Contract Horizon, Severance Pay, and Turnover

Vladimir Vladimirov*

June 22, 2018

Abstract

Firm executives are often hired with renewable fixed-term contracts. This paper asks why and what determines the length of such contracts. The paper develops a model in which the fit between a firm and its managers changes over time. Stipulating severance pay for premature dismissal in a given period mitigates managers' incentives to conceal a deteriorating fit but increases these incentives in preceding periods. By optimally choosing the length of renewable fixed-term contracts, boards can manage the use and effectiveness of severance pay. The relation between contract length, severance pay, and managers' outside options helps explain several puzzling stylized facts.

Keywords: contract length, contract horizon, severance pay, renewable fixed-term contracts, voluntary and forced turnover, asymmetric information.

JEL Classification: G30, G34, D82

---

*University of Amsterdam. E-mail: vladimirov@uva.nl. I thank Patrick Bolton, Arnoud Boot, Murillo Campello, Tolga Caskurlu, Jason Donaldson, Ruediger Fahlenbrach, Daniel Garrett, Stuart Gillan, Juan Pedro Gomez, Moqi Groen-Xu, Sebastian Gryglewicz, Michel Habib, Marina Halac, Roman Inderst, Dirk Jenter, Boyan Jovanovic, Oguzhan Karakas, Nicolas Klein, Tomislav Ladika, Chen Lin, Alexander Ljungqvist, Richard Lowery, Andrey Malenko, Frédéric Malherbe, David Martinez-Miera, Konstantin Milbradt, Martin Oehmke, Per Östberg, Daniel Paravisini, Enrico Perotti, Florian Peters, Sebastian Pfeil, Giorgia Piacentino, Joshua Pierce, Uday Rajan, Adriano Rampini, Merih Sevilir, Fenghua Song, Jan Starmans, Spyros Terovitis, Ed Van Wesep, David Yermack, John Zhu, as well as conference and seminar participants at the American Finance Association annual meeting (Philadelphia), European Economic Association annual meeting (Geneva), European Winter Finance Summit, Executive Compensation Conference (Rotterdam), University of Amsterdam, and University of Zurich for constructive feedback.

1 Introduction

A first-order problem for boards evaluating managers is that even high-quality managers may not always be a good fit for the organization. Consider the example of Ronald Boire. He was deemed to be an "ideal chief executive for Barnes & Noble," but he had to step down a year after being hired because he "was not a good fit."[1] Fit reflects the complex match between the manager's collection of skills and the organization's assets and growth options (Lazear, 2009). However, even a good fit could change over time due to a changing environment, inadequate firm-specific human capital investments, family or health issues, or conflicts with colleagues or subordinates. A key problem is that managers can withhold information from boards or can engage in actions, such as creative accounting, that help them appear a good fit even if that hurts the firm's subsequent profitability.

This paper argues that boards can address this problem by optimally choosing the length of contracts and the severance pay they offer executives for premature dismissals. In practice, 46% of S&P 500 firms and 68% of S&P 1500 firms hire managers with explicit contracts that typically cover a fixed period, allow for renewal, and foresee severance pay upon premature termination (Gillan et al., 2009; Rau and Xu, 2013). Yet, despite the prevalence of such contracts, many of their features do not seem well-understood. Barnes & Noble had offered Mr. Boire a three-year contract that could be renewed for another two years. However, given that fixed-term contracts can be terminated at any time, existing research offers little insight into why, for example, offering a five-year contract straight away would not be better and, more fundamentally, why the length of contracts matters in the first place.

The contribution of this paper is to rationalize the use and length of renewable fixed-term contracts by liking them to the use of severance pay in employment relationships. The benefit of severance pay (\$10.5m in Mr. Boire's case) is that it can mitigate managers' desire to withhold information or take actions to conceal a deteriorating fit (Almazan and Suarez, 2003; Inderst and Mueller, 2010). This reduces the board's need to rely on the firm's noisy performance when evaluating that fit and making dismissal decisions. However, stipulating severance pay for premature termination in *future* periods increases the incentives to conceal a bad fit and, thus, the severance pay needed to prevent that *today*. This problem is especially acute for longer contracts. Thus, contract length, severance pay, and the sensitivity of executive turnover to performance will be closely intertwined. The model's implications can shed light on various stylized facts that have been difficult to reconcile with prior theory, such as why CEOs are dismissed more often in industry-wide bad times, appearing to be punished for bad luck beyond their control (Jenter and Kanaan, 2015).

[1] "Barnes & Noble Says CEO Boire 'Not a Good Fit' and Will Step Down," WSJ, 16 August 2016.

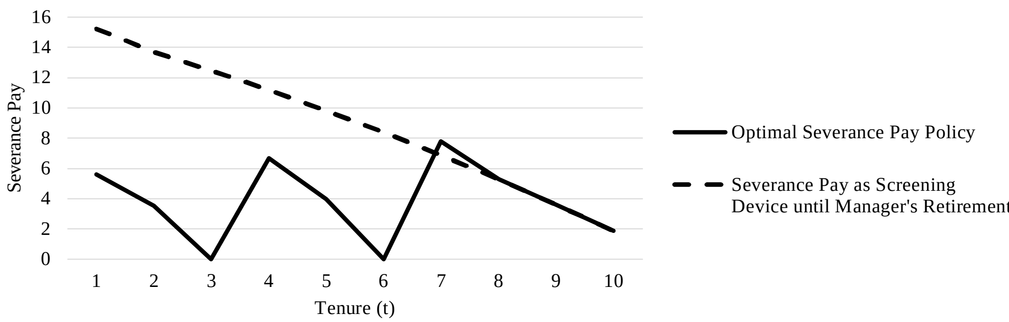

Figure 1: Relation between severance pay for premature dismissal and the length of contracts stipulating such pay. The optimal long-term contract (solid line) allows the board to dismiss managers without severance pay in some periods. It can be implemented with renewable fixed-term contracts whose renewal coincides with such periods. Allowing for renewal dates reduces the severance pay for premature dismissal in all preceding periods.

The paper develops a model in which a board repeatedly appoints finitely-lived managers. Once hired, a manager can increase the likelihood of being of good match quality (henceforth, fit) for the organization by investing in firm-specific human capital. Being a good fit makes it not only more likely to achieve high cash flows, but also to be a good fit next period. However, only the manager observes the extent of her firm-specific human capital investments and how they impact her fit. The paper's main results arise from investigating the extent to which the board should rely on adequate severance pay that incentivizes managers not to conceal a deteriorating fit by, e.g., withholding information or taking actions detrimental for future profitability. An important, but natural, restriction on contracting is that the board cannot commit to retain the manager if its beliefs indicate that her fit has deteriorated.[2]

The first main result (illustrated in Figure 1) is that it is optimal for the board to offer managers a contract that allows the board to terminate that contract in some periods without paying severance. In such periods, managers are more likely to withhold information or act to conceal a deteriorating fit, forcing the board to rely more heavily on the firm's noisy performance. An off-the-shelf implementation of such a contract can be achieved with renewable fixed-term contracts of the type offered to Mr. Boire. Such contracts allow for dismissal at any time, but dismissal is costless on renewal dates. If not terminated on such dates, the contract continues as originally agreed upon. The length of the fixed-term contracts corresponds then to the time between renewal dates at which the board has the option to dismiss the manager without severance pay.

The optimality of renewable fixed-term contracts derives from the trade-off between of-

[2]This is key, as with such limited commitment, it is often suboptimal to screen an agent's private information (Hart and Tirole, 1988; Laffont and Tirole, 1990; Malcomson, 2016). Note that none of the following arguments requires that the manager agrees that someone else might be a better fit for the job.

fering severance pay over multiple periods and reducing the manager's ability to extract rent. Specifically, the advantage of screening managers by offering severance pay is that it helps the board to make more-informed replacement decisions. This is beneficial, as a manager who is a good fit today is also more likely to be a good fit tomorrow, and vice versa. However, severance pay needs to compensate managers for forgoing future wages, typically above their outside options, paid for their investment in firm-specific human capital. This leads to two crucial disadvantages in a dynamic setting. First, severance pay reduces the benefit of deferring compensation (otherwise key in dynamic models). Since the board may retain a manager even after low cash flow realizations, attributing them to bad luck, deferring pay makes managers more willing to conceal a deteriorating fit. Second, offering severance pay in a given period increases the manager's incentives to conceal a deteriorating fit in preceding periods. From the perspective of such periods, staying longer with the firm would pay off even if she is a bad fit. Hence, when a contract offers severance pay over multiple periods, the severance pay must increase with the remaining number of such periods (dashed line in Figure 1). This increase leads to a stark mismatch between the costs and benefits of offering severance pay. In particular, if a manager is replaced in her second year, the benefits of having a contract that would have screened out managers with a deteriorating fit between year two and year ten are never realized. However, the manager's severance pay in year two still needs to compensate her for forgoing wages and severance pay until year ten.

In light of this cost-benefit mismatch, the advantage of having the option to fire the manager without severance pay (implemented with renewal dates) is twofold: It allows for cheap dismissal and, since the latter reduces the manager's continuation payoff, it reduces the necessary severance pay in all preceding periods. The trade-off is that not offering severance pay increases the risk that the manager may withhold information or take actions to conceal her fit. This forces the board to rely more heavily on the firm's performance as a noisy indicator of the manager's fit and increases the risk of incorrect replacement and retention decisions.[3] This trade-off is best resolved by including the option to fire the manager without severance pay at regular intervals. As illustrated by the solid line in Figure 1, this strategy significantly reduces the need for severance pay, while exposing the board to the relatively low risk of making less-efficient replacement decisions in years three and six, which the manager reaches with relatively low probability.

The paper's main implications arise from studying the determinants of the length of renewable contracts, i.e., the time between periods in which the board may fire the manager

[3]In line with the model, the evidence is that ex ante severance agreements are associated with more-truthful managers (Rau and Xu, 2013; Brown, 2015), and CEOs with longer remaining tenure have higher severance agreements (Rau and Xu, 2013). Furthermore, the sensitivity of CEO turnover to performance spikes close to renewal dates (Cziraki and Groen-Xu, 2015).

without severance pay following underperformance. A key factor is the manager's outside option conditional on leaving the firm. Low-paying alternative employment opportunities increase the manager's reluctance to reveal information that may lead to her dismissal and, thus, would require offering higher severance pay. Avoiding this cost means relying more heavily on firm performance to evaluate the manager, which goes hand in hand with offering shorter contracts. Hence, Jenter and Kanaan's (2015) findings that CEO dismissals following underperformance are more likely in industry-wide downturns (when outside options are low) could be due to an optimal evaluation and replacement policy rather than to a lack of relative performance evaluation. Naturally, firms that stand to lose more from not having the right manager in charge, such as firms with better growth options, would benefit more from incentivizing the manager not to conceal her fit. To reduce their dependence on noisy performance, these firms would rely more on severance pay and would offer longer contracts.

Though renegotiations to offer severance pay at renewal dates can be avoided, if they occur, they may restrict the board to offer contracts that look like the dashed line in Figure 1. In this case, the manager's age becomes important, as it offers a natural commitment to a maximum contract horizon. Then, the board will offer higher severance packages to younger managers, since such manager have more to lose from early termination (in line with the findings of Rau and Xu, 2013). Furthermore, the board will hire older managers if the managers' outside options are low (as in downturns) since, then, the severance pay needed to incentivize the manager not to conceal her fit will be higher.

The paper's main contribution is that it rationalizes the use of renewable fixed-term contracts and analyzes key determinants of the length of such contracts. Building on prior static models rationalizing severance pay as an incentive and screening device (Levitt and Snyder, 1997; Inderst and Mueller, 2010; Almazan and Suarez, 2003; Van Wesep, 2010), the results add to our understanding of the dynamic trade-offs and implications of offering severance pay to induce managers not to conceal a deteriorating fit. The insights regarding contract length and the performance sensitivity of turnover further add to He (2012) and Van Wesep and Wang (2014) who have also noted that outside options may affect severance pay. Expanding on this insight is important, as it highlights that the pervasive view that higher outside options make it more expensive to employ managers may not always apply, especially when incentive pay requires managers to be paid above their outside option.[4] More broadly, this would imply that the board might appear to choose from a "select club" of managers with better outside options (even if they are not better), as they would be less desperate

[4]The outside option of an incumbent is typically not another CEO position, and CEOs typically take a big pay cut in their new employment following dismissal (Fee and Hadlock, 2004; Nielsen, 2017). Indeed, in most models with unobservable effort, agents are paid above their outside option.

to conceal a deteriorating fit. For the same reason, the board may tolerate investments in general human capital, even if they come at the expense of firm-specific human capital.

The paper's novel implications for contract horizon and the dynamic use and structure of severance agreements also distinguish it from models in which the board learns the managers' quality from firm performance over time (Hermalin and Weisbach, 1998; Taylor, 2010). Particularly related are Jenter and Lewellen (2017) and Garrett and Pavan (2012). Both papers consider dynamically changing types but take polar opposite approaches. In Jenter and Lewellen (2017), the board does not screen managers and, thus, must rely on the firm's most recent performance to infer their productivity. By contrast, Garrett and Pavan (2012) analyze full-commitment contracts that always incentivize managers to report their private information. In the present paper, the decision of how to evaluate the manager can be seen as optimally relying on both approaches, while relaxing the assumption that the board can commit to (sometimes) retaining managers with a deteriorating fit. The resulting suboptimality of full revelation has been analyzed also in the literature on relational contracts (Halac, 2012; Malcomson, 2016). However, this literature focuses on "at will" employment relationships lacking an explicit contract, while in the current paper contract horizon and contractual incentives are the main focus.[5]

Also related are Anderson et al. (2018) and Eisfeldt and Kuhnen (2013) in which a shock that decreases industry returns prompts the firm to look for a manager who is better suited to the new environment. This provides one explanation for Jenter and Kanaan's (2015) findings that turnover is more likely in industry-wide bad times. Instead, in the present paper, managers are more likely to be fired following underperformance, as boards rely less on severance pay and more on performance when screening managers. This could help explain why Fee et al. (2015) find no evidence for a lack of relative performance evaluation, when also considering turnover that seems "voluntary" and that may have been eased by a severance payment. In Eisfeldt and Rampini (2008), CEO turnover is procyclical. However, managers in their model live for only one period, which does not allow for an analysis of contract horizon. The literature analyzes turnover also as a threat to discipline managers (Sannikov, 2008) and reduce myopia (Varas, 2017) in which case severance pay helps prevent shirking when agents can save (He, 2012). The key difference in the present paper is that the aim of turnover is to appoint a better manager, which raises the question of whether severance pay should be offered to screen managers whose fit has deteriorated.[6]

[5] A renewable fixed-term contract is an implementation of a long-term contract with costless termination options. Though the literature comparing long- with short-term contracts (e.g., Hart and Tirole, 1988) and emphasizing the benefits of laxer control (Aghion and Tirole, 1997; Crémer, 1995) is related, it neither discusses severance pay nor explains the use and the determinants of the length of renewable contracts.

