column is the average candidate's productivity under that rule; it differs across rules because candidates inherit from their predecessors. Comparing the first two rows isolates the value of the board's discretion: when the board chooses, the insiders it promotes average 2.2% of assets, and when internal promotion is imposed they average -0.8% .

Tables 8 and 9 also bear directly on the interpretation of the insider performance premium documented in Section 2. In the estimated model, appointed insiders are more productive than appointed outsiders in the year of appointment, 2.2% of assets against 0.6% . The gap can come from two sources: internal candidates may be more productive than external ones before any selection takes place, or the board may choose better among internal candidates than among external ones. The first source is small and favors outsiders. Without the board's pre-hire signal, appointed insiders average 0.7% against 0.9% for outsiders (Table 8); with the composition of appointments dictated, so that neither candidate is selected, they average -0.8% to -0.6% against 0.0% (Table 9). Selection on what the board knows accounts for the whole advantage, and more. The advantage of appointed insiders is therefore informational

rather than an endowment of the internal pipeline. This offers a resolution to the inconclusive insider-versus-outsider performance debate introduced in Section 5: the two candidate pools are similar, and insiders' observed advantage is manufactured by what boards know about them at the time of hire. It also explains why outsider appointments are the higher-variance bet emphasized by Quigley et al. (2019): outsiders are only weakly screened draws.

7 Conclusion

Boards are genuinely but partially informed when deciding whether to dismiss a CEO and who to replace him with. However, only the dismissal margin has a standard measure of appropriate board responses to information: the sensitivity of termination to firm performance. We find that boards resolve roughly sixty percent of the uncertainty about a candidate's productivity before an internal appointment, enough to reject both the view that internal promotion is uninformed or fully informed. Before an external appointment they resolve almost none of the uncertainty. After an appointment boards learn slowly, because a year of firm performance is a noisy signal of CEO productivity: the pre-hire signal about an internal candidate is worth roughly nine years of firm performance signals, and the signal about an external candidate only a few months. Additionally, firms filter most, though not all, of industry performance out of their assessments: personnel decisions respond to industry conditions at roughly a quarter of the rate at which they respond to comparable firm-specific news.

Boards' use of what they know is worth about ten percent of firm value during candidate selection and seven percent while monitoring incumbents, so boards' informed decisions in the less-studied margin of candidate selection are worth more. The information value not yet captured is also larger at selection, and the two margins are strong substitutes: perfect information at either margin alone (selection or monitoring) would deliver most of what full information offers. Boards' attribution of industry performance, by contrast, carries little firm value.

We find that the performance advantage of insider CEOs reflects different levels of information at candidate selection rather than an inherently superior internal pool. Internal candidates are no more productive on average than external ones. Appointed insiders outperform because boards select sharply among internal candidates who they know well, which they are unable to do among external prospects who they do not know nearly as well. The insider performance advantage disappears once the board's pre-hire information or its ability to act on it is removed.

Two limitations bound our conclusions. We model no compensation contract, so the

board's problem here is one of information rather than of incentive provision, and the value we assign to information is the value of better personnel decisions alone. Additionally, we treat the board as a single decision maker, setting aside the board composition and independence that a large governance literature connects to personnel outcomes.

A board that rarely dismisses a poor performer may be unwilling to act or unable to tell, and the turnover-performance sensitivity cannot separate the two, although they call for different remedies: one addressed to the incentives of directors, the other to the information reaching them. Our estimates say that both bind, and that the dismissal wedge (the board's disutility from firing) is large enough to suppress a substantial part of what monitoring information would otherwise be worth to shareholders. A weak turnover-performance sensitivity is therefore not by itself evidence that a board is failing, and distinguishing the two readings requires a measure of what boards observe rather than a better measure of what they do.

References

- Agrawal, Anup, Charles R. Knoeber, and Theofanis Tsoulouhas**, “Are outsiders handicapped in CEO successions?,” *Journal of Corporate Finance*, 2006, 12 (3), 619–644.
- Allgood, Sam and Kathleen A. Farrell**, “The Match between CEO and Firm,” *The Journal of Business*, 2003, 76 (2), 317–341.
- Altonji, Joseph G. and Charles R. Pierret**, “Employer Learning and Statistical Discrimination,” *The Quarterly Journal of Economics*, 2001, 116 (1), 313–350.
- **and Lewis M. Segal**, “Small-Sample Bias in GMM Estimation of Covariance Structures,” *Journal of Business & Economic Statistics*, 1996, 14 (3), 353–366.
- Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro**, “Measuring the Sensitivity of Parameter Estimates to Estimation Moments,” *The Quarterly Journal of Economics*, 2017, 132 (4), 1553–1592.
- Baker, George, Michael Gibbs, and Bengt Holmstrom**, “The Internal Economics of the Firm: Evidence from Personnel Data,” *The Quarterly Journal of Economics*, 1994, 109 (4), 881–919.
- Bertrand, Marianne and Antoinette Schoar**, “Managing with Style: The Effect of Managers on Firm Policies,” *The Quarterly Journal of Economics*, 2003, 118 (4), 1169–1208.
- **and Sendhil Mullainathan**, “Are CEOs Rewarded for Luck? The Ones Without Principals Are,” *The Quarterly Journal of Economics*, 2001, 116 (3), 901–932.
- Bidwell, Matthew**, “Paying More to Get Less: The Effects of External Hiring versus Internal Mobility,” *Administrative Science Quarterly*, 2011, 56 (3), 369–407.
- Bushman, Robert, Zhonglan Dai, and Xue Wang**, “Risk and CEO turnover,” *Journal of Financial Economics*, 2010, 96 (3), 381–398.
- Celentano, Francesco and Antonio Mello**, “Why Do Firms Often Not Have a CEO Succession Plan?,” *Working Paper*, 2024.
- Chan, William**, “External Recruitment versus Internal Promotion,” *Journal of Labor Economics*, 1996, 14 (4), 555–570.