[6] The paper also contributes to prior work on human capital investments (Jovanovic, 1979 a,b; and Felli and

2 Model

Consider an infinitely lived firm in which the board maximizes shareholder wealth and is in charge of hiring and replacing the firm's manager (she). The firm operates in an economy, in which every period t consists of three dates. At the first date of every period, $\tau_t = 0$, an incumbent manager can invest in firm-specific human capital. Such an investment carries a non-monetary cost c, but it increases the likelihood that the manager's fit with the firm in the current period, $\theta_t \in \{\theta_G, \theta_N\}$, is good. Specifically, if the manager invests in firm-specific human capital, her fit is θ_G with probability e_t. With probability $1 - e_t$ or, respectively, if she does not invest in firm-specific human capital, her fit is $\theta_N < \theta_G$.

At the interim date $\tau_t = 1$, the manager learns her fit and can report it, and the board can decide whether or not to replace her with a new manager. All cash flows from the period are realized at the final date $\tau_t = 2$. If the board has not already replaced the manager at the interim date, it can choose again whether or not to keep her for the next period. Cash flows are verifiable and can take values $x_t \in \{x, x + \Delta x\}$. The manager's fit $0 \leq \theta_t \leq 1$ corresponds to the likelihood of achieving the higher cash flow $x + \Delta x$, where $x, \Delta x \geq 0$. All parties are risk neutral, and the common discount factor between two neighboring periods is $\delta \in (0, 1)$.

Neither the board nor potential managers have private information when a new manager is hired. Furthermore, the managers from which the board can choose have zero wealth and are identical in all respects except for their age, i.e., managers are not infinitely-lived and leave the labor market once they reach their retirement age. However, the key assumptions are that a manager's investments in firm-specific human capital, as well as the realizations of θ_t at the interim date $\tau_t = 1$ of every period, are known only to the manager.

The probability that a manager's investment in firm-specific human capital results in a good fit in period t depends on her fit θ_{t-1} in the previous period. Specifically, there is a positive correlation with $e_t(\theta_G) > e_1 > e_t(\theta_N)$, where e_1 is the likelihood of θ_G in the manager's first (complete) period after being hired, and $e_t(\theta_{t-1})$ makes the dependence on θ_{t-1} explicit. One interpretation of the negative fit persistence $e_t(\theta_N) < e_1$ is as the negative consequences of myopic actions taken by the manager to conceal a deteriorating fit in t, which make it even less likely that the firm is subsequently profitable under her management. In this Markov environment, the t-subscripts in $e_t(\theta_G)$ and $e_t(\theta_N)$ are not necessary, but they are helpful for keeping track of the intertemporal forces affecting contracting. Initially, $\{e_1, e_t(\theta_G), e_t(\theta_N)\}$ and the manager's outside option, which pays $\overline{U}$ per period, are fixed,

Harris, 1996) by analyzing a setting in which a worker's fit changes over time and is her private information. Tenure limits also reduce agents' ability to extract rents in Lazear (1979), Prescott and Townsend (2006), and Hertzberg et al. (2010). Managers leave also when taking better outside options (Wang, 2011, 2015) but Section 3.3 relaxes these assumptions.

Contracting At the beginning of the employment relationship, the board offers the manager a contract $\mathbf{w} = (w_t, \Delta w_t, w_{s,t}, \psi_t^1, \psi_t^2)_{t=1}^T$. The contract components characterizing any given period t can depend on the current and past cash flow realizations $(x_i)_{i=1}^t$, as well as on the manager's reports $(\widehat{\theta}_i)_{i=0}^t$ about her fit (if she makes such reports). Unless this leads to confusion, the history dependence is not made explicit but is captured only by the subscript t. In this contract, w_t stands for the manager's wage in the low cash flow state, and Δw_t stands for how much she receives in addition (i.e., her "bonus") in the case of a high cash flow realization; $w_{s,t}$ is the manager's severance pay if she is replaced at the interim date $\tau_t = 1$, *prior* to the cash flow realization $x_t \in \{x, x + \Delta x\}$ in that period, with ψ_t^1 being the probability of such an interim replacement; ψ_t^2 stands for the probability of replacing the manager at the end of the period, i.e., in $\tau_t = 2$, *after* the cash flow realization x_t.[7] While the contract does not explicitly consider a payment to the manager for leaving the firm at the beginning of a period before she obtains private information or a payment at the interim date $\tau_t = 1$ for staying with the firm, it is shown that such payments will not arise.

The manager is penniless and protected by limited liability, which requires that $w_t, w_{s,t} \geq 0$, and should have no incentives to destroy cash flows, i.e., $\Delta w_t \geq 0$ (Innes, 1990).[8] Contracts that satisfy these requirements are labeled as "feasible." Furthermore, it is assumed that the manager cannot be prevented from leaving the firm at any time during the employment relationship. Thus, the contract should at least compensate her for her outside employment opportunity, which would pay her $\overline{U}$ at the end of every period until her retirement in T. It is assumed that if the manager leaves the firm at the interim date $\tau_t = 1$ of a period, she still obtains $\overline{U}$ from her outside employment opportunity for that period.[9] Figure 2 summarizes the timing of events in each period.

Performance Evaluation and Replacement If the board replaces the incumbent at the interim date, the new manager is paid $\overline{U}$ to complete the period. In this period, her likelihood of success is $\overline{\theta}$ with $\theta_N < \overline{\theta} < \theta_G$; this requires no firm-specific human capital investment; and does not give rise to private information. Then, at the beginning of the following period, the board makes the manager an offer covering the whole potential relationship. If a manager is replaced, she is not rehired.

[7] The payment to the manager at the end of the period could also be reinterpreted as the manager's severance pay if she is fired at the end of that period.

[8] It is assumed that the manager does not save. Though all parties are risk neutral and use the same discount factor and cash flows are verifiable, the assumption is not innocuous. This is because a manager whose contract is terminated could otherwise offer the board a payment to be kept on the job.

[9] Assuming that the manager receives only a fraction of $\overline{U}$ leads to the same qualitative results.

$\tau_t = 0$	$\tau_t = 1$	$\tau_t = 2$
If new manager: board offers contract & manager accepts/rejects. Manager invests in firm-specific human capital or not – With investment: prob. of good fit θ_G is $e_t(\theta_{t-1})$ – With prob. $1 - e_t(\theta_{t-1})$ or if no investment: fit is θ_N.	Manager privately learns $\theta_t \in \{\theta_G, \theta_N\}$, can report $\widehat{\theta}$. Board replaces or keeps the manager: $\psi_t^1 \in \{keep, replace\}$ depending on $\widehat{\theta}_t$ and history. If $\psi_t^1 = replace$, manager paid $w_{s,t}$	$x_t \in \{x, x + \Delta x\}$ realized. Probability of $x + \Delta x$ is θ_t. Manager paid w_t or $w_t + \Delta w_t$ Board replaces or keeps manager: $\psi_t^2 \in \{keep, replace\}$ depending on $\widehat{\theta}_t$, x_t, history. If new manager hired at $\tau_t = 1$, the prob. of $x + \Delta x$ is $\overline{\theta}$ and new manager paid $\overline{U}$ in t.

Figure 2: Timing of events in period t.

Definition 1. *An equilibrium with optimal termination is a perfect Bayesian equilibrium in pure strategies in which replacement and retention decisions are efficient from the board's perspective given its information at this point in time. In the case of full revelation, these decisions should coincide with those of the subgame perfect equilibrium of the full information game at* $\tau_t = 1$.

Following Malcomson (2016), Definition 1 restricts attention to pure strategies and rules out the commitment not to renegotiate retention and replacement strategies that are suboptimal, given the board's ex post information. In a period in which the manager's fit is revealed at $\tau_t = 1$, the only replacement strategy consistent the board's ex post information is to replace a manager if and only if her fit is θ_N — i.e., $\psi_t^1 = 1$ if $\theta_t = \theta_N$; $\psi_t^1 = 0$ if $\theta_t = \theta_G$, in which case $\psi_t^2 = 0$ regardless of x_t (unless the period coincides with the manager's retirement age). This is because, by replacing a manager whose fit is θ_N, the board can increase the likelihood of high cash flows in both the current and the subsequent period, as $\overline{\theta} > \theta_N$ and $e_1 > e_{t+1}(\theta_N)$. By contrast, it would be ex post suboptimal to replace an incumbent who is a good fit, as $\theta_G > \overline{\theta}$ and $e_{t+1}(\theta_G) > e_1$.

If the board optimally retains a manager at the end of some period $t - 1$, it is clearly optimal to retain the manager also at the interim date τ_t of period t if the manager's fit is not revealed at that date. At the end of period t, the board uses then the firm's cash flows to decide whether to keep the manager also in the subsequent period. It is assumed that, if the cash flows in t are high, the board expects that the incumbent's likelihood of being a good fit in the following period will be higher than that of a new manager; the board assumes the opposite if the firm has produced low cash flows

$$\sum_{\theta_t \in \{\theta_B, \theta_G\}} \Pr(\theta_t | x)\, e_{t+1}(\theta_t) < e_1 < \sum_{\theta_t \in \{\theta_B, \theta_G\}} \Pr(\theta_t | x + \Delta x)\, e_{t+1}(\theta_t), \tag{1}$$

where $\Pr(\theta_t | x_t)$ is the board's posterior about the fit realization θ_t, conditional on the cash flow realization in t and the prior history. A sufficient condition for (1) to hold in terms of primitives is given in the Appendix (expression (B.15)). By Definition 1, this would imply that in periods in which the manager's fit is not revealed, she is retained if and only if the cash flows are high, i.e., $\psi_t^1 = 0$ and $\psi_t^2 = 0$ if $x_t = x + \Delta x$, but $\psi_t^2 = 1$ if $x_t = x$. Since the replacement strategies are deterministic, $\psi_t^\tau \in \{0, 1\}$, they are labeled for transparency as $\{keep, replace\}$.

Assumption (1) leads to a stark replacement policy that helps to focus on the novel economic forces, while abstracting from issues related to learning from the firm's cash flow history over an extended period how the manager's fit is changing. Section 3.3.3 discusses an extension of the model with continuous types, in which the board learns both from whether the contractually stipulated severance pay is more attractive than trying to avoid dismissal (the board learns whether the manager's fit is above a cutoff) and from the firm's cash flow performance in every period. In this extension, there can be more gradual changes to the manager's fit, and the board's decision to dismiss the manager following low cash flows will also depend on the history of cash flow realizations.

Finally, note that there is no truly voluntary turnover in this model. However, in practice, a smooth transition, eased by a severance package, might appear voluntary to outsiders, even if the board and the manager disagree behind the scenes as to whether a replacement could do a better job. Indeed, one could extend the model by assuming that the manager overestimates her fit or that her vision how to run the firm differs from the board's.[10]

[10] One could assume that the board believes that $\overline{\theta} > \theta_N$, while the manager believes that the board's preferred way has a success probability of only $\overline{\theta}^m < \theta_N$. Such disagreement has been motivated by heterogeneous priors and overconfidence (Goel and Thakor, 2008; Gervais et al., 2011; Huang et al., 2016). Interestingly, it has also been applied to explain short-term debt contracts (Zhu, 2018).

3 A Multi-Period Employment Relationship

Given a contract offer $\mathbf{w} = \left(w_t, \Delta w_t, w_{s,t}, \psi_t^1, \psi_t^2\right)_{t=1}^T$, let $\omega_t \in \{w_t, w_t + \Delta w_t, w_{s,t} + \overline{U}\}$ denote the firm's wage bill in period t.[11] The board's expected payoff in the first period is

$$V_1(\mathbf{w}) = \mathbb{E}\left[\sum_{i=1}^{T} \delta^{i-1}\left(q_i\left(x_i - \omega_i\right) + \widetilde{q}_i \delta V^*\right)\right], \tag{2}$$

where E is the expectation over the future θ_t and x_t realizations; and q_i and $\widetilde{q}_i$ are the endogenous probabilities that the incumbent manager is still with the firm in period i and, respectively, leaves the firm by the end of that period. V^* denotes the board's equilibrium expected payoff from hiring a new manager, starting from that manager's first complete period. Note that, since managers are ex ante identical, their information evolves independently, and time is infinite, the board's contracting problem when making an offer to a replacement manager is identical to that faced with her predecessor. Thus, in equilibrium, the board's expected payoff in (2) must be equal to V^*. For convenience, t is reset to one for every new manager, so that t could be interpreted as her tenure at the firm. The board's promise-keeping constraint implies that the manager's expected payoff in any given period t during her tenure is

$$U_t(\theta_{t-1}, \mathbf{w}) = \mathbb{E}\left[\sum_{i=t}^{T} \delta^{i-t}\left(\overline{U} + q_i\left(\omega_i - c_i - \overline{U}\right)\right) | \theta_{t-1}\right]. \tag{3}$$

Expression (3) states that the manager can obtain $\overline{U}$ in every period until she leaves the labor market in T, but she might receive something different from $\overline{U}$ while she is employed by the firm. What is crucial to the analysis is that the manager's fit persistence implies that her payoff in t, $U_t(\theta_{t-1}, \mathbf{w})$, depends on her fit realization in $t-1$ (because e_t depends on θ_{t-1}). There is no such prior realization when she is hired, so for period one, we write $U_1(\mathbf{w})$.

Using (3) to plug into (2), the board's objective when hiring a manager is to choose $\mathbf{w}$ to maximize

$$\max_{\mathbf{w}} \mathbb{E}\left[\sum_{i=1}^{T} \delta^{i-1}\left(q_i\left(x_i - c_i - \overline{U}\right) + \widetilde{q}_i \delta V^*\right)\right] - U_1(\mathbf{w}) + \sum_{i=1}^{T} \delta^{i-1}\overline{U}, \tag{4}$$

subject to the constraints that the contract $\mathbf{w}$ is feasible, incentive-compatible, and individ-

[11] Recall that if the board replaces a manager at $\tau_t = 1$, it pays $w_{s,t}$ to the departing manager and $\overline{U}$ to the replacement manager to complete the period.

ually rational for the manager in every period. Hence, the board trades off maximizing the surplus generated from employing a manager with minimizing the manager's rent

$$\nu_t(\theta_{t-1}, \mathbf{w}) := U_t(\theta_{t-1}, \mathbf{w}) - \sum_{j=t}^{T} \delta^{j-t} \overline{U}. \tag{5}$$

in the first period. We now state the relevant constraints.

Severance Pay as a Screening Device Suppose that the board seeks truthful reporting at the interim date of some period t. The manager truthfully reveals her fit in t and stays if it is θ_G or leaves with a severance package if it is θ_N if the following incentive constraints are satisfied

$$w_t + \theta_G \Delta w_t + \delta U_{t+1}(\theta_G, \mathbf{w}) \geq w_{s,t} + \sum_{j=t}^{T} \delta^{j-t} \overline{U} \tag{6}$$

$$w_{s,t} + \sum_{j=t}^{T} \delta^{j-t} \overline{U} \geq w_t + \theta_N \Delta w_t + \delta U_{t+1}(\theta_N, \mathbf{w}). \tag{7}$$

In terms of the practical interpretation of a manager's "reports," the question is whether the severance pay $w_{s,t}$ is high enough that the manager does not withhold information or take myopic actions, leading to $e_{t+1}(\theta_N) < e_1$, to appear a good fit in t.[12] Note that the manager's continuation payoff $U_{t+1}(\theta_t, \mathbf{w})$ can take on two values depending on her fit realization (θ_G or θ_N) in t, which means that both payoffs will play the role of state variables for characterizing the manager's contract.

To induce a manager to invest in firm-specific human capital in period t, the contract must further satisfy

$$\left(\begin{array}{c} e_t(\theta_{t-1})(w_t + \theta_G \Delta w_t + \delta U_{t+1}(\theta_G, \mathbf{w})) \\ + (1 - e_t(\theta_{t-1})) \left(w_{s,t} + \sum_{j=t}^{T} \delta^{j-t} \overline{U} \right) - c \end{array} \right) \geq w_{s,t} + \sum_{j=t}^{T} \delta^{j-t} \overline{U}, \tag{8}$$

where the right-hand-side of (8) captures that a manager who does not invest in firm-specific human capital does not have a good fit with certainty and, thus, is replaced at date $\tau_t = 1$

[12] Eliciting the manager's fit in t at the beginning of $t+1$ is suboptimal. It requires offering the manager the same information rent as when learning her fit in t, but without allowing for the benefit of replacing the manager earlier. Furthermore, note that the advantage of a Markov environment is that, if the manager truthfully reports θ_{t+1} on the equilibrium path in $t+1$, she does the same off the equilibrium path (i.e., after misreporting in t). Finally, note that, since the cash flow realization x_t carries no additional information about θ_t, it is without loss of generality not to condition the continuation payoff U_{t+1} on x_t. Such history dependence becomes relevant in Section 3.3.3.

of the period. If (8) is satisfied, the first incentive constraint (6) is lax.

Without offering incentives for firm-specific human capital investments, the board could satisfy (6) and (7), without leaving any rent to the manager by offering $w_t = \overline{U}$, $w_{s,t} = \Delta w_t = 0$. However, the incumbent would then add no value to the firm, as $\theta_N < \overline{\theta}$ and $e_{t+1}(\theta_N) < e_1$. Thus, the need to stimulate investment in firm-specific human capital leads to paying the manager an "efficiency wage" above her outside option (i.e., a positive rent).

Evaluating the Manager Based on Noisy Firm Performance Suppose, next, that the board does not offer incentives for truthful reporting (thus, accepting the risk that the manager withholds information or takes actions to conceal that her fit is θ_N) and relies on the firm's cash flows to evaluate the manager in period t. In such a period, the constraint that the manager prefers investing in firm-specific human capital to not investing and forgoing the chance of being a good fit is

$$\begin{aligned} & w_t + (\theta_N + e_t(\theta_{t-1})\Delta\theta)\Delta w_t + \delta \mathbb{E}_{\theta_{t-1}}\left[U^e_{t+1}(\theta_t, \mathbf{w})\right] - c \\ \geq \;& w_t + \theta_N \Delta w_t + \delta U^e_{t+1}(\theta_N, \mathbf{w}), \end{aligned} \tag{9}$$

where $\Delta\theta \equiv \theta_G - \theta_N$ and where, taking into account that the manager is replaced if and only if the firm's cash flows are low (which occurs with probability $1 - \theta_t$), the expected continuation payoffs are defined as

$$\begin{aligned} \mathbb{E}_{\theta_{t-1}}\left[U^e_{t+1}(\theta_t, \mathbf{w})\right] &\equiv e_t(\theta_{t-1}) U^e_{t+1}(\theta_G, \mathbf{w}) + (1 - e_t(\theta_{t-1})) U^e_{t+1}(\theta_N, \mathbf{w}) \\ U^e_{t+1}(\theta_t, \mathbf{w}) &\equiv \theta_t U_{t+1}(\theta_t, \mathbf{w}) + (1-\theta_t)\sum_{j=t+1}^{T} \delta^{j-t-1}\overline{U}. \end{aligned}$$

Since the manager prefers to stay even if her fit is θ_N, it should hold that

$$w_t + \theta_N \Delta w_t + \delta U^e_{t+1}(\theta_N, \mathbf{w}) \geq \sum_{j=t}^{T} \delta^{j-t}\overline{U} + w_{s,t}. \tag{10}$$

Condition (10) will not be binding in the optimal contract, as that would imply that the board needs to increase the manager's pay, so she conceals that her fit is θ_N.

3.1 Dynamics of the Manager's Contract

We can use now conditions (6)–(10) to derive the dynamics of the manager's contract for any given reporting strategy. The key state variables in any given period are the manager's continuation payoffs, depending on her fit realizations, time, and the history of periods without incentives for truthful reporting.

Proposition 1 *The board pursues one of two replacement strategies in any given period t. (i) The first abstains from offering severance pay at $\tau_t = 1$ and stipulates dismissal if the firm's cash flows are low at the end of the period ($\psi_t^1 =$ keep, $\psi_t^2 =$ keep if $x_t = x + \Delta x$, $\psi_t^2 =$ replace if $x_t = x$). Implementing this strategy in period $t < T$ goes hand in hand with deferring bonus payments and offering*

$$
\begin{aligned}
\Delta w_t &= 0 && (11)\\
w_t &= \max\left\{0, \sum_{j=t}^{T} \delta^{j-t}\overline{U}_j - \theta_N \Delta w_t - \delta U_{t+1}^e\left(\theta_N, \mathbf{w}\right)\right\} && (12)\\
w_{s,t} &= 0. && (13)
\end{aligned}
$$

(ii) The second strategy involves offering severance pay at $\tau_t = 1$, which incentivizes truthful reporting (tr.rep.), in which case the manager is replaced in period t if and only if her fit is θ_N ($\psi_t^1 =$ replace if $\widehat{\theta}_t = \theta_N$, $\psi_t^1 =$ keep if $\widehat{\theta}_t = \theta_G$, $\psi_t^2 =$ keep). Implementing this alternative requires that

$$
\Delta w_t = \begin{cases} \max\left\{ \begin{array}{c} \frac{c}{e_t(\theta_G)\Delta\theta} + \delta\frac{U_{t+1}(\theta_N,\mathbf{w}) - U_{t+1}(\theta_G,\mathbf{w})}{\Delta\theta}, \\ \frac{c}{e_t(\theta_G)\theta_G} + \frac{\sum_{j=t}^{T}\delta^{j-t}\overline{U} - \delta U_{t+1}(\theta_G,\mathbf{w})}{\theta_G}, 0 \end{array} \right\} & \text{if tr.rep. in } t-1 \\ \max\left\{ \begin{array}{c} \frac{c}{e_t(\theta_t^*)\Delta\theta} + \delta\frac{U_{t+1}(\theta_N,\mathbf{w}) - U_{t+1}(\theta_G,\mathbf{w})}{\Delta\theta}, \\ \frac{c}{e_t(\theta_t^*)\theta_G} + \frac{\sum_{j=t}^{T}\delta^{j-t}\overline{U} - \delta U_{t+1}(\theta_G,\mathbf{w})}{\theta_G}, \Delta w_{t-n}^{-1}\left(0\right), 0 \end{array} \right\} & \text{otherwise} \end{cases} \quad (14)
$$

$$
\begin{aligned}
w_t &= 0 && (15)\\
w_{s,t} &= \max\left\{0, \theta_N \Delta w_t + \delta U_{t+1}\left(\theta_N, \mathbf{w}\right) - \sum_{j=t}^{T} \delta^{j-t}\overline{U}\right\}, && (16)
\end{aligned}
$$

where $\Delta w_{t-n}^{-1}(0)$ is the minimum bonus required in t to compensate the manager for $n \geq 1$ preceding periods with bonus deferral (i.e., periods without incentives for truthful reporting from $t - n$ to $t - 1$), while satisfying (9) in these periods. The board always pursues truthful reporting in the manager's retirement period T. In the first period, $e_t(\theta_G)$ must be replaced by e_1 in (14).

Part (i) of Proposition 1 considers the case in which the board relies only on firm performance to evaluate the manager's fit. Since severance pay is inadequate to stimulate truthful reporting, it can be set to $w_{s,t} = 0$. To stimulate investment in firm-specific human capital, the board needs to punish the manager for signals indicating no such investment (i.e.,

$w_t = 0$),[13] and to reward her for signals indicating the opposite. However, reminiscent of Lazear's (1979) classical result, once a bonus is paid out, it ceases to have an incentive effect. Hence, it is optimal to defer bonus payments and make them conditional on future success, i.e., we have $\Delta w_t = 0$. The deferral remains in force until a period in which the manager is offered incentives to truthfully report her fit or until she reaches her retirement age. Since, absent truthful reporting, a manager whose fit has deteriorated could stay with the firm, the board needs to decide whether to make the deferred bonus sufficiently large to incentivize firm-specific human capital investment by such a manager in the following period. If so, θ^*_{t-1} in expression (14) should be set to θ_N.[14]

The are two major changes in the contract if the board tries to elicit truthful reporting from the manager (part (ii) of Proposition 1). The first one is that the manager needs to be offered adequate severance pay.[15] The second one is that the attractiveness of deferring compensation is dramatically reduced. The reason is that the signals that the board relies on to infer whether the manager has invested in firm-specific human capital are the manager's reports, and not the firm's cash flows. Thus, incentivizing the manager to invest in firm-specific human capital requires that she is paid more for reporting θ_G. The problem is that the incentive constraint (7) for truthfully reporting a bad fit θ_N requires that the manager is paid severance $w_{s,t}$, which compensates her for the wage she would forgo upon dismissal. This limits the ability to incentivize investment in firm-specific human capital (condition (8)) by deferring compensation into the future, as that increases the manager's continuation payoff also when her fit is θ_N, i.e., deferring increases the manager's incentives to conceal θ_N. This is trivial to see when the manager's continuation payoffs are $U_{t+1}(\theta_G, \mathbf{w}) \approx U_{t+1}(\theta_G, \mathbf{w})$. In this case, only a sufficiently high bonus Δw_t can create incentives for investment in firm-specific human capital (plug (7) into (8)). That bonus might be even higher if it needs to account for bonus deferrals in previous periods. Finally, recall that the source of the manager's rent is the need to incentivize investments in firm-specific human capital. Since the manager is not concerned with forgoing future rent in her retirement period T and pay can no longer be deferred, incentivizing truthful reporting in that period brings no additional cost.

[13] For completeness, note that w_t in (12) might be positive if setting $w_t = 0$ makes it impossible to satisfy (10). However, as noted, (10) woud not be binding in equilibrium.

[14] The second expression in (14) states that the manager's bonus in t is determined by the more stringent of the conditions: (i) that the manager invests in firm-specific human capital in t and (ii) that she does the same in all preceding period(s) in which she is not paid a bonus.

[15] The first term in the max-operators in (14) is larger (less) than the respective second term if $w_{s,t} > 0$ ($w_{s,t} = 0$).

3.2 The Board's Choice of Severance Pay, Contract Horizon, and Turnover

The objective of maximizing the board's payoff (4) can be stated now as choosing the optimal reporting policy for every period, subject to (11)–(16). In this problem, the two continuation payoffs $U_{t+1}(\theta_G, \mathbf{w})$ and $U_{t+1}(\theta_N, \mathbf{w})$, time, and the history of periods with incentives for truthful reporting play the role of state descriptors, completely characterizing the dynamics of the manager's contract. Since a manager's continuation payoffs at retirement are zero ($U_{T+1}(\theta_T, \cdot) = 0$), her payoff can be derived recursively in every period for any truthful reporting policy that the board can choose from. This can then be used to calculate the board's payoff for any such policy and to select the one that maximizes (4).

3.2.1 Illustration of Main Results With Two-Period Employment

Consider the case in which managers retire after two periods, i.e., $T = 2$. Assume, for illustration, that the primitives take the values from the description of Figure 3. Suppose, first, that the board seeks truthful reporting in both periods and offers $\{w_1, w_2\} = \{0, 0\}$, $\{\Delta w_1, \Delta w_2\} = \{6.1, 5.1\}$, and $\{w_{s,1}, w_{s,2}\} = \{2.4, 1.1\}$, which satisfy Proposition 1 for $\overline{U} = 1$. From (8), we can recursively calculate the manager's payoff in a period in which the board offers severance pay as[16]

$$U_t(\theta_{t-1}, \mathbf{w}) = w_{s,t} + \sum_{j=t}^{T} \delta^{j-t}\overline{U} + \max\left\{ e_t(\theta_{t-1}) \begin{pmatrix} w_t + \theta_G \Delta w_t + \delta U_{t+1}(\theta_G, \mathbf{w}) \\ -w_{s,t} - \sum_{j=t}^{T} \delta^{j-t}\overline{U} \end{pmatrix} - c, 0 \right\}. \tag{17}$$

This gives us $U_2(\theta_G, \mathbf{w}) = 2.1$, $U_2(\theta_N, \mathbf{w}) = 2.1$, and $U_1(\mathbf{w}) = 4.4$ (note that $e_t(\theta_{t-1}) = e_1$ in $t = 1$ and that $U_{T+1} = 0$). We can now verify that condition (6) is lax, while (7) and (8) are satisfied with equality. However, if the manager's outside option would be lower, $\overline{U} = 0$, the manager's severance pay would need to be increased to $\{w_{s,1}, w_{s,2}\}_{\overline{U}=0} = \{4.4, 2.1\}$ to satisfy the constraint that she truthfully reports θ_N. Two insights follow immediately. First, the severance pay in $t = 1$ must be higher than in $t = 2$, as the manager needs to be compensated not only for forgoing wages in period one, but also wages and severance pay in period two. Second, *a manager with a higher outside option may be preferable, as she requires less severance pay to be honest.*

Suppose, now, that the board does not seek truthful reporting in the first period and offers $\{w_1, w_2\} = \{0, 0\}$, $\{\Delta w_1, \Delta w_2\} = \{0, 12.4\}$, and $\{w_{s,1}, w_{s,2}\} = \{0, 4\}$. In particular, note

[16] The max-operators in (17) and (18) take into account that the manager invests in firm-specific human capital only if it is optimal to do so. For example, $U_t(\theta_N, \mathbf{w})$ may be an out-of-equilibrium payoff, in an equilibrium in which the manager invests in firm-specific human capital only if her fit in $t-1$ is θ_G.