- Chen, Guoli and Donald C. Hambrick**, “CEO Replacement in Turnaround Situations: Executive (Mis)Fit and Its Performance Implications,” *Organization Science*, 2012, *23* (1), 225–243.
- Cremers, Martijn and Yaniv Grinstein**, “Does the Market for CEO Talent Explain Controversial CEO Pay Practices?,” *Review of Finance*, 2014, *18* (3), 921 – 960.
- Custódio, Cláudia, Miguel A. Ferreira, and Pedro Matos**, “Generalists versus specialists: Lifetime work experience and chief executive officer pay,” *Journal of Financial Economics*, 2013, *108* (2), 471–492.
- Cvijanović, Dragana, Nickolay Gantchev, and Rachel Li**, “CEO Succession Roulette,” *Management Science*, 2023, *69* (10), 5794–5815.
- Cziraki, Peter and Dirk Jenter**, “The Market for CEOs,” *European Corporate Governance Institute - Finance Working Paper No. 831/2022*, 2021.
- Demarzo, Peter M. and Yuliy Sannikov**, “Learning, Termination, and Payout Policy in Dynamic Incentive Contracts,” *The Review of Economic Studies*, 07 2017, *84* (1), 182–236.
- DeVaro, Jed and Hodaka Morita**, “Internal Promotion and External Recruitment: A Theoretical and Empirical Analysis,” *Journal of Labor Economics*, 2013, *31* (2), 227–269.
- Engel, Ellen, Rachel M. Hayes, and Xue Wang**, “CEO turnover and properties of accounting information,” *Journal of Accounting and Economics*, 2003, *36* (1-3), 197–226.
- Fee, C. Edward and Charles J. Hadlock**, “Raids, Rewards, and Reputations in the Market for Managerial Talent,” *The Review of Financial Studies*, 2003, *16* (4), 1315–1357.
- Ferreira, Daniel and Radosława Nikolowa**, “Talent Discovery and Poaching under Asymmetric Information,” *The Economic Journal*, 2023, *133* (649), 201–234.
- Frydman, Carola**, “Rising Through the Ranks: The Evolution of the Market for Corporate Executives, 1936–2003,” *Management Science*, November 2019, *65* (11), 4951–4979.
- Gayle, George-Levi and Robert A. Miller**, “Has Moral Hazard Become a More Important Factor in Managerial Compensation?,” *American Economic Review*, December 2009, *99* (5), 1740–69.
- , **Limor Golan, and Robert A. Miller**, “Promotion, Turnover, and Compensation in the Executive Labor Market,” *Econometrica*, 2015, *83* (6), 2293–2369.

- Geelen, Thomas and Jakub Hajda**, “Succession,” *Working Paper*, 2024.
- Gentry, Richard J., Joseph S. Harrison, Timothy J. Quigley, and Steven Boivie**, “A Database of CEO Turnover and Dismissal in S&P 1500 Firms, 2000–2018,” *Strategic Management Journal*, 2021, 42 (4), 968–991.
- Gibbons, Robert and Kevin J. Murphy**, “Relative Performance Evaluation for Chief Executive Officers,” *Industrial and Labor Relations Review*, 1990, 43 (3), 30S–51S.
- Graham, John R., Hyunseob Kim, and Mark Leary**, “CEO-board dynamics,” *Journal of Financial Economics*, 2020, 137 (3), 612–636.
- Greenwald, Bruce C. N.**, *Adverse Selection in the Labor Market* 1979.
- He, Li and Enrique J. Schroth**, “Why are CEOs rarely succeeded by external candidates?,” *Working Paper*, 2024.
- Hermalin, Benjamin and Michael Weisbach**, “Endogenously Chosen Boards of Directors and Their Monitoring of the CEO,” *The American Economic Review*, 1998, 88 (1), 96–118.
- Hermalin, Benjamin E.**, “Trends in Corporate Governance,” *The Journal of Finance*, 2005, 60 (5), 2351–2384.
- **and Michael S. Weisbach**, “Assessing Managerial Ability: Implications for Corporate Governance,” in Benjamin E. Hermalin and Michael S. Weisbach, eds., *The Handbook of the Economics of Corporate Governance*, Vol. 1 of *The Handbook of the Economics of Corporate Governance*, North-Holland, 2017, pp. 93–176.
- Holmström, Bengt**, “Moral Hazard and Observability,” *The Bell Journal of Economics*, 1979, 10 (1), 74–91.
- Huson, Mark R., Robert Parrino, and Laura T. Starks**, “Internal Monitoring Mechanisms and CEO Turnover: A Long-Term Perspective,” *The Journal of Finance*, 2001, 56 (6), 2265–2297.
- Jenter, Dirk and Fadi Kanaan**, “CEO Turnover and Relative Performance Evaluation,” *The Journal of Finance*, 2015, 70 (5), 2155–2184.
- **and Katharina Lewellen**, “Performance-Induced CEO Turnover,” *The Review of Financial Studies*, 06 2020, 34 (2), 569–617.