$\{x, \Delta x\}$	$\overline{U}$	$V^*_{n,r}$	$V^*_{r,r}$
$\{500, 25\}$	0	10,246	10,227
	1	10,240	10,232
$\{500, 250\}$	0	12,858	12,965
	1	12,852	12,971

Table 1: Board's expected payoff. The table compares the board's expected payoff $V^*_{r,r}$ from offering incentives for truthful reporting in both periods and from abstaining from offering such incentives in the first period, $V^*_{n,r}$, for different values of Δx and $\overline{U}$.

that this contract offers no severance pay in $t = 1$ and that the first-period bonus is deferred and paid, conditional on the firm also performing well, in $t = 2$. By plugging in for these contract parameters, we can now verify that conditions (6)–(10) are satisfied. Furthermore, the manager's payoff in $t = 2$ can be calculated again from (17) as $U_2(\theta_G, \mathbf{w}) = 6.4$ and $U_2(\theta_N, \mathbf{w}) = 5.6$. From (9), the manager's payoff in a period in which the board does not offer severance pay is

$$U_t(\theta_{t-1}, \mathbf{w}) = \max \left\{ \begin{array}{c} w_t + (\theta_N + e_t(\theta_{t-1}) \Delta\theta) \Delta w_t - c + \delta \mathbb{E}_{\theta_{t-1}} \left[U^e_{t+1}(\theta_t, \mathbf{w}) \right], \\ w_t + \theta_N \Delta w_t + \delta U^e_{t+1}(\theta_N, \mathbf{w}) \end{array} \right\}. \tag{18}$$

Hence, for an outside option of $\overline{U} = 1$, the manager's payoff in $t = 1$ is $U_1(\mathbf{w}) = 2.7$, which is 38% lower than with truthful reporting in both periods. Thus, there is a *trade-off between lowering the manager's rent and risking an inefficient replacement* decision by evaluating the manager based on the firm's cash flows.

Table 1 compares the board's residual payoff (4) when sticking to each of the two strategies described above for every new hire. It illustrates two of the paper's main results: (i) *If the board seeks truthful reporting, it may prefer hiring a manager with a higher outside option*, especially if pursuing this policy in both periods (last column); (ii) *a lower outside option makes offering severance pay less attractive (thus, making the board reliant only on the firm's noisy performance)*, especially if the cash flow upside is small ($V^*_{n,r} > V^*_{r,r}$ if $\Delta x = 25$).[17] The rest of the paper generalizes and discusses the intuition for these insights. It further shows that, with longer time to retirement ($T > 2$), even firms with a high cash flow upside regularly abstain from offering incentives for truthful reporting (Figure 3). Readers interested mainly in the implications of the analysis can skip to Implications 1–6.

[17] The derivations are in Appendix B. If $\overline{U} = 0$, we need to set $\Delta w_2 = 10.9$ and $w_{s,2} = 4.4$ and we have $U_1(w) = 1.8$.

3.2.2 The Cost-Benefit Mismatch of Offering Severance Pay

The first question that should be clarified is why the board might choose not to offer incentives for truthful reporting in all periods. The first-best replacement policy would require seeking truthful reporting in every period and hiring the manager with the longest time until retirement T, as the positive correlation of fit between periods ($e(\theta_G) > e_1$) implies that a manager who stays a good fit should be kept as long as possible. However, this policy might require giving up too much information rent to the manager. In the special case in which the board offers severance pay in every period, the manager's expected payoff in (17) reduces simply to

$$U_t(\theta_{t-1}, \mathbf{w}^r) = w_{s,t}^r + \sum_{j=t}^{T} \delta^{j-t}\overline{U}, \tag{19}$$

as the constraint that the manager invests in firm-specific human capital given that she was a good fit last period binds in every period (the superscript r stands for truthful reporting). Thus, offering the manager severance pay not to withhold information revealing her as a bad fit leads to information rent of proportionate size.[18] Plugging in for the severance pay $w_{s,t}^r$ (while neglecting for a moment its zero lower bound), the manager's expected rent (5) when she is hired in period one is

$$\begin{aligned}
\nu_1(\mathbf{w}^r) = w_{s,1}^r &= \frac{\theta_N c}{e_1 \Delta\theta} + \delta \frac{\theta_G U_2(\theta_N, \mathbf{w}^r) - \theta_N U_2(\theta_G, \mathbf{w}^r)}{\Delta\theta} - \sum_{j=1}^{T} \delta^{j-1}\overline{U} \\
&= \frac{\theta_N}{e_1 \Delta\theta} c + \delta \left(w_{s,2}^r + \sum_{j=2}^{T} \delta^{j-2}\overline{U} \right) - \sum_{j=1}^{T} \delta^{j-1}\overline{U} \\
&= \sum_{j=1}^{T} \delta^{j-1} \left(\frac{\theta_N c}{e_j(\theta_G)\Delta\theta} - \overline{U} \right),
\end{aligned} \tag{20}$$

Proposition 2 *The board follows the first-best policy of hiring the manager with the longest time until retirement and incentivizing truthful reporting in all periods if*

$$\frac{\theta_N c}{e_t(\theta_G)\Delta\theta} \leq \overline{U}. \tag{21}$$

If the reverse inequality holds, there is a threshold $\widehat{\delta}$, such that the board will hire a manager with less time to retirement (lower T) or abstain from seeking truthful reporting in some

[18] Note that the manager's rent in t does not depend on her type in $t-1$. This is because the manager's rent is defined by her rent when not investing in firm-specific human capital (cf. (8)), and the latter coincides with the out-of-equilibrium rent of a manager who was a bad fit in $t-1$, but stays until t without investing in firm-specific human capital in t.

periods if $\delta > \widehat{\delta}$. In these periods, dismissal is triggered by low cash flow performance.

Together, Proposition 1 and 2 imply:

Corollary 1 ***Dynamic effects of severance pay:*** *Offering severance pay in t (i) reduces the attractiveness of deferring compensation beyond t and (ii) increases the incentives to conceal θ_N and, thus, the necessary severance pay in the preceding periods.*

If condition (21) does not hold, the manager can expect an "efficiency" wage above her outside option in all periods until she is replaced. This leads to two effects specific to a dynamic setting that increase her incentive to conceal a deteriorating fit. First, because the board retains the manager even following low cash flow realizations when she reports θ_G, the manager can enjoy a positive continuation payoff regardless of the cash flows in the present period. As noted above, this dramatically reduces the attractiveness of deferring compensation.[19] Second, offering severance pay in period t, increases the manager's reluctance to report truthfully in all *preceding* periods since, from the perspective of these periods, staying with the firm would reward the manager regardless of her fit. Hence, the more periods with contractually specified severance pay remain on the manager's contract, the higher the severance pay the board would need to offer the manager to compensate her for forgoing future pay. This insight highlights the mismatch between the costs and benefits of offering severance pay in multiple subsequent periods. For illustration, consider again the policy of offering severance pay in all periods until the manager's retirement age T (dashed line in Figure 3). With such a policy, if the manager is replaced in period t, the board does not realize the benefit of truthful reporting in periods $t+1$ until T. However, because the severance payments in $t+1$ until T compensate the manager for the forgone future wages and severance payments, the severance pay in t compensates the manager as if she would have run the firm until retirement. For this reason, it may be optimal not to seek truthful reporting at least in some periods (despite the risk of inefficient retention and replacement decisions) or to hire an older manager.

3.2.3 Relying on Noisy Performance Measures Instead of Severance Pay

In all that follows, suppose that the first-best condition (21) is not satisfied. As illustrated in Section 3.2.1, not offering severance pay comes with the compelling advantage that it can reduce the rent that must be promised to the manager. The manager can enjoy then her

[19] Interestingly, Anderson et al. (2018) also show that boards front-load compensation when a manager expects that she may be replaced in the future by a new manager who is a better match to the firm's changing environment. However, in their full-commitment framework, severance pay is never optimal, whereas in the present framework, severance pay is the main reason for front loading.

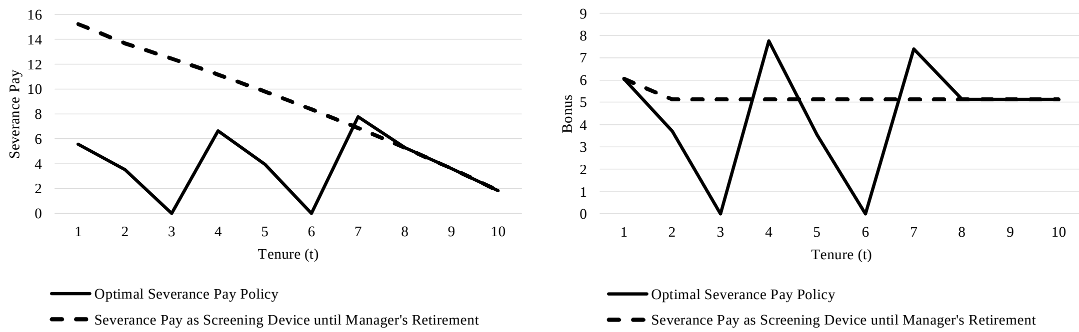

Figure 3: Optimal Screening with Severance Pay vs. Offering Severance Pay in All Periods. The dips in bonus and severance pay correspond to periods in which the board does not offer incentives for truthful reporting, but relies on firm performance. In terms of implementation, the dips would correspond to the end dates of renewable fixed-term contracts. The simulations are performed with $T = 10$, $e_1 = 0.55$, $e(\theta_G) = 0.65$, $e(\theta_N) = 0.45$, $c = 1$, $\theta_G = 0.7$, $\theta_N = 0.4$, $\overline{\theta} = 0.48$, $\delta = 0.95$, $x = 500$ and $\Delta x = 250$, $\overline{U} = 0$. The figure illustrates that, even though the manager's pay is small relative to the firm's size, it is optimal not to offer severance pay at regular intervals.

continuation payoff (above her outside option) only if she generates high cash flows, while in a period with severance pay, she can enjoy such continuation payoff regardless of the cash flows in that period. The prospect of lower future rent, in turn, implies that the manager is truthful about her fit even when offered lower severance pay in all preceding periods.

Proposition 3 *Reducing the manager's rent in period t by abstaining from offering severance pay in that period, reduces the manager's rent and severance pay in all preceding periods.*

Proposition 3 implies that, by choosing whether to seek truthful reporting, the board faces a trade-off between making more efficient replacement decisions and minimizing the manager's rent. This trade-off is at the heart of Figure 3, which plots the optimal contract offered to the manager when the board determines the optimal sequence of truthful reporting periods that maximizes its expected payoff and compares it to a contract that always offers incentives for truthful reporting. Solving for the optimal truthful reporting policy for $T > 2$ is not tractable. Numerically, this can easily be done by recursively deriving the manager's payoff from (17)–(18) in every period for any truthful reporting policies that the board can choose from, and then selecting the policy that maximizes (4).

Figure 3 illustrates that the manager's severance pay increases in the remaining periods stipulating severance pay for premature dismissal (Proposition 2). However, by introducing a period without incentives for truthful reporting, the board not only saves on the cost of offering severance pay in that period, but can also afford to offer lower severance pay in all preceding periods (Proposition 3). Figure 3 illustrates that it is optimal to abstain from offering severance pay in periods three and six, even though the manager's wage is only a very small fraction of the firm's value. This is because such a policy halves the manager's expected rent, while exposing the board to the relatively mild risk of making a wrong replacement decision in these periods, which the incumbent manager reaches with relatively low probability. Finally, note that the manager is promised a non-trivial bonus for achieving high cash flows in periods in which her contract offers severance pay, with bonus increases following periods without severance pay to account for bonus deferrals in such periods (Proposition 1).

3.2.4 Implementation with Renewable Fixed-Term Contracts

A simple way to implement policies that alternate between offering and not offering incentives for truthful reporting (and, thus, severance pay) in regular intervals is with renewable fixed-term contracts (i) that stipulate severance pay for premature terminations as a multiple of the manager's bonus and remaining tenure; (ii) for which severance is not paid when not renewing the manager's contract on renewal dates; and (iii) for which, absent termination, the contract continues as originally agreed upon. The period between two renewal dates defines then the length of each renewable fixed-term contract. Note that, in practice, an executive cannot claim severance pay without a "good reason," such as a change in duties, diminution of pay, or relocation (Rau and Xu, 2013). However, if the board learns that the manager's fit has deteriorated, it would trigger such termination. As will be discussed in Section 3.3.3, this implementation extends to a more general model in which the board updates its beliefs about the manager's fit over an extended period. Also in this case, renewal corresponds to a period in which the board would optimally replace the manager without paying severance following a history of low cash flows.

In practice, about 45% of S&P 500 CEOs and 68% of S&P 1500 executives are hired with explicit contracts that typically cover a prespecified period, stipulate severance pay for premature termination, and renew automatically, unless one of the parties objects (Gillan et al., 2009; Rau and Xu, 2013; Brown, 2015). In line with Proposition 1, the severance agreements are usually a multiple of managers' salary and bonus and can depend on their remaining tenure. Also in line with the model, boards pay more attention to performance, and the relation between turnover and performance becomes stronger close to renewal dates (Liu and Xuan, 2016; Cziraki and Groen-Xu, 2017). Thus, the paper provides a simple intuition for the widespread use of such renewable fixed-term contracts.

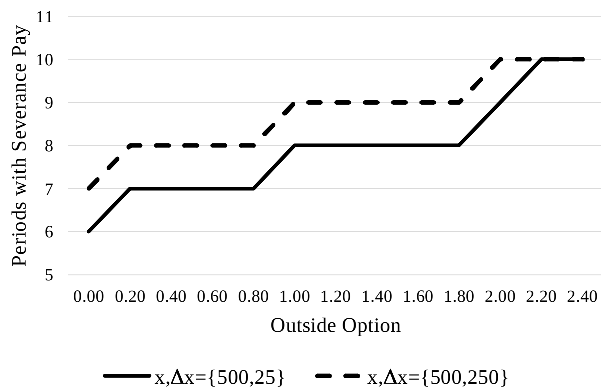

Figure 4: Comparative Statics in $\overline{U}$ and Δx. The figure plots how the number of periods in which the board incentivizes truthful reporting through severance pay changes in $\overline{U}$ and Δx. The primitives are the same as in Figure 3, with $T = 10$ and $\overline{U}$ taking values from zero to two and $\{x, \Delta x\}$ taking values $\{500, 25\}$ and $\{500, 250\}$, respectively. The figure illustrates that the board seeks more truthful reporting when the manager's outside option $\overline{U}$ and the cash flow upside Δx are higher.

3.2.5 Determinants of Contract Horizon

The next step is to highlight the wide-ranging implications of the factors determining whether the board offers severance pay and, thus, the length of renewable contracts. Figure 4 illustrates that the board seeks more truthful reporting when the manager's outside option $\overline{U}$ is higher. The intuition is that a higher outside option makes the manager less reluctant to report that her fit is θ_N and to seek alternative employment. Hence, the severance pay that the board needs to promise the manager to reveal that her fit has deteriorated is lower. Furthermore, Figure 4 shows that incentivizing truthful reporting becomes more attractive as Δx increases. Intuitively, firms with higher growth prospects have more to lose from not having the right manager in charge. These results can be derived analytically for the case in which managers leave the labor market after at most two periods (i.e., $T = 2$). An optimal contract that does not offer severance pay in the first period would then be implemented with two renewable fixed-term contracts, each having a length of one period.

Proposition 4 *Suppose that managers retire after two periods $T = 2$. (i) Offering severance pay to incentivize truthful reporting becomes less attractive for the board as $\overline{U}$ and Δx decrease. (ii) The board can implement an optimal policy of not offering severance pay in the first period by offering a fixed-term contract that goes for one period and that can be renewed for one more period.*

Implication 1 ***Determinants of contract horizon and turnover-performance sensitivity:*** *(i) The length of renewable fixed-term contracts will be shorter and turnover-*

performance sensitivity will be higher when managers' outside options are low. (ii) Firms with a higher cash flow upside will offer contracts with longer horizons. This will go hand in hand with higher (average) severance pay and a higher likelihood of a timely turnover preempting underperformance.

To the extent that managers' outside options are lower in industry downturns — e.g., because more firms are going bankrupt; fewer firms are being started; and more competition for available positions exists within the labor force — an immediate corollary is:

Implication 2 ***Turnover in downturns****: (i) Dismissal decisions rely more strongly on performance measures in industry-wide downturns. This offers an alternative explanation (compared to a lack of relative performance evaluation) for the stronger turnover-performance sensitivity in downturns. (ii) The higher reliance on performance measures to infer the manager's fit increases the risk of making a wrong retention or replacement decision, which might exacerbate downturns.*

Section 3.3 returns to these determinants of contract horizon, as they shape also other key aspects of employment relationships.

3.2.6 Renegotiations and Hiring Older Managers

The preceding results show that it might be optimal for the board to abstain from offering severance pay in some periods in order to limit the manager's rent. However, once reaching such a period, the board could renegotiate the existing contract and offer severance pay that would incentivize truthful reporting. There are several reasons why pursuing this strategy can become optimal ex post, even if it is not optimal ex ante. First, one benefit not offering severance pay in t is that it decreases the manager's rent not only in t, but also in all preceding periods. However, the latter benefit ceases to exist once both parties arrive at t.[20] Second, there is scope for renegotiations after the manager has invested in firm-specific human capital and the cost c in period t is sunk since, at that point, incentivizing firm-specific human capital investment is no longer an objective. As is usual, although such renegotiations might be beneficial ex post, they limit the board's contracting options ex ante.

A commitment to avoid offering severance pay on renewal dates or holding up the manager can be achieved in the present context. In particular, if the board engages in renegotiations once, all future managers would expect the same and demand only renegotiation-proof contracts from then on. Since the firm is infinitely-lived, this "trigger strategy" would prevent

[20] As in Malcomson (2016), Definition 1 does not rule out that there might be scope for renegotiations if the manager's type has not been revealed.

the board from deviating from its commitment to avoid renegotiations if δ is sufficiently high.

Still, it is interesting to consider the extreme case in which the potential for renegotiations forces the board to offer a contract that offers severance pay in all periods. This case has clear implications for the manager's age, as then the contract length coincides with her time to retirement T.[21,22] In this case, the board's problem is equivalent to that discussed in Proposition 2. Then, if the first-best condition (21) is not satisfied, the board might have to choose a manager with less time to retirement, as her rent might otherwise become too high.

To gain some intuition about the factors affecting whether to hire a younger or an older manager, it is helpful to consider again condition (21). It suggests that a higher per-period outside option $\overline{U}$ decreases the manager's rent. A more attractive outside opportunity makes the manager less reluctant to leave the firm, which reduces the severance pay she needs to be promised to truthfully report her fit. Thus, given that the cost of employing the manager longer is lower, while the benefit is unchanged, the board finds it optimal to offer contracts with longer horizons. Also similar to before, longer contracts are optimal if Δx is higher, as there is more to gain from holding on longer to a manager who is a good fit.

Proposition 5 *Suppose that the first-best condition (21) is not satisfied and that the board chooses to stimulate truthful reporting by offering severance pay in all periods. Then, the board chooses a manager with a longer time to retirement T if her per-period outside option $\overline{U}$ and the cash flow upside Δx are higher.*

Implication 3 ***CEO age****: If the board cannot commit not to offer severance pay to the manager upon dismissal, ex ante severance agreements must be higher for younger managers. (ii) Furthermore, the board prefers hiring an older manager (i.e., T is lower) in industry downturns and by firms with low growth potential.*

While the second part of Implication 3 has not been tested, the first part finds empirical support in Rau and Xu (2013).

3.3 Discussion and Extensions

The following section discusses several extensions of the model, including a generalization with continuous types in which the board relies on both cash flow performance and severance

[21] There are no clear predictions regarding T from the previous section except that firms with a higher cash flow upside Δx tend to choose higher T.

[22] There is evidence that discretionary severance pay (not required by a CEO's contract) is often granted to managers. However, while Yermack (2006) documents that such pay is typically modest, Goldman and Huang (2015) find that often also the opposite is true.

pay to infer the manager's fit.

3.3.1 Hiring Managers with Better Outside Options

Suppose that the pool of potential managers differs according to their likelihood of success e_t and their outside options $\overline{U}$. Assume further that condition (21) is not satisfied for any $\overline{U}$ and e_t. All remaining parameters of the model remain the same.

Clearly, if all information were common knowledge, the board would prefer hiring the manager with the highest likelihood e_t of being a good fit and with the lowest outside option $\overline{U}$. However, when the board needs to incentivize the manager to invest in firm-specific human capital, it needs to pay her an "efficiency wage" that is above her outside option. In this case, hiring a manager with a higher outside option could *lower* the board's expected wage bill since it reduces the need for generous severance pay $w_{s,t}$. This can be seen especially clearly in the extreme case in which the board stimulates truthful reporting in all periods (as in Section 3.2.6). Then, the board strictly prefers hiring a manager with a higher outside option. Thus, if the board faces the choice of selecting between a manager that has a higher likelihood e_t of being a good fit or one with a higher outside option $\overline{U}$, it might prefer the manager with the higher outside option.

Proposition 6 *(i) Take any given policy of stimulating truthful reporting through severance pay. If* $\left(p\left(\mathbf{w}\right)-\delta^{T}\right)>\left(1-\delta\right)\frac{\partial}{\partial\overline{U}}U_{1}\left(\mathbf{w}\right)$ *(*$p\left(\mathbf{w}\right)$ *is defined in the Appendix), the board's payoff increases in the manager's outside option* $\overline{U}$ *for that policy. (ii) When offering severance pay in all periods to a manager who is paid above her outside option, the board always prefers hiring a manager with a higher outside option.*

Investments in general human capital, such as taking board seats at different firms, increase CEOs' outside options. Thus, it might seem surprising that boards would tolerate such behavior given that it might distract managers, and they could then use it as a bargaining chip to get a higher salary. However, CEOs rarely leave the firm to become CEOs elsewhere (Fee and Hadlock, 2004), and their labor income declines, on average, by 40% following termination (Nielsen, 2017). Thus, when managers are paid above their outside options anyway, the more relevant effect might be:

Implication 4 ***Investments in general human capital:*** *Investments in general human capital, increasing a manager's outside options, make it cheaper to offer severance pay that compensates the manager for leaving prematurely and not trying to avoid termination. Thus, boards might tolerate such investments even if they come at the expense of investments in firm-specific human capital (lower* e_t*).*

The flip side of the preceding implication is that managers would be more reluctant to leave firms that require high investments in firm-specific human capital that leave them little opportunity to invest in general human capital. This may necessitate relying more strongly on firm performance to identify managers with deteriorating fit.

3.3.2 Outside Options and Experience

Up until now, we have assumed that outside options are fixed over time. However, as managers stay longer with the firm, their reputation in the labor market may improve, and, as a result, their outside options may increase. One of the main points of the present paper is that such increases do not necessarily make the manager more expensive to the firm. In fact, the exact opposite might be the case. If the manager is paid an efficiency wage above her outside option, increases in that outside option imply that it becomes easier to keep the manager honest. Thus, one could expect the board to rely more on severance pay and less on firm performance measures to judge the manager's fit. Further relating to the implementation of the optimal contract, we have:

Implication 5 ***Experience****: If a manager's outside options increase with her tenure: (i) the length of renewable contacts will increase with the manager's tenure; and (ii) the relation between managerial turnover and firm performance will weaken with the manager's tenure.*

Another implication of the analysis is that boards will avoid damaging departing CEOs' outside options, as this would necessitate higher severance pay. Indeed, there is evidence that severance pay is higher when firms replace CEOs with a reputation for firm mismanagement (Goldman and Huang, 2015).

Implication 6 ***Reputation****: When replacing CEOs, boards will avoid damaging their reputation, as replacing CEOs with a reputation for mismanagement requires offering higher severance pay.*

3.3.3 Relying on Both Truthful Reporting and Performance Measures

To highlight the main novel economic insights, the baseline model made a number of simplifying assumptions, which reduced the board's problem to choosing between incentivizing truthful reporting, in which case the manager is replaced if her fit is θ_N, and relying exclusively on firm performance, in which case replacement follows low cash flows. In practice, boards may rely on both strategies in every period. Incorporating this into the model, does not change the qualitative insights.

Specifically, consider an extension of the model in which the manager's fit realization θ_t is a draw from the continuous support $\left[\underline{\theta}, \overline{\overline{\theta}}\right]$ according to the probability distribution $F(\theta_t|\theta_{t-1})$, which depends on the manager's fit realization from the previous period. In this extension, the severance pay offered by the board would determine a cutoff $\widetilde{\theta}_t \geq \underline{\theta}$ in every period, such that for all types below the cutoff the ex ante specified severance package is more attractive than trying to avoid dismissal, while for all types above the cutoff, staying with the firm is better.[23] Inefficiency results if the cutoff is different from the lowest fit that the board would keep under symmetric information. In this extension, if the manager stays, the board knows only that her fit is distributed over $[\widetilde{\theta}_t, \overline{\overline{\theta}}]$, and the cash flow realization at the end of the period would affect the posterior belief over this set. Clearly, the lower is $\widetilde{\theta}_t$ (i.e., the lower is severance pay in period t), the more informative the firm's cash flows are about whether the manager's fit is low. The board's main control variables are then the cutoffs $\widetilde{\theta}_t$ in every period, as well as the bonuses Δw_t (both $\widetilde{\theta}_t$ and Δw_t can be history-dependent) that determine which managers the board would like to stimulate to invest in firm-specific human capital. In this extension, contract renewal is again deterministic, with the first renewal period corresponding to the first period in which the board would want to replace the manager without severance pay following the worst-possible history of cash flow realizations. The occurrence of this period depends on the chosen cutoffs $\widetilde{\theta}_t$. High cutoffs would imply a weaker relation between turnover and underperformance. Intuitively, retention and dismissal decisions are based on more information: not only on noisy performance, but also on the additional information reported (or not withheld) by the manager. Formally, conditional on a manager staying with the firm, her fit would be above a higher cutoff. Hence, there would be less-negative updating about her fit following low cash flows since the manager's fit would be coming from a "better" distribution. This does not mean that there would be less turnover, but that turnover is more likely to be accompanied with a severance package and, because dismissal is based on more information, it may come earlier before the firm has underperformed dramatically due to a bad fit.[24]

Solving numerically this extension yields similar insights to those in Figures 3 and 4. Specifically, the manager's severance pay increases in her remaining tenure when the board implements high cutoff types $\widetilde{\theta}_t$ over that tenure. High cutoffs lead to longer contracts with a lower sensitivity of turnover to cash flow performance, and the frequency of renewal periods

[23] See Inderst and Mueller (2010) for a one-period formalization.

[24] The second renewal period would correspond then to the period in which the board would like to replace the manager without severance pay following the second-worst history of cash flow realizations (conditional on the manager not being dismissed with a severance package until then). Note that the contract may stipulate severance pay also at the *interim* dates of renewal periods, i.e., as long as premature dismissal reduces the likelihood of low cash flows in that period.

tends to decrease in the manager's outside option $\overline{U}$ and the cash flow upside Δx.[25]

4 Conclusion

This paper analyzes optimal contract horizon, severance pay, and turnover in a model in which managers' fit with the firm evolves over time, and managers are better informed about such changes. The board can minimize the likelihood that a manager whose fit has deteriorated will stay with the firm, but this may require generous severance packages to incentivize such a manager not to mislead the board about the quality of her fit. The costs of severance pay in a given period, specific to a dynamic setting, are that it limits the ability to defer compensation and that it increases the incentives to conceal a bad fit in preceding periods. Thus, a board may prefer, instead, to rely more heavily on the firm's noisy cash flow performance when evaluating a manager. Although such a policy might not preempt bad performance and might lead to less-efficient replacement and retention decisions, it reduces the manager's ability to extract rent from the firm.

The main predictions from the analysis are as follows. It is optimal for the board to allow for periods in which it has the option to fire the manager without paying severance. In these periods, the relation between dismissals and underperformance will be strongest, and managers will sometimes be dismissed for bad luck and not for being a bad fit. The resulting optimal contract can be implemented with renewable fixed-term contracts, stipulating severance pay upon early termination, but allowing the board to costlessly replace the manager when her term expires. These contracts are widely used in practice (Gillan et al., 2009; Rau and Xu, 2013), but have not been addressed by prior theory.