- Jovanovic, Boyan**, “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 1979, 87 (5), 972–90.
- Kaplan, Steven N. and Bernadette A. Minton**, “How Has CEO Turnover Changed?,” *International Review of Finance*, 2012, 12 (1), 57–87.
- Lange, Fabian**, “The Speed of Employer Learning,” *Journal of Labor Economics*, 2007, 25 (1), 1–35.
- Lazear, Edward P.**, “Hiring Risky Workers,” in Isao Ohashi and Toshiaki Tachibanaki, eds., *Internal Labour Markets, Incentives and Employment*, Palgrave Macmillan, 1998, chapter 7, pp. 143–158.
- Lyman, Noah**, “Reputation, Entrenchment, and Dynamic Managerial Incentives,” *Working Paper*, 2024.
- Margiotta, Mary M and Robert A Miller**, “Managerial Compensation and the Cost of Moral Hazard,” *International Economic Review*, 2000, 41 (3), 669–719.
- McConnell, John J. and Qianru Qi**, “Does CEO Succession Planning (Disclosure) Create Shareholder Value?,” *Journal of Financial and Quantitative Analysis*, 2022, 57 (6), 2355–2384.
- Mobbs, Shawn and Charu G. Raheja**, “Internal managerial promotions: Insider incentives and CEO succession,” *Journal of Corporate Finance*, 2012, 18 (5), 1337 – 1353.
- Murphy, Kevin J. and Jan Zabojsnik**, “Managerial Capital And The Market For Ceos,” *Working Paper*, October 2006.
- Pan, Yihui**, “The Determinants and Impact of Executive-Firm Matches,” *Management Science*, 2017, 63 (1), 185–200.
- , **Tracy Yue Wang, and Michael S. Weisbach**, “Learning About CEO Ability and Stock Return Volatility,” *The Review of Financial Studies*, 02 2015, 28 (6), 1623–1666.
- Parrino, Robert**, “CEO turnover and outside succession A cross-sectional analysis,” *Journal of Financial Economics*, 1997, 46 (2), 165–197.
- Quigley, Timothy J., Donald C. Hambrick, Vilmos F. Misangyi, and G. Alessandra Rizzi**, “CEO selection as risk-taking: A new vantage on the debate about the consequences of insiders versus outsiders,” *Strategic Management Journal*, 2019, 40 (9), 1453–1470.

- Schwartz-Ziv, Miriam and Michael S. Weisbach**, “What do boards really do? Evidence from minutes of board meetings,” *Journal of Financial Economics*, 2013, 108 (2), 349–366.
- Shen, Wei and Albert A. Cannella**, “Revisiting the Performance Consequences of CEO Succession: The Impacts of Successor Type, Postsuccession Senior Executive Turnover, and Departing CEO Tenure,” *Academy of Management Journal*, 2002, 45 (4), 717–733.
- Taylor, Lucian**, “Why Are CEOs Rarely Fired? Evidence from Structural Estimation,” *The Journal of Finance*, 2010, 65 (6), 2051–2087.
- Taylor, Lucian A.**, “CEO wage dynamics: Estimates from a learning model,” *Journal of Financial Economics*, 2013, 108 (1), 79–98.
- Warner, Jerold B., Ross L. Watts, and Karen H. Wruck**, “Stock prices and top management changes,” *Journal of Financial Economics*, 1988, 20, 461–492. The Distribution of Power Among Corporate Managers, Shareholders, and Directors.
- Weisbach, Michael S.**, “Outside directors and CEO turnover,” *The Journal of Financial Economics*, 1988, 20, 431–460.
- Zhang, Yan and Nandini Rajagopalan**, “Once an outsider, always an outsider? CEO origin, strategic change, and firm performance,” *Strategic Management Journal*, 2010, 31 (3), 334–346.

A Data Appendix

This appendix describes the construction of the estimation sample: the data sources, the performance measures and the industry benchmark, the CEO panel and the classification of successions, and the filters that produce the final sample.

Sources. We combine four sources. Compustat North America provides firm fundamentals for the universe of publicly traded firms in the United States and Canada. CRSP provides monthly stock returns, which we compound to annual returns aligned with the fiscal year. ExecuComp provides executive-level records for S&P 1500 firms, from which we build the CEO panel; we merge the full executive file rather than the CEO file alone, because an executive’s non-CEO records at the firm are what identify prior experience at the time of appointment. Finally, we classify CEO departures as forced or voluntary using the CEO dismissal database of Gentry et al. (2021), which codes the circumstances of each departure of an S&P 1500 chief executive from press coverage.

Performance measures. Our measure of profitability is operating return on assets: operating income before depreciation scaled by total assets. Market-adjusted stock returns are compound annual firm returns net of the compound value-weighted market return over the same fiscal year. Both raw measures are winsorized at the 1st and 99th percentiles.

The industry benchmark. Industry conditions are measured at the level of GICS industry groups (24 groups). For profitability and for market-adjusted returns we compute industry-year means over the entire Compustat/CRSP universe, before the merge with ExecuComp restricts the panel to S&P 1500 firms, so that the benchmark for an industry does not depend on which of its members happen to appear in the estimation sample. A firm’s industry-adjusted performance is its raw performance minus this universe mean, winsorized again at the 1st and 99th percentiles. Computing the benchmark on the universe rather than within the sample matters for the industry moments targeted in estimation: sample-only industry cells are small enough that a within-sample mean would carry substantial sampling noise into the industry-adjusted series.

The CEO panel. From the merged Compustat-CRSP-ExecuComp panel we retain firm-years with non-missing operating income, assets, and fiscal year, and drop firms with non-contiguous records. We identify each firm-year’s CEO from ExecuComp’s annual CEO flag and drop the small number of firm-years with more than one executive so flagged. Appointment and departure years come from ExecuComp’s service dates; an executive who becomes

CEO in the second half of the year is assigned the following year as the first year of the spell. We classify CEOs as insiders or outsiders using prior tenure with the firm at the time of appointment: insiders are CEOs who are founders or who have more than one year of prior experience at the firm, and all other CEOs are outsiders. This threshold follows Parrino (1997).