The paper further offers novel insights regarding the determinants of the length of contracts. A key factor is a manager's outside option. The severance pay needed to incentivize a manager to leave is higher if her outside option is low. Thus, in such cases, the board will rely less on offering adequate severance pay and more on the firm's cash flow performance to screen out managers with deteriorating fit. This would make it optimal to offer shorter renewable fixed-term contracts, since such contracts make it possible to terminate the man-

[25] The economic forces would be similar in an extension in which firm performance brings additional information relative to the manager's information at the interim date. Another extension involves allowing for mixed strategies. Specifically, it is conceivable that there is an equilibrium in which the manager randomizes between reporting θ_N or θ_G if her fit is θ_N, and the board subsequently randomizes between replacing and keeping the manager depending on the cash flow realization. Such stochastic replacement reduces the manager's on- and off-equilibrium continuation payoffs and, thus, the need for severance pay at the expense of increasing the likelihood of making a wrong replacement decision following a low cash flow realization. In the presence of limited commitment, randomization generates interesting dynamics even with constant types (Hart and Tirole, 1988; Laffont and Tirole, 1990).

ager's contract costlessly at the end of its term. Another implication is that managers will have less-adequate incentives to reveal a bad fit in industry downturns (when their outside options are low), leading boards to rely more on firm performance and, in turn, to a tighter link between CEO turnover and underperformance. There is, indeed, evidence for this prediction, but it has hitherto been interpreted as a lack of relative performance evaluation (Jenter and Kanaan, 2015). Furthermore, the stronger reliance on noisy performance in downturns is more likely to leave firms with managers who are not a good fit, which might exacerbate downturns. Factors such as the firm's growth opportunities also matter for contract horizon. Contract length will be longer and, thus, severance pay will be higher for firms with better growth prospects who stand to gain more from having the right manager in charge.

The manager's age also plays a key role as a commitment device to shorter contracts, especially if there is scope for renegotiations to offer severance pay at contract expiration. Hiring an older manager helps then to keep the manager's pay from growing too large, as older managers would need to be offered lower severance packages.

The insight that a higher outside option makes a manager potentially cheaper to replace and, thus, potentially cheaper to employ has several broader implications for employment relationships. One is that a board might appear to hire from a "select club," i.e., a manager with high outside options, even if it is unlikely that she is a better fit. This is particularly true when managers are paid above their outside option anyway. For the same reason, boards might tolerate investments in general human capital even if they come at the expense of firm-specific human capital investments. Further work may generalize the model, endogenize the managers' outside options, and differentiate between the effect of temporary and permanent shocks to the manager's fit.

References

[1] Aghion, Philippe, and Jean Tirole, 1997, Real and formal authority in organizations, *Journal of Political Economy* 150(1), 1–29.

[2] Almazan, Andres, and Javier Suarez, 2003, Entrenchment and severance pay in optimal governance structures, *Journal of Finance* 63(2), 519–547.

[3] Anderson, Ronald W., M. Cecilia Bustamante, Stéphane Guibaud, and Mihail Zervos, 2018, Agency, firm growth, and managerial turnover, *Journal of Finance* 73(1), 419–464.

[4] Bernardo, Antonio, and Ivo Welch, 2001, On the evolution of overconfidence and entrepreneurs, *Journal of Economics and Management Strategy* 10(3), 301–330

[5] Bester, Helmut, and Roland Strausz, 2001, Contracting with imperfect commitment and the revelation principle: the single agent case, *Econometrica* 69(4), 1077–1098.

[6] Brown, Kareen E., 2015, Ex ante severance agreements and earnings management, *Contemporary Accounting Research* 32(3), 897–940.

[7] Crémer, Jacques, 1995, Arm's length relationships, *Quarterly Journal of Economics* 110(2), 275–295.

[8] Cziraki, Peter, and Moqi Groen-Xu, 2015, CEO career concerns and risk taking, Working Paper, LSE and University of Toronto.

[9] Eisfeldt, Andrea L., and Adriano Rampini, 2008, Managerial incentives, capital reallocation, and the business cycle, *Journal of Financial Economics* 87(1), 177–199.

[10] Eisfeldt, Andrea L., and Camelia M. Kuhnen, 2013, CEO turnover in a competitive assignment framework, *Journal of Financial Economics* 109(2), 351–372.

[11] Fee, C. Edward, and Charles J. Hadlock, 2004, Management turnover across the corporate hierarchy, *Journal of Accounting and Economics* 37(1) 3–38.

[12] Fee, C. Edward, Charles J. Hadlock, Jing Huang, and Joshua R. Pierce, 2015, Robust models of CEO turnover: new evidence on relative performance evaluation, Working Paper, Tulane University, Michigan State University, University of South Carolina, and University of Kentucky.

[13] Felli, Leonardo, and Christopher Harris, 1996, Learning, wage dynamics, and firm-specific human capital, *Journal of Political Economy* 104(4), 838–868.

[14] Garrett, Daniel, and Alessandro Pavan, 2012, Managerial turnover in a changing world, *Journal of Political Economy* 120(5), 879–925.

[15] Gervais, Simon, J. B. Heaton, and Terrance Odean, 2011, Overconfidence, compensation contracts, and capital budgeting, *Journal of Finance* 66(5), 1735–1777.

[16] Gillan, Stuart L., Jay C. Hartzell, and Robert Parrino, Explicit versus implicit contracts: evidence from CEO employment agreements, *Journal of Finance* 64(4), 1629–1655

[17] Goel, Anand M., and Anjan Thakor, 2008, Overconfidence, CEO selection, and corporate governance, *Journal of Finance* 63(6), 2737–2784.

[18] Goldman, Eitan M., and Peggy P. Huang, 2015, Contractual vs. actual separation pay following CEO turnover, *Management Science* 61(5), 1108–1120.

[19] Halac, Marina, 2012, Relational contracts and the value of relationships, *American Economic Review* 102(2), 750–779.

[20] Hart, Oliver D., and Jean Tirole, 1988, Contract renegotiation and Coasian dynamics, *Review of Economic Studies* 55(4), 509–540.

[21] He, Zhiguo, 2012, Dynamic compensation contracts with private savings, *Review of Financial Studies* 25(5), 1494–1549.

[22] He, Zhiguo, Bin Wei, Jianfeng Yu, and Feng Gao, 2017, Optimal long-term contracting with learning, *Review of Financial Studies*, forthcoming.

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

[24] Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini, 2010, Information and incentives inside the firm: evidence from loan officer rotation, *Journal of Finance* 65(3), 795–828.

[25] Huang, Sheng, Johan Maharjan, and Anjan Thakor, 2016, Disagreement-induced CEO turnover, Working Paper, Washington University in St. Louis.

[26] Inderst, Roman, and Holger M. Mueller, 2010, CEO replacement under private information, *Review of Financial Studies* 23(8), 2935–2969.

[27] Innes, Robert D., 1990 Limited liability and incentive contracting with ex-ante choices, *Journal of Economic Theory* 52(1), 45–67.

[28] Jenter, Dirk, and Fadi Kanaan, 2015, CEO turnover and relative performance evaluation, *Journal of Finance* 70(5), 2155–2184.

[29] Jenter, Dirk, and Katharina Lewellen, 2017, Performance-induced CEO turnover, Working Paper, LSE and Darthmouth College.

[30] Jovanovic, Boyan, 1979 a, Job matching and the theory of turnover, *Journal of Political Economy* 87(5), 972–990.

[31] Jovanovic, Boyan, 1979 b, Firm-specific capital and turnover, *Journal of Political Economy* 87(6), 1246–1260.

[32] Laffont, Jean-Jacques, and Jean Tirole, 1990, Adverse selection and renegotiations in procurement, *Review of Economic Studies* 57(4), 597–925.

[33] Lazear, Edward P., 1979, Why is there mandatory retirement, *Journal of Political Economy* 87(6), 1261–1284.

[34] Lazear, Edward P., 2009, Firm-specific human capital: a skill-weights approach, *Journal of Political Economy 117(5), 914–940.*

[35] Levitt, Steven D., and Christopher M. Snyder, 1997, Is no news bad news? Information transmission and the role of "early warning" in the principal agent model, *RAND Journal of Economics* 28(4), 641–661.

[36] Liu, Ping, and Yuhai Xuan, 2016, The contract year phenomenon in the corner office: an analysis of firm behavior during CEO contract renewals, Working Paper, University of Illinois at Urbana-Champaign.

[37] Malcomson, James M., 2016, Relational incentive contracts with persistent private information, *Econometrica* 84(1), 317–346.

[38] Nielsen, Kasper M., 2017 Personal Costs of Executive Turnovers, Working Paper, Hong Kong University of Science and Technology.

[39] Prescott, Edward S., and Richard Townsend, 2006, Private information and intertemporal job assignments, *Review of Economic Studies* 73(2), 531–548.

[40] Rau, P. Raghavendra, and Jin Xu, 2013, How do ex ante severance pay contracts fit into optimal executive incentive schemes?, *Journal of Accounting Research* 51(3), 631–671.

[41] Sannikov, Yuliy, 2008, A continuous-time version of the principal-agent problem, *Review of Economic Studies* 75(3), 957–984.

[42] Taylor, Lucian A., 2010, Why are CEOs rarely fired? Evidence from structural estimation, *Journal of Finance* 65(6), 2051–2087.

[43] Van den Steen, Eric, 2004, Rational overoptimism (and other biases), *American Economic Review* 94(4), 1141–1151.

[44] Van Wesep, Edward D., 2010, Pay (be)for(e) performance: the signing bonus as an incentive device, *Review of Financial Studies* 23(10), 3812–3848.

[45] Van Wesep, Edward D., and Sean Wang, 2014, The prevention of excess managerial risk taking, *Journal of Corporate Finance* 29, 579–593.

[46] Varas, Felipe, 2017, Managerial short-termism, turnover policy, and the dynamics of incentives, *Review of Financial Studies*, forthcoming.

[47] Wang, Cheng, 2011, Termination of dynamic contracts in an equilibrium labor market model, *Journal of Economic Theory* 146(1), 74–110.

[48] Wang, Cheng, and Youzhi Yang, 2015, Outside opportunities and termination, *Games and Economic Behavior* 91, 207–228.

[49] Yermack, David, 2006, Golden handshakes: separation pay for retired and dismissed CEOs, *Journal of Accounting and Economics* 41(3), 237–256.

[50] Zhu, John Y., 2018, Disagreement and short-term contracts, Working Paper, University of Pennsylvania.

Appendix A Omitted Proofs

Proof of Proposition 1. The proof boils down to determining which constraints bind and determining the dynamics of the manager's compensation contract. The technical derivations are relegated to Lemma B.1 in Appendix B, as they bring little further insight relative to the sketch of the proof briefly presented here and the intuition contained in the main text.

Consider, first, the case in which the board does not offer incentives for truthful reporting in period $t < T$. Since w_t does not affect the incentives to invest in firm-specific human capital, it is optimally set to zero, unless the interim participation constraint (10) becomes binding. Severance pay is (weakly) suboptimal in such a period (i.e., $w_{s,t} = 0$), as it does not improve the manager's effort incentives, but could make condition (10) more difficult to satisfy. Finally, deferring the manager's bonus by a period (i.e., satisfying constraint (9) by promising the manager a higher continuation payoff in return for setting $\Delta w_t = 0$) helps relax the constraint that the manager invests in firm-specific human capital in the subsequent period(s).

Consider, next, the case in which the board incentivizes truthful reporting in period t. Then, the manager's severance pay $w_{s,t}$ is determined by (7). If this constraint were not binding, it would be optimal to decrease $w_{s,t}$, as this would reduce the manager's pay, while improving her incentives to invest in firm-specific human capital. Condition (7) is not binding only if it is satisfied for $w_{s,t} = 0$. If there is truthful reporting in $t-1$, Δw_t is determined from the constraint that the manager invests in firm-specific human capital (8). Note that depending on whether the zero lower bound of $w_{s,t}$ is binding, we have a different expression for the minimum bonus required to satisfy (8). This is captured by the first two terms in the max-operators in (14). Furthermore, note that it is without loss of generality that Δw_t does not depend on the cash flow history when there is truthful reporting in $t-1$. Since the only thing that determines whether manager is retained in $t-1$ is whether her fit is θ_G (which is what a successful investment in firm-specific human capital leads to), such history-dependence would be irrelevant for satisfying the incentive constraints in $t-1$.[26] If, instead, t follows $n \geq 1$ periods in which the board does not offer incentives for truthful reporting (from $t-n$ to $t-1$), the bonus and respective continuation payoffs in t must be high enough to satisfy (9) in the preceding n periods (for which $\Delta w_{t-n} = ... = \Delta w_{t-1} = 0$). We state this condition as $\Delta w_t \geq \Delta w_{t-n}^{-1}(0)$. Then, Δw_t is determined by the more stringent of condition (8), $\Delta w_t \geq \Delta w_{t-n}^{-1}(0)$, and $\Delta w_t \geq 0$. Though Δw_t depends on the history of periods with truthful reporting, the history-dependence on cash flows is again trivial, as the manager is dismissed when generating low cash flows in $t-n$ until $t-1$.

It remains to argue that the board always seeks truthful reporting in the manager's retirement period T. Observe, first, that if the board seeks truthful reporting in period T, we have

$$\begin{aligned}
\Delta w_T &= \max\left\{0, \Delta w_{T-n}^{-1}(0), \frac{c}{e_T\left(\theta_{T-1}^*\right)\Delta\theta}\right\} \\
w_{s,T} &= \max\left\{0, \theta_N \Delta w_T - \overline{U}\right\}.
\end{aligned}$$

Instead, if the board seeks no truthful reporting in the final period, it can no longer delay payments, and it must offer a payment that satisfies (9)–(10), while also making it optimal to set $\Delta w_t = 0$ in the immediately preceding periods without truthful reporting (in case

[26]Naturally, history dependence becomes important in the model's extension in Section 3.3.3.

there are such periods). Thus,

$$\begin{aligned} \Delta w_T &= \max\left\{0, \Delta w_{T-n}^{-1}(0), \frac{c}{e_T\left(\theta_{T-1}^*\right)\Delta\theta}\right\} \\ w_T &= \max\left\{0, \overline{U} - \theta_N \Delta w_T\right\}. \end{aligned}$$

Plugging into the manager's payoff U_T, we obtain that this payoff is identical in both cases. From (11)–(16), this implies that the manager's payoff is the same in all $t < T$ regardless of the truthful reporting policy in T. However, the board's payoff is higher with truthful reporting in T, making this policy optimal. **Q.E.D.**

Proof of Proposition 2. In what follows, Step 1 introduces some notation, and Step 2 argues to a contradiction.