Forced and voluntary turnover. A succession is classified as forced when the departing CEO's exit is coded as a dismissal in Gentry et al. (2021), and as voluntary when the CEO is replaced without a dismissal code. The database codes each departure from contemporaneous press reports, classifying it as a dismissal when the board removed the executive rather than the executive leaving of their own accord. Because press coverage understates board-initiated exits that are presented as retirements, this classification is conservative. The database's coverage ends in 2019, which sets the end of our sample window.

Final sample. We drop CEO-years in which the CEO cannot be classified as an insider or an outsider; most of these are interim appointments. The final sample covers fiscal years 1995 through 2019 and contains 38,213 firm-year observations on 3,090 distinct firms and 6,307 CEO spells. For estimation we export one observation per firm-year containing operating ROA, the market-adjusted return, their industry-year benchmark means and industry-adjusted values, the insider indicator, the forced-turnover indicator, an indicator for a new CEO's first year, and CEO tenure. All regressions and moments used in estimation are computed from this file, on the simulated panel and the data identically.

B Learning Dynamics: Proof of Proposition 1

The filtering problem. Stack the latent states $q_t = (\theta_{it} - \mu_h, g_{nt})'$ and the signals $z_t = (u_{jt} - (1 - \rho)\mu_h, s_{nt}^g)'$. From (4), (3), (6), and (7), the system has the linear state-space form

$$q_{t+1} = F q_t + w_{t+1}, \quad z_t = H q_t + e_t, \quad (22)$$

$$F = \begin{pmatrix} \phi & 0 \\ 0 & \phi_g \end{pmatrix}, \quad H = \begin{pmatrix} 1 - \rho & 1 \\ 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\zeta^2 \end{pmatrix}, \quad R_e = \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_\omega^2 \end{pmatrix},$$

with $w_t \sim N(0, Q)$ and $e_t \sim N(0, R_e)$ mutually independent. Let $\hat{q}_{t|t-1}$ and $\Sigma_{t|t-1}$ denote the one-step-ahead mean and covariance of q_t given the signal history. The Kalman recursions

are

$$r_t = z_t - H\hat{q}_{t|t-1}, \quad S_t = H\Sigma_{t|t-1}H' + R_e, \quad (23)$$

$$\hat{q}_{t|t} = \hat{q}_{t|t-1} + \Sigma_{t|t-1}H'S_t^{-1}r_t, \quad \Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma_{t|t-1}H'S_t^{-1}H\Sigma_{t|t-1}, \quad (24)$$

$$\hat{q}_{t+1|t} = F\hat{q}_{t|t}, \quad \Sigma_{t+1|t} = F\Sigma_{t|t}F' + Q. \quad (25)$$

Because $\phi, \phi_g < 1$, the state is stationary and the Riccati iteration implied by (24)–(25) converges to a unique fixed point Σ^* from any initial covariance (Anderson and Moore, 1979). Steady-state learning evaluates the recursions at Σ^* , with stationary gain

$$K = \Sigma^*H'(H\Sigma^*H' + R_e)^{-1} \equiv \begin{pmatrix} K_u & K_g \\ L_u & L_g \end{pmatrix}. \quad (26)$$

whose first row (K_u, K_g) collects the loadings of the board's belief about θ on the performance and benchmark signals, and whose second row (L_u, L_g) collects those of its belief about g . Write $p_\theta = \Sigma_{11}^*$, $p_g = \Sigma_{22}^*$, $p_{\theta g} = \Sigma_{12}^*$, and define the signal loadings

$$a_1 \equiv (1 - \rho)p_\theta + p_{\theta g}, \quad a_2 \equiv (1 - \rho)p_{\theta g} + p_g, \quad (27)$$

which are the covariances of the performance innovation u with the two latent states given the board's information. The θ -row of (26) is then

$$\begin{pmatrix} K_u & K_g \end{pmatrix} = \begin{pmatrix} a_1 & p_{\theta g} \end{pmatrix} \begin{pmatrix} (1 - \rho)a_1 + a_2 + \sigma_\varepsilon^2 & a_2 \\ a_2 & p_g + \sigma_\omega^2 \end{pmatrix}^{-1}. \quad (28)$$

The updating rule. The θ -component of the prediction step in (25), evaluated at the stationary gain and with the type mean added back, reads $\tilde{\theta}_{it+1} - \mu_h = \phi[(\tilde{\theta}_{it} - \mu_h) + K_u r_{u,t} + K_g r_{g,t}]$, where $r_{u,t} = u_{jt} - \hat{u}_{jt}$ and $r_{g,t} = s_{nt}^g - \hat{g}_{nt}$ are the two components of the innovation r_t in (23). Defining the benchmark-adjustment share $\psi_g \equiv -K_g/K_u$ and factoring K_u out of the bracket yields equation (8).

Signs of the gains. The sign claims in part (i) follow from the signs of the stationary covariance, which we establish first.

Lemma 1. *Suppose $\phi_g \leq \phi$. At the stationary solution, $p_{\theta g} < 0$, $a_1 > 0$, and $a_2 > 0$.*

PROOF: Because the two signals are conditionally independent, the update (24) can be computed sequentially: first on the benchmark signal s^g (loading $(0, 1)$), then on the performance innovation u (loading $(1 - \rho, 1)$). We show that the properties $\{p_{\theta g} \leq 0, a_1 > 0, a_2 > 0\}$

are preserved by each step of the filter iteration; since the iteration converges to the unique fixed point from the diagonal initialization $\Sigma_0 = \text{diag}(\sigma_\theta^2, \sigma_\zeta^2/(1 - \phi_g^2))$, which satisfies them, the fixed point inherits them.

Benchmark update. With loading $(0, 1)$ and noise σ_ω^2 , the update maps $p_g \mapsto p_g \sigma_\omega^2 / (p_g + \sigma_\omega^2)$, $p_{\theta g} \mapsto p_{\theta g} \sigma_\omega^2 / (p_g + \sigma_\omega^2)$, and $p_\theta \mapsto p_\theta - p_{\theta g}^2 / (p_g + \sigma_\omega^2)$. Direct computation gives $a'_2 = a_2 \sigma_\omega^2 / (p_g + \sigma_\omega^2) > 0$ and $a'_1 = a_1 + (-p_{\theta g}) a_2 / (p_g + \sigma_\omega^2) \geq a_1 > 0$; the sign of $p_{\theta g}$ is preserved.

Performance update. With loading $(1 - \rho, 1)$ and noise σ_ε^2 , write $D = (1 - \rho)a_1 + a_2 + \sigma_\varepsilon^2 > 0$. The update maps $p_{\theta g} \mapsto p_{\theta g} - a_1 a_2 / D < 0$ (strictly, since $a_1, a_2 > 0$), and direct computation gives $a_1 \mapsto a_1 \sigma_\varepsilon^2 / D > 0$ and $a_2 \mapsto a_2 \sigma_\varepsilon^2 / D > 0$.

Prediction. The prediction step maps $p_{\theta g} \mapsto \phi \phi_g p_{\theta g} \leq 0$ and, using $\phi_g \leq \phi$, $a_1 \mapsto (1 - \rho)(\phi^2 p_\theta + \sigma_\eta^2) + \phi \phi_g p_{\theta g} \geq \phi \phi_g a_1 + (1 - \rho)\sigma_\eta^2 > 0$. For a_2 , the prediction step gives $a_2 \mapsto \phi_g^2 a_2 + (1 - \rho)\phi_g(\phi - \phi_g) p_{\theta g} + \sigma_\zeta^2$, which is positive whenever $\sigma_\zeta^2 \geq (1 - \rho)\phi_g(\phi - \phi_g) |p_{\theta g}|$; this inequality holds automatically when $\phi_g = \phi$, and we verify that it holds at the stationary solution for every parameter configuration considered in the paper. ■

Given Lemma 1, part (i) follows from (28): writing $\det S > 0$ for the determinant of the innovation variance,

$$K_u = \frac{a_1(p_g + \sigma_\omega^2) - p_{\theta g} a_2}{\det S} > 0, \quad K_g = \frac{p_{\theta g}[(1 - \rho)a_1 + a_2 + \sigma_\varepsilon^2] - a_1 a_2}{\det S} < 0, \quad (29)$$

since every term in the numerator of K_u is positive and every term in the numerator of K_g is negative. Hence $K_u > 0$ and $\psi_g = -K_g/K_u > 0$, establishing part (i).

Limits. As $\sigma_\omega \rightarrow \infty$, the benchmark signal carries no information: $K_g \rightarrow 0$ at rate σ_ω^{-2} while K_u converges to the positive gain of the one-signal filter, so $\psi_g \rightarrow 0$. As $\sigma_\omega \rightarrow 0$, the benchmark reveals g_{nt} exactly. The board's information is then equivalent to $(u_{jt} - s_{nt}^g, s_{nt}^g)$, where $u_{jt} - s_{nt}^g = (1 - \rho)\theta_{it} + \varepsilon_{jt}$ is a signal about the CEO alone. Updating on it with coefficient k and using $\hat{u}_{jt} = (1 - \rho)\tilde{\theta}_{it} + \hat{g}_{nt}$,

$$k[(u_{jt} - s_{nt}^g) - (1 - \rho)\tilde{\theta}_{it}] = k[(u_{jt} - \hat{u}_{jt}) - (s_{nt}^g - \hat{g}_{nt})], \quad (30)$$

so the coefficients on the performance and benchmark surprises are equal and opposite: $K_g = -K_u$ and $\psi_g = 1$. ■

C Moment Construction and Estimation Details

This appendix defines each targeted moment, the auxiliary regressions and samples from which they are computed, the weight matrix, the simulation and numerical routine, and the standard-error formulas. Every moment is computed by the same code on the data and on each simulated panel, so any approximation in a moment’s construction is shared between the two and nets out of the comparison.

Inputs and timing. Performance enters the moments in industry-adjusted form: profitability is operating ROA net of the industry-year universe mean, and returns are market-adjusted returns net of their industry-year mean, as described in Appendix A. In the hiring and firing regressions (19)–(20), the information set is dated relative to the decision year t : recent performance R^{rec} is the firm-specific market-adjusted return in year $t - 1$, older performance R^{old} is the average of years $t - 2$ and $t - 3$, and the industry return R^{ind} is the industry-year mean of market-adjusted returns in year $t - 1$. All three covariates are standardized to unit variance within the estimation sample, so the targeted coefficients are responses to a one-standard-deviation movement.

Profitability dynamics and variances (8 moments). The pooled autocorrelations are the two slope coefficients from a regression of operating ROA on its first and second lags with a constant only; we deliberately omit fixed effects here to avoid the short-panel bias they would induce, and instead let the permanent firm component enter the moment on both the data and the simulation side. Mean ROA is the sample mean. The variance of the firm effect is the variance across firms of firm-average ROA. The within-spell variance is the variance of industry-adjusted profitability around CEO-spell means. The one-year and two-year change variances are the variances of $y_{jt}^{fs} - y_{jt-1}^{fs}$ and $y_{jt}^{fs} - y_{jt-2}^{fs}$ within CEO spells; differencing removes the permanent firm component, and the pair of horizons separates transitory noise from persistent movements in quality. The between-CEO variance is the variance across CEOs of spell-average industry-adjusted profitability net of the firm’s average, pooled across types. Because the mean of a short spell is a noisy estimate of that spell’s productivity, the raw variance overstates the dispersion of CEO quality; we remove each spell’s sampling variance, its own within-spell variance divided by its length, accounting for the firm demeaning. Using each spell’s own variance matters because volatile firms dismiss more often, so short spells and noisy spells coincide.