Step 1. *Notation.* Let the likelihood that the manager retains her job in period t depending on whether there is truthful reporting and depending on the manager's fit from the previous and current period be defined as

$$\mathbf{e}_t = \begin{cases} \begin{array}{ccc} \text{if } t = 1 & \text{if } 1 < t \leq T & \\ \begin{pmatrix} e_1 & 0 \end{pmatrix} & \begin{pmatrix} e_t(\theta_G) & 0 \\ e_t(\theta_N) & 0 \end{pmatrix} & \text{if tr.rep.} \\ \begin{pmatrix} e_1\theta_G & (1-e_1)\theta_N \end{pmatrix} & \begin{pmatrix} e_t(\theta_G)\theta_G & (1-e_t(\theta_G))\theta_N \\ e_t(\theta_N)\theta_G & (1-e_t(\theta_N))\theta_N \end{pmatrix} & \text{otherwise} \end{array} \end{cases}.$$

The vector/matrix representation will be useful to minimize notation. Analogously, let the probability of replacement in any given period be defined as

$$\mathbf{p}_t = \begin{cases} \begin{array}{ccc} \text{if } t = 1 & \text{if } 1 < t \leq T & \\ 1 - e_1 & \begin{pmatrix} 1 - e_t(\theta_G) \\ 1 - e_t(\theta_N) \end{pmatrix} & \text{if tr.rep.} \\ \begin{array}{c} e_1(1-\theta_G) \\ +(1-e_1)(1-\theta_N) \end{array} & \begin{pmatrix} e_t(\theta_G)(1-\theta_G) + (1-e_t(\theta_G))(1-\theta_N) \\ e_t(\theta_N)(1-\theta_G) + (1-e_t(\theta_N))(1-\theta_N) \end{pmatrix} & \text{otherwise} \end{array} \end{cases}.$$

We can define now the (discounted) likelihood of replacement over the course of the entire potential employment relationship (for $T > 2$) as[27]

$$p(\mathbf{w}) := \delta \mathbf{p}_1 + \sum_{i=2}^{T-1} \delta^i \left(\Pi_{k=1}^{i-1} \mathbf{e}_k\right) \mathbf{p}_i + \delta^T \left(\Pi_{k=1}^{T-1} \mathbf{e}_k\right) \mathbf{1}, \tag{A.1}$$

which corresponds to $\mathbb{E}\left[\sum_{i=1}^{T} \delta^{i-1} \widetilde{q}_i\right] \delta = \mathbb{E}\left[\delta \widetilde{q}_1 + \sum_{i=2}^{T-1} \delta^j \widetilde{q}_j + \delta^T \widetilde{q}_T\right]$ in expression (4). Note that the replacement probability in the manager's retirement period T is $\mathbf{1} = (1; 1)$.

[27]For $T = 1$, we have $p(\mathbf{w}) = \delta$, and for $T = 2$, we have $p(\mathbf{w}) = \delta p_1 + \delta^2 e_1$.

Similarly, we can define the expected amount that the outgoing manager would be paid by the outside labor market from the period after her dismissal onwards (for $T > 2$) as[28]

$$h\left(\overline{U}, \mathbf{w}\right) := \delta \mathbf{p}_1 \sum_{j=2}^{T} \delta^{j-2} \overline{U} + \sum_{i=2}^{T-1} \left(\delta^i \left(\Pi_{k=1}^{i-1} \mathbf{e}_k \right) \mathbf{p}_i \sum_{j=i+1}^{T} \delta^{j-i-1} \overline{U} \right), \tag{A.2}$$

which corresponds to $\sum_{i=1}^{T} \delta^{i-1} \overline{U} - \mathbb{E}\left[\sum_{i=1}^{T} \delta^{i-1} q_i \overline{U}\right]$ in expression (4). It is convenient to rewrite this expression as

$$\begin{aligned}
h\left(\overline{U}, \mathbf{w}\right) &= \frac{\overline{U}}{1-\delta} \left(\delta \mathbf{p}_1 \left(1 - \delta^{T-1}\right) + \sum_{i=2}^{T-1} \delta^i \left(\Pi_{k=1}^{i-1} \mathbf{e}_k \right) \mathbf{p}_i \left(1 - \delta^{T-i}\right) \right) \\
&= \frac{\overline{U}}{1-\delta} \left(-\mathbf{p}_1 \delta^T + \underbrace{\delta \mathbf{p}_1 + \sum_{i=2}^{T-1} \delta^i \left(\Pi_{k=1}^{i-1} \mathbf{e}_k \right) \mathbf{p}_i}_{p(\mathbf{w}) - \delta^T \left(\Pi_{k=1}^{T-1} \mathbf{e}_k \right) \mathbf{1}} - \delta^T \sum_{i=2}^{T-1} \left(\Pi_{k=1}^{i-1} \mathbf{e}_k \right) \mathbf{p}_i \right) \\
&= \frac{\overline{U}}{1-\delta} \left(-\mathbf{p}_1 \delta^T + p(\mathbf{w}) - \delta^T \Pi_{k=1}^{T-1} \mathbf{e}_k \mathbf{1} - \delta^T \left(\mathbf{e}_1 - \Pi_{k=1}^{T-1} \mathbf{e}_k \right) \mathbf{1} \right) \\
&= \frac{\overline{U}}{1-\delta} \left(-\delta^T + p(\mathbf{w}) \right).
\end{aligned} \tag{A.3}$$

Denoting further $s(\mathbf{w}) := \mathbb{E}\left[\sum_{i=1}^{T} \delta^{i-1} q_i (x_i - c)\right]$ in the board's equilibrium expected payoff given by expression (4) and expressing $U_1(\mathbf{w})$ as $\nu_1(\mathbf{w}) + \sum_{j=1}^{T} \delta^{j-1} \overline{U}_j$ using (5), expression (4) can be stated as

$$V^* = -\nu_1(\mathbf{w}) - \sum_{j=1}^{T} \delta^{j-1} \overline{U}_j + s(\mathbf{w}) + h\left(\overline{U}, \mathbf{w}\right) + p(\mathbf{w}) V^*, \tag{A.4}$$

Using (A.3), we can simplify (A.4) to

$$V^* = \frac{s(\mathbf{w}) - v_1(\mathbf{w})}{1 - p(\mathbf{w})} - \frac{\overline{U}}{1-\delta}. \tag{A.5}$$

The functional dependence on $\mathbf{w}$ makes explicit that p, h, and s depend on the contract offered by the board and, hence, on the truthful reporting policy that goes along with this

[28] To be precise, in case the manager reveals that she is a bad fit, she receives $\overline{U}$ from the outside labor market in the period in which she is fired, but the firm hires a new manager for $\overline{U}$ for the remainder of the period, and the two terms cancel out in board's expected payoff. For $T = 1$, $h(\mathbf{w}) = 0$; for $T = 2$, $h(\mathbf{w}) = \delta p_1 \overline{U}$.

contract.

Step 2. *Optimality of Hiring an Older Manager or Abstaining from Incentives for Truthful Reporting.* We argue to a contradiction. Suppose that the board incentivizes truthful reporting in all periods. We have

$$
\begin{aligned}
s(\mathbf{w}) &= x + \left(\overline{\theta} + e_1\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c \\
&\quad + \left(\delta e_1 + \sum_{j=2}^{T-1} \delta^j e_1 \Pi_{i=2}^{j} e_i\left(\theta_G\right)\right)\left(x + \left(\overline{\theta} + e\left(\theta_G\right)\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c\right) \\
&= x + \left(\overline{\theta} + e_1\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c \\
&\quad + \left(\delta e_1 \frac{\left(1 - e\left(\theta_G\right)^{T-1}\delta^{T-1}\right)}{1 - e\left(\theta_G\right)\delta}\right)\left(x + \left(\overline{\theta} + e\left(\theta_G\right)\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c\right).
\end{aligned}
\tag{A.6}
$$

where, given the Markov structure, it is without loss to omit the subscripts of $e(\theta_G)$. Furthermore

$$
\begin{aligned}
1 - p(\mathbf{w}) &= 1 - \left(\delta\left(1 - e_1\right) + \sum_{j=2}^{T-1} \delta^j e_1 e\left(\theta_G\right)^{j-2}\left(1 - e\left(\theta_G\right)\right) + \delta^T e_1 e\left(\theta_G\right)^{T-2}\right) \\
&= (1-\delta)\left(1 + \frac{\delta e_1\left(1 - e\left(\theta_G\right)^{T-1}\delta^{T-1}\right)}{1 - e\left(\theta_G\right)\delta}\right)
\end{aligned}
\tag{A.7}
$$

Plugging in for $s(\mathbf{w})$, $p(\mathbf{w})$, as well as for $h(\mathbf{w})$ from (A.3) and $v_1(\mathbf{w})$ from (20), (A.5) becomes

$$
V^* = \frac{\left(\begin{array}{c} x + \left(\overline{\theta} + e_1\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c \\ + \left(\delta e_1 \frac{\left(1 - e(\theta_G)^{T-1}\delta^{T-1}\right)}{1 - e(\theta_G)\delta}\right)\left(x + \left(\overline{\theta} + e\left(\theta_G\right)\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c\right) \\ - \left(\frac{\theta_N c}{e_1 \Delta\theta} - \overline{U} + \frac{\delta - \delta^T}{1-\delta}\left(\frac{\theta_N c}{e(\theta_G)\Delta\theta} - \overline{U}\right)\right) \end{array}\right)}{(1-\delta)\left(1 + \frac{\delta e_1\left(1 - e(\theta_G)^{T-1}\delta^{T-1}\right)}{1 - e(\theta_G)\delta}\right)} - \frac{\overline{U}}{1-\delta}.
$$

It is now sufficient to show that increasing $T \to \infty$, can make the board's expected payoff V^* negative. This is the case if

$$
\begin{aligned}
& \frac{\theta_N c}{e\left(\theta_G\right)\Delta\theta} - \overline{U} \\
> \; & \frac{1-\delta}{\delta}\left(\begin{array}{c} \left(x + \left(\overline{\theta} + e_1\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c\right) - \frac{\theta_N c}{e_1 \Delta\theta} \\ + \frac{\delta e_1}{1 - e(\theta_G)\delta}\left(x + \left(\overline{\theta} + e\left(\theta_G\right)\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c - \overline{U}\right) \end{array}\right)
\end{aligned}
\tag{A.8}
$$

The key observation now is that the RHS of (A.8) decreases towards zero as δ increases towards one (in particular, note that $\frac{1-\delta}{\delta}$ and $\frac{\delta e_1}{1-e(\theta_G)\delta}$ do not cancel out). By contrast, the LHS of (A.8) is independent of δ and positive if (21) is not satisfied. Thus, for any parameter constellation, there is a threshold $\widehat{\delta}$ such that for $\delta > \widehat{\delta}$, this condition is satisfied. Hence, in these cases the board will deviate from a policy of pursuing truthful reporting in all periods and choosing T as high as possible. **Q.E.D.**

Proof of Proposition 3. Consider a period in which the manager is offered adequate severance pay to report truthfully in t. Substituting for $w_{s,t}$ in (19) from (14) and (16), we have

$$\begin{aligned} U_t\left(\theta_{t-1}, \mathbf{w}\right) &= w_{s,t} + \sum_{j=t}^{T} \delta^{j-t}\overline{U} + \max\left\{ e_t\left(\theta_{t-1}\right) \left(\begin{array}{c} w_t + \theta_G \Delta w_t + \delta U_{t+1}\left(\theta_G, \mathbf{w}^r\right) \\ -w_{s,t} - \sum_{j=t}^{T} \delta^{j-t}\overline{U} \end{array} \right) - c, 0 \right\} \\ &= \theta_N \Delta w_t + \delta U_{t+1}\left(\theta_N, \mathbf{w}\right) + \max\left\{ \frac{e_t\left(\theta_t\right)}{e_t\left(\theta^*_{t-1}\right)} - 1, 0 \right\} c \end{aligned} \tag{A.9}$$

The last term in (A.9) would be positive if the board incentivizes investments in firm-specific human capital also by $\theta^*_{t-1} = \theta_N$. This may occur if there is no truthful reporting in $t-1$.

Suppose, instead that there is no truthful reporting in t. Then, the manager's expected payoff is simply

$$U_t\left(\theta_{t-1}, \mathbf{w}^{nr}\right) = \max\left\{ -c + \delta \mathbb{E}_t\left[U^e_{t+1}\left(\theta_t, \mathbf{w}^{nr}\right) | \theta_{t-1} \right], U^e_{t+1}\left(\theta_N, \mathbf{w}^{nr}\right) \right\}, \tag{A.10}$$

where the max-operator captures that the manager may not invest in firm-specific human capital (in particular off the equilibrium path). From (A.9) and (A.10), it is immediate that reducing the manager's continuation payoff $U_{t+1}\left(\theta_t, \mathbf{w}\right)$ reduces her payoff in period t, which in turn reduces the manager's severance pay in t (cf. (16)).[29] **Q.E.D.**

The following straightforward result is helpful for the proof of Proposition A.1.

Lemma A.1 *Consider a contract offer $\mathbf{w}_n$ that gives the board an expected payoff of V^*_n and features n periods with truthful reporting. Compare this offer to an alternative $\mathbf{w}_{n+k}$ that gives the board V^*_{n+k} and features $n+k$ periods with truthful reporting. The attractiveness of*

[29] Using Proposition 1, we can plug in for all contract parameters to derive case-by-case conditions for when modifying a contract by abstaining from severance pay in a given period reduces the manager's rent in that period.

offer V_{n+k} *for the board increases in* $\overline{U}$ *and* Δx *if*

$$\frac{\partial}{\partial \overline{U}}\left(V_{n+k}^{*}-V_{n}^{*}\right) = \frac{\partial}{\partial \overline{U}}\left(\frac{-v_{1}\left(\mathbf{w}_{n+k}\right)}{1-p\left(\mathbf{w}_{n+k}\right)}-\frac{-v_{1}\left(\mathbf{w}_{n}\right)}{1-p\left(\mathbf{w}_{n}\right)}\right)>0 \tag{A.11}$$

$$\frac{\partial}{\partial \Delta x}\left(V_{n+k}^{*}-V_{n}^{*}\right) = \frac{\partial}{\partial \Delta x}\left(\frac{s\left(\mathbf{w}_{n+k}\right)}{1-p\left(\mathbf{w}_{n+k}\right)}-\frac{s\left(\mathbf{w}_{n}\right)}{1-p\left(\mathbf{w}_{n}\right)}\right)>0. \tag{A.12}$$

Proof of Lemma A.1. The claim follows by standard comparative statics arguments. The increasing differences in (A.11) and (A.12) imply that the V_{n+k}^{*} offer becomes increasingly more attractive for the board as $\overline{U}$ and Δx increase. To obtain the equalities in (A.11) and (A.12), plug in for V^{*} from (A.5). **Q.E.D.**

Proof of Proposition 4. Suppose that the manager lives for at most two periods and that the first-best condition (21) is not satisfied, in which case $w_{s,t}>0$ for all t. If the board employs the manager for one period only, it always seeks truthful reporting (Proposition 1). Thus, consider the case in which $T=2$. We start by deriving the separate components of (A.5) and then plug into (A.11) and (A.12).