Insider premium profile (11 moments). The premium regression (18) is estimated on all CEO-years with firm and year fixed effects. The targeted coefficients are the six insider-by-

tenure-bin interactions (bins 1, 2, 3, 4–6, 7–9, 10+), the coefficient on the turnover indicator (the succession dip), and the four baseline tenure-bin main effects for bins 3 and later, which pin the tenure profile of performance under outsider CEOs; the turnover indicator absorbs the first year of a spell, so the profile is measured relative to the second year.

Hiring and firing margins (8 moments). The hiring regression (20) is estimated on succession years only: the dependent variable is an indicator that the incoming CEO is an insider, the covariates are the departing regime’s performance measures defined above, and the regression includes firm and year fixed effects. The targeted moments are the three performance coefficients and the unconditional insider share of successions. The firing regression (19) is estimated on all CEO-years with firm and year fixed effects; the dependent variable is the forced-turnover indicator. The targeted moments are the three performance coefficients and the unconditional forced-turnover rate per firm-year.

Dismissal timing (10 moments). The dismissal-timing regression (21) is estimated on all CEO-years with spell and year fixed effects. The dependent variable is the forced-turnover indicator, and the covariates are indicators for tenure years 2, 3, 4–6, 7–9 and 10 and above, each on its own and interacted with the insider indicator; the first year of a spell is the base, and the insider main effect is absorbed by the spell fixed effects. The spell fixed effects absorb every difference in dismissal rates across spells, so these coefficients describe when within a spell dismissal occurs, while the forced-turnover rate pins how often.

Industry factor (3 moments). Let $\bar{y}_{Jt}$ denote the industry-year universe mean of operating ROA. After removing industry means, the lag-one and lag-two autoregression coefficients of $\bar{y}_{Jt}$ and its variance are targeted. In the data, $\bar{y}_{Jt}$ blends the industry factor with the averaged firm-level components; the same blend arises in the simulated panel, where the moment is computed from the simulated industry means in exactly the same way, so the contamination is common to both sides and the moments remain informative about (ϕ_g, σ_ζ) jointly with ρ .

Weighting, simulation, and the objective. The weight matrix is diagonal, with each moment weighted by the inverse of its variance; we use the diagonal rather than the full inverse covariance matrix to avoid the small-sample bias of fully optimal weighting (Altonji and Segal, 1996). The simulated moments are averaged over $S = 20$ simulated economies.

Standard errors. Standard errors account for simulation error and for the covariance across moments. Writing G for the Jacobian of the simulated moments in the parameters

(computed by central finite differences under common random numbers), W for the diagonal weight matrix used in estimation, and Σ for the covariance matrix of the empirical moments, the estimator’s asymptotic variance is the sandwich:

$$\widehat{\text{avar}}(\hat{\theta}) = \left(1 + \frac{1}{S}\right) (G'WG)^{-1} G'W \Sigma WG (G'WG)^{-1} \quad (31)$$

The factor $(1+1/S)$ accounts for simulation noise. Σ is estimated from the influence functions of the moments, clustered by firm, with the three industry-factor moments clustered by industry; it enters only the standard errors, the estimation weight being the diagonal matrix above. The retirement hazard is estimated in a first stage, so the simulated moments depend on its coefficients $\hat{\eta}$ as well as on θ . We account for this by replacing each firm’s influence contribution to the moment condition with $\psi_{m,i} - G_{\eta} \psi_{\eta,i}$, where G_{η} is the Jacobian of the simulated moments in the hazard coefficients and $\psi_{\eta,i}$ is the firm’s influence contribution to the first-stage logit estimate. This captures both the sampling variation of $\hat{\eta}$ and its covariance with the empirical moments. The adjustment changes the standard errors in Table 5 by less than one percent. Standard errors for derived quantities (σ_{θ} , λ , ψ_I and $\psi_I - \psi_O$) follow from the delta method applied to the full covariance matrix of the estimates.

D First Stage: Voluntary Turnover

Table 10: First Stage: Voluntary Separation Hazard

Coefficients		Implied annual hazard $\delta(\tau)$	
ω_0 (constant)	−3.902	$\tau = 1$	0.025
ω_1 (tenure)	0.259	$\tau = 5$	0.053
ω_2 (tenure ²)	−0.0116	$\tau = 10$	0.089
ω_3 (tenure ³)	1.43×10^{-4}	$\tau = 15$	0.105
		$\tau = 20$	0.099
		$\tau = 30$	0.064

Notes: Maximum-likelihood estimates of the logistic hazard (32) on the CEO-year panel. The right panel evaluates the estimated hazard at selected tenures.

Voluntary separation is modeled as an exogenous hazard in CEO tenure, estimated in a first stage and held fixed in the structural estimation. In each year an incumbent of tenure τ departs voluntarily with probability

$$\delta(\tau) = \frac{1}{1 + \exp(-(\omega_0 + \omega_1\tau + \omega_2\tau^2 + \omega_3\tau^3))}, \quad (32)$$

and the same hazard applies in the model’s solution and in the simulated panels. Treating this margin as exogenous reflects its role in the model: the board’s decision problem concerns dismissal and appointment, and voluntary departures (retirements and outside opportunities) arrive independently of the board’s beliefs.

In the data, a voluntary exit is a CEO-year that is the last observed year of the spell, at a firm that remains in the panel afterward, and that carries no dismissal flag; requiring the firm to continue distinguishes a CEO departure from the firm itself exiting the sample, and the dismissal flag routes forced exits to the model’s endogenous firing margin. We estimate (32) by maximum likelihood on the CEO-year panel. Parameter estimates and the implied hazard profile are reported in Table 10. The hazard rises from about 2.5% in a CEO’s first year to a peak of roughly 10% around fifteen years of tenure and declines thereafter, consistent with retirement timing concentrating in the second decade of a spell and with very long-tenured CEOs (disproportionately founders) being unlikely to leave.

E Sensitivity of Parameter Estimates to Moments

This appendix quantifies the dependence of the parameter estimates on the targeted moments using the sensitivity measure of Andrews et al. (2017). Let G denote the Jacobian of the simulated moments with respect to the parameters and W the weight matrix, both evaluated at the estimated parameter vector. The sensitivity matrix is:

$$\Lambda = -(G'WG)^{-1}G'W \quad (33)$$

Entry Λ_{pm} measures the local effect of a perturbation of moment m on the estimate of parameter p : it answers the question of which features of the data are responsible for each estimate. To place entries on a common scale, we standardize by the precision of the moments and of the estimates, reporting $\tilde{\Lambda}_{pm} = \Lambda_{pm} s_m / s_p$, where s_m is the standard error of moment m and s_p the standard error of parameter p . A standardized entry of 0.5 thus indicates that a one-standard-error perturbation of the moment moves the parameter estimate by half of its standard error.

Table 11 reports, for each parameter, the three most influential moments and their standardized sensitivities. The patterns line up with the identification discussion in Section 4. The dispersion of the permanent firm component is determined almost entirely by the across-firm variance of profitability, and its mean by average profitability. The idiosyncratic shock volatility σ_ε draws on the three within-spell variance moments, the variances of profitability and of its one- and two-year changes, while firm persistence ρ combines the pooled auto-

Table 11: Sensitivity of parameter estimates to targeted moments

Parameter	First		Second		Third	
	Moment	$\tilde{\Lambda}$	Moment	$\tilde{\Lambda}$	Moment	$\tilde{\Lambda}$
$\bar{\iota}$	Mean profitability	-0.44	Timing, yrs 10+	-0.29	Premium, ten. 2	+0.22
σ_{ι}	Var. firm effect	-0.94	Profit. autocorr. (2)	-0.13	Profit. autocorr. (1)	-0.12
ρ	Profit. autocorr. (1)	-0.41	Profit. autocorr. (2)	+0.41	Var. 1-yr change	+0.29
σ_{ε}	Within-spell var.	-0.35	Var. 2-yr change	-0.33	Var. 1-yr change	-0.28
ϕ	Firing on older ret.	+0.33	Premium, ten. 1	-0.28	Timing, yrs 7-9	+0.28
σ_{η}	Timing, yrs 7-9	-0.42	Ins. timing, yrs 7-9	+0.37	Timing, yrs 4-6	+0.28
ψ_g	Firing on industry ret.	-0.81	Var. industry mean	-0.24	Hiring on industry ret.	+0.20
κ	Hiring on older ret.	-0.38	Premium, ten. 1	-0.36	Hiring on industry ret.	-0.33
μ_I	Timing, yr 2	-0.37	Ins. timing, yr 2	+0.29	Insider share	-0.23
c	Succession dip	+0.85	Firing on recent ret.	-0.23	Premium, ten. 1	+0.22
π	Timing, yrs 10+	+0.40	Between-CEO var.	-0.25	Ins. timing, yrs 10+	-0.23
$\sigma_{\nu,I}$	Timing, yr 2	-0.35	Premium, ten. 2	+0.32	Firing on recent ret.	+0.27
ϕ_g	Ind. autocorr. (1)	-0.45	Ind. autocorr. (2)	-0.43	Profit. autocorr. (2)	-0.20
σ_{ζ}	Var. industry mean	-0.94	Ind. autocorr. (1)	+0.38	Ind. autocorr. (2)	+0.38
γ	Timing, yrs 7-9	-0.37	Ins. timing, yrs 7-9	+0.33	Timing, yr 3	+0.29
ψ_O	Timing, yrs 4-6	+0.49	Timing, yrs 10+	+0.29	Ins. timing, yr 2	+0.28

Standardized sensitivities $\tilde{\Lambda}_{pm} = \Lambda_{pm} s_m / s_p$ from (33), evaluated at the estimated parameter vector with the estimation weight matrix: the movement of parameter p , in units of its standard error, induced by a one-standard-error perturbation of moment m . For each parameter the table reports the three largest entries in absolute value. Moment labels: “Profit. autocorr.” are the pooled autocorrelations of industry-adjusted profitability; “Var. 1-yr (2-yr) change” the within-spell variances of one- and two-year profitability changes; “Var. firm effect” the across-firm variance of average profitability; “Ind. autocorr.” and “Var. industry mean” the autocorrelations and variance of the industry benchmark; “Premium, ten. b ” the insider premium coefficient for tenure bin b ; “Firing (Hiring) on recent/older/industry ret.” the coefficients of the firing (hiring) regressions on recent, longer-lagged, and industry returns.