In the final period, the board always seeks truthful reporting by Proposition 1. By (17), this implies that

$$\begin{aligned} U_{2}\left(\theta_{1}, \mathbf{w}\right) &= \max\left\{e_{2}\left(\theta_{1}\right)\left(w_{2}+\theta_{G} \Delta w_{2}\right)+\left(1-e_{2}\left(\theta_{1}\right)\right)\left(w_{s,2}+\overline{U}\right)-c, w_{s,2}+\overline{U}\right\} \\ &= \max\left\{\left(\theta_{N}+e_{2}\left(\theta_{1}\right) \Delta \theta\right) \Delta w_{2}-c, \theta_{N} \Delta w_{2}\right\}. \end{aligned}$$

where we use that $w_{s,2}=\theta_{N} \Delta w_{2}-\overline{U}$ and $w_{2}=0$. Furthermore, we have that $\Delta w_{2}= \max\left\{\Delta w_{1}^{-1}(0), \frac{c}{e_{2}\left(\theta_{t-1}^{*}\right) \Delta \theta}\right\}$, where $\Delta w_{1}^{-1}(0)$ is the minimum bonus in $t=2$ such that the manager invests in firm-specific human capital in $t=1$ even though $w_{1}=\Delta w_{1}=0$, i.e., $\Delta w_{1}^{-1}(0)$ is defined by

$$\begin{aligned} 0 &= \frac{c}{e_{1} \Delta \theta}+\delta \frac{U_{2}^{e}\left(\theta_{N}, \mathbf{w}\right)-U_{2}^{e}\left(\theta_{G}, \mathbf{w}\right)}{\Delta \theta} && \text{(A.13)} \\ &= \frac{c}{e_{1} \Delta \theta}+\delta \frac{\left(\begin{array}{c} \theta_{N} \max\left\{\left(\theta_{N}+e_{2}\left(\theta_{N}\right) \Delta \theta\right) \Delta w_{2}-c, \theta_{N} \Delta w_{2}\right\}+\left(1-\theta_{N}\right) \overline{U} \\ -\theta_{G} \max\left\{\left(\theta_{N}+e_{2}\left(\theta_{G}\right) \Delta \theta\right) \Delta w_{2}-c, \theta_{N} \Delta w_{2}\right\}-\left(1-\theta_{G}\right) \overline{U} \end{array}\right)}{\Delta \theta} && \text{(A.14)} \end{aligned}$$

To avoid the max-operator in (A.14), let $\widetilde{e}_{2}\left(\theta_{N}\right)=e_{2}\left(\theta_{N}\right)$ (and $\widetilde{c}=c$) if the manager invests in firm-specific human capital in $t=2$ if her fit in $t=1$ is θ_{N}, and $\widetilde{e}_{2}\left(\theta_{N}\right)=\widetilde{c}=0$ otherwise.[30] To further simplify notation, let $e\left(\theta_{G}\right)=e_{G}$. From (A.14), we obtain that

[30]Note that the board will always offer sufficient incentives for θ_{G} to invest in firm-specific human capital,

$\Delta w_1^{-1}(0) = \frac{\frac{c}{\delta e_1 \Delta\theta} + \frac{\theta_G c - \theta_N \widetilde{c}}{\Delta\theta} + \overline{U}}{\delta(e_G\theta_G + \theta_N(1-\widetilde{e}_N))}$. Consider, first, the case in which $\Delta w_2 = \Delta w_1^{-1}(0)$. We have

$$\begin{aligned} \nu_1(\mathbf{w}^{nr}) &= U_1(\mathbf{w}^{nr}) - \sum_{j=1}^{2} \delta^{j-1}\overline{U} = -c + \delta\mathbb{E}_1\left[U_2^e(\theta_t, \mathbf{w})\right] - \overline{U}(1+\delta) \\ &= -c + \delta\left[U_2^e(\theta_N, \mathbf{w}) + e_1\left(U_{t+1}^e(\theta_G, \mathbf{w}) - U_2^e(\theta_N, \mathbf{w})\right)\right] - \overline{U}(1+\delta) \\ &= \delta\left(\theta_N\left((\widetilde{e}_N\theta_G + (1-\widetilde{e}_N)\theta_N)\frac{\frac{c}{\delta e_1\Delta\theta} + \frac{\theta_G c - \theta_N\widetilde{c}}{\Delta\theta} + \overline{U}}{\delta(e_G\theta_G + \theta_N(1-\widetilde{e}_N))} - c\right) + (1-\theta_N)\overline{U}\right) \\ &\quad -\overline{U}(1+\delta) \\ &= A_1 + \left(\frac{\theta_N(\widetilde{e}_N\theta_G + (1-\widetilde{e}_N)\theta_N)}{(\theta_G e_G + (1-\widetilde{e}_N)\theta_N)} - 1 - \delta\theta_N\right)\overline{U}. \end{aligned} \tag{A.15}$$

where A_1 stands for terms that do not depend on $\overline{U}$; the third equality follows after using from (A.13) that $U_2^e(\theta_N, \mathbf{w}) - U_2^e(\theta_G, \mathbf{w}) = \frac{c}{e_1\delta}$ and after plugging for $\Delta w_2 = \Delta w_1^{-1}(0)$ (cf. (A.14)).

If the board seeks truthful reporting in the first period, then, by (20), the manager's rent is simply

$$\nu_1(\mathbf{w}^r) = \frac{\theta_N}{e_1\Delta\theta}c - \overline{U} + \delta\left(\frac{\theta_N}{e_G\Delta\theta}c - \overline{U}\right),$$

By plugging $\nu_1(\mathbf{w}^r)$ and $\nu_1(\mathbf{w}^{nr})$ into (A.11), we obtain that

$$\begin{aligned} &\frac{\partial}{\partial\overline{U}}\left(\frac{-\nu_1(\mathbf{w}^r)}{1-p(\mathbf{w}^r)} - \frac{-\nu_1(\mathbf{w}^{nr})}{1-p(\mathbf{w}^{nr})}\right) \\ &= \frac{1+\delta}{1-\delta(1-e_1)-\delta^2 e_1} - \frac{-\left(\frac{\theta_N(\widetilde{e}_N\theta_G + (1-\widetilde{e}_N)\theta_N)}{(\theta_G e_G + (1-\widetilde{e}_N)\theta_N)} - 1 - \delta\theta_N\right)}{1-\delta(1-e_1\theta_G - (1-e_1)\theta_N) - \delta^2(e_1\theta_G + (1-e_1)\theta_N)} \\ &= \frac{(1+\delta)\delta(\mathbb{E}_1\theta - e_1) + \left(\theta_N\frac{(\widetilde{e}_N\theta_G + (1-\widetilde{e}_N)\theta_N)}{(\theta_G e_G + (1-\widetilde{e}_N)\theta_N)} + \delta(1-\theta_N)\right)(1+\delta e_1)}{(1-\delta)(1+\delta e_1)(1+\delta\mathbb{E}_1\theta)} \\ &> \delta\frac{(1+\delta)(\mathbb{E}_1\theta - e_1) + (1-\theta_N)(1+\delta e_1)}{(1-\delta)(1+\delta e_1)(1+\delta\mathbb{E}_1\theta)} \end{aligned} \tag{A.16}$$

where $\mathbb{E}_1\theta \equiv e_1\theta_G + (1-e_1)\theta_N$. After some transformations, the last expression becomes

$$\frac{\delta(\theta_N + \theta_G e_1 - 2\theta_N e_1) + (\theta_G - \theta_N)e_1 + 1 - e_1}{(1-\delta)(1+\delta e_1)(1+\delta\mathbb{E}_1\theta)} > 0.$$

Hence, (A.16) is positive.

Consider, second, the case in which $\Delta w_2 = \frac{c}{e_2\left(\theta_{t-1}^*\right)\Delta\theta}$. As in (A.15), we have then

$$\begin{aligned} \nu_1(\mathbf{w}^{nr}) &= -c + \delta\left[U_2^e(\theta_N, \mathbf{w}) + e_1\left(U_{t+1}^e(\theta_G, \mathbf{w}) - U_2^e(\theta_N, \mathbf{w})\right)\right] - \overline{U}(1+\delta) \\ &= A_2 + (\delta e_1\Delta\theta - 1 - \delta\theta_N)\overline{U} \end{aligned}$$

as, otherwise, it would be better off hiring a new manager where A_2 stands for terms that do not depend on $\overline{U}$. Following the same steps as above, we have again $\frac{\partial}{\partial \overline{U}}\left(\frac{-\nu_1(\mathbf{w}^r)}{1-p(\mathbf{w}^r)}-\frac{-\nu_1(\mathbf{w}^{nr})}{1-p(\mathbf{w}^{nr})}\right)>0$. By Lemma A.1, this proves the claim.

Approaching (A.12) similarly, we have

$$
\begin{aligned}
&\frac{s\left(\mathbf{w}^r\right)}{1-p\left(\mathbf{w}^r\right)}-\frac{s\left(\mathbf{w}^{nr}\right)}{1-p\left(\mathbf{w}^{nr}\right)} \\
=\;& \frac{x+\left(e_1\theta_G+\left(1-e_1\right)\overline{\theta}\right)\Delta x-c+\delta e_1\left(x+\left(e_G\theta_G+\left(1-e_G\right)\overline{\theta}\right)\Delta x-c\right)}{\left(1-\delta\right)\left(1+\delta e_1\right)} \\
&-\frac{x+\left(e_1\theta_G+\left(1-e_1\right)\theta_N\right)\Delta x-c+\delta\begin{pmatrix} e_1\theta_G\left(x+\left(e_G\theta_G+\left(1-e_G\right)\overline{\theta}\right)\Delta x-c\right) \\ +\left(1-e_1\right)\theta_N\left(x+\left(\widetilde{e}_N\theta_G+\left(1-\widetilde{e}_N\right)\overline{\theta}\right)\Delta x-c\right)\end{pmatrix}}{\left(1-\delta\right)\left(1+\delta E_1\theta\right)}.
\end{aligned}
$$

Defining $\Delta e = e_G - \widetilde{e}_N$, after some transformations, the terms dependent on Δx become

$$
\frac{\Delta x}{\left(1-\delta\right)\left(1+e_1\delta\right)\left(1+\delta\mathbb{E}_1\theta\right)}\begin{pmatrix}\delta\left(\theta_G-\overline{\theta}\right)\left(\left(e_G-e_1\right)\left(e_1-\mathbb{E}_1\theta\right)+\theta_N\Delta e\left(1-e_1\right)\left(1+\delta e_1\right)\right) \\ +\left(\overline{\theta}-\theta_N\right)\left(1-e_1\right)\left(1+\delta e_1\right)\end{pmatrix}
$$

which is strictly positive as

$$
\begin{aligned}
& \left(e_G-e_1\right)\left(e_1-\mathbb{E}_1\theta\right)+\theta_N\Delta e\left(1-e_1\right)\left(1+\delta e_1\right) \\
> \;& \left(e_G-e_1\right)\left(e_1-e_1\theta_G-\left(1-e_1\right)\theta_N+\theta_N\left(1-e_1\right)\right) \\
= \;& \left(e_G-e_1\right)\left(e_1-e_1\theta_G\right)>0
\end{aligned}
$$

Hence, expression (A.12) is positive, proving also the second statement. **Q.E.D.**

Proof of Proposition 5. Using (A.4) to express V^*, we have to show that

$$
\frac{\partial}{\partial \overline{U}}\left(V^*\left(\mathbf{w}'\right)-V^*\left(\mathbf{w}\right)\right)=\frac{\frac{\partial}{\partial \overline{U}}h\left(\mathbf{w}'\right)\left(1-p\left(\mathbf{w}\right)\right)-\frac{\partial}{\partial \overline{U}}h\left(\mathbf{w}\right)\left(1-p\left(\mathbf{w}'\right)\right)}{\left(1-p\left(\mathbf{w}\right)\right)\left(1-p\left(\mathbf{w}'\right)\right)}>0 \qquad \text{(A.17)}
$$

where the offer $\mathbf{w}'$ is made to a manager with time to retirement $T' = T + 1$. To obtain the equality in (A.17), we use that when the first-best condition (21) does not hold, the manager's expected payoff $U_1(\mathbf{w})$ is independent of $\overline{U}$ (cf. (A.9)). The increasing difference in (A.17) will imply that the T'-offer becomes increasingly more attractive as $\overline{U}$ increases.

Recalling from (A.3) that $h(\mathbf{w}) = \frac{\overline{U}}{1-\delta}(-\delta^T + p(\mathbf{w})) = \frac{\overline{U}}{1-\delta}(1-\delta^T + p(\mathbf{w}) - 1)$ and plugging in from (A.7), we can express the numerator in (A.17) as

$$
\begin{aligned}
& \frac{1}{1-\delta}\left(\left(1-\delta^{T+1}+p\left(\mathbf{w}'\right)-1\right)\left(1-p\left(\mathbf{w}\right)\right)-\left(1-\delta^{T}+p\left(\mathbf{w}\right)-1\right)\left(1-p\left(\mathbf{w}'\right)\right)\right) \\
= & \frac{1}{1-\delta}\left(\left(1-\delta^{T+1}\right)\left(1-p\left(\mathbf{w}\right)\right)-\left(1-\delta^{T}\right)\left(1-p\left(\mathbf{w}'\right)\right)\right) \\
= & \left(1-\delta^{T+1}\right)\left(1+\frac{\delta e_1\left(1-\left(\delta e_G\right)^{T-1}\right)}{1-\delta e_G}\right)-\left(1-\delta^{T}\right)\left(1+\frac{\delta e_1\left(1-\left(\delta e_G\right)^{T}\right)}{1-\delta e_G}\right)
\end{aligned}
$$

which after some transformations becomes

$$
\begin{aligned}
& \frac{\delta^T}{1-\delta e_G}\left(\left(1-\delta\right)\left(1-\delta e_G\right)+e_1\left(\delta\left(1-\delta\right)+\left(e_G\right)^{T-1}\left(\delta^{T+1}\left(1-e_G\right)-\left(1-\delta e_G\right)\right)\right)\right) \\
\geq & \ 0.
\end{aligned} \tag{A.18}
$$

To see the last inequality, observe that expression