correlations with the one-year change variance. The industry parameters are pinned by the industry block, the autocorrelations of the industry benchmark for ϕ_g and its variance for σ_{ζ} , and the monetary turnover cost by the succession dip. The industry attribution parameter ψ_g is dominated by the response of firing to industry returns, the moment that directly measures how far industry conditions enter personnel decisions. Continuity κ loads on the dependence of hiring on the predecessor’s record and on the entry premium, its two distinguishing implications, and the persistence of productivity ϕ on the response of firing to older performance together with the entry premium. The information and wedge parameters are identified mainly by the dismissal-timing block. The precision of the outsider signal ψ_O loads on the tenure profile of outsider dismissals, most strongly in years four to six; the insider signal noise $\sigma_{\nu,I}$ on the year-two timing coefficient and the second-year premium; and the type mean μ_I on the year-two coefficients and the insider share. The two wedges respond

to the later part of the profile, the dismissal cost π to dismissals at ten or more years of tenure and the outsider hiring cost γ to the coefficients for years seven to nine, as does the productivity innovation σ_η , since how quickly dismissal risk builds with tenure depends on how fast productivity, and beliefs about it, move.

Because the parameters are estimated jointly, the sensitivity matrix also reflects the covariation of estimates through $(G'WG)^{-1}$: a moment can influence a parameter it does not directly inform when the two are connected through a third parameter. The appearance of the between-CEO variance among the sensitivities of the dismissal cost is of this kind: the dispersion of productivity sets the stakes of a dismissal. Similarly, the two wedges draw on the timing block as a whole rather than on any single coefficient, reflecting that they shift the dismissal profile at every tenure jointly with the information parameters.

F Non-Stationary Learning

Section 3 holds the board’s posterior covariance at its stationary value at every tenure. Here we relax that restriction and re-estimate, letting the posterior variance depend on elapsed tenure and on the type of the executive appointed. Insiders then enter better informed than outsiders, with an entry variance of 0.46 of σ_θ^2 against 0.86 at the estimates, and the two paths converge as tenure accumulates. The filter introduces no additional parameters.

Table 12 reports the estimates. No parameter moves by more than two standard errors of the difference. The largest movements are in the precision of the board’s signal about outsiders, which rises from 0.06 to 0.14, and in the outsider hiring cost, which rises from 0.060 to 0.082; the information advantage at hiring narrows from 0.54 to 0.37 and remains more than four standard errors from zero. The fit is worse under the non-stationary filter: the criterion rises from 40.7 to 65.0, and the shortfall is concentrated in the tenure profile of the insider premium, which the non-stationary model places too low at every tenure.

Two quantities become functions of tenure rather than constants. The share of a candidate’s uncertainty the board has resolved is 0.54 for an insider and 0.14 for an outsider at appointment, 0.57 and 0.25 after one year of performance, and 0.67 for both at long tenure, against 0.72 at every tenure under the stationary filter. Industry filtering ψ_g rises from about 0.60 in an executive’s first year to 0.71 at long tenure, because a board still uncertain about a new CEO attributes more of a common industry movement to that executive; the values reported in Section 5 are the stationary limits. Given that the stationary filter fits better and that no parameter changes materially, we retain it in the baseline specification.

Table 12: Estimates under Stationary and Non-Stationary Learning

Parameter	Description	Baseline	Non-stationary	t
<i>CEO productivity and succession</i>				
ϕ	Productivity persistence	0.987 (0.007)	0.987 (0.006)	0.0
σ_η	Productivity-innovation s.d.	0.008 (0.001)	0.008 (0.001)	-0.4
σ_θ	Implied stationary dispersion	0.051 (0.009)	0.047 (0.009)	-0.3
μ_I	Insider productivity premium	-0.005 (0.002)	-0.004 (0.003)	0.1
κ	Continuity (as fraction of ϕ)	0.491 (0.082)	0.426 (0.075)	-0.6
$\lambda = \kappa\phi$	Implied inheritance	0.484 (0.082)	0.420 (0.075)	-0.6
<i>Board information</i>				
ψ_I	Insider signal precision	0.597 (0.129)	0.516 (0.091)	-0.5
$\sigma_{\nu,I}$	Implied signal noise	0.036	0.042	
ψ_O	Outsider signal precision	0.061 (0.015)	0.143 (0.052)	1.5
$\sigma_{\nu,O}$	Implied signal noise	0.199	0.116	
$\psi_I - \psi_O$	Information advantage at hiring	0.536 (0.127)	0.373 (0.078)	-1.1
ψ_g	Industry-performance filtering	0.756 (0.126)	0.713 (0.201)	-0.2
<i>Firm profitability</i>				
$\bar{\iota}$	Firm-effect mean	0.095 (0.004)	0.096 (0.005)	0.3
σ_ι	Firm-effect s.d.	0.098 (0.004)	0.099 (0.004)	0.3
ρ	Operational persistence	0.423 (0.041)	0.479 (0.043)	1.0
σ_ε	Transitory shock s.d.	0.061 (0.002)	0.061 (0.002)	0.0
ϕ_g	Industry-factor persistence	0.509 (0.094)	0.415 (0.108)	-0.7
σ_ζ	Industry-factor innovation s.d.	0.022 (0.002)	0.020 (0.003)	-0.5
c	Monetary turnover cost	0.0023 (0.0026)	0.0010 (0.0025)	-0.4
<i>Board preferences</i>				
π	Non-pecuniary firing cost	0.108 (0.011)	0.120 (0.017)	0.6
γ	Outsider hiring cost	0.060 (0.008)	0.082 (0.020)	1.0
Objective		40.74	65.02	

Notes: Standard errors in parentheses. Both columns are estimated on the same 40 moments with the same weighting matrix. The baseline column reproduces Table 5. The last column is the difference between the two estimates divided by the square root of the sum of their squared standard errors; it treats the two estimates as independent, although both are computed from the same sample, and so understates the precision of the difference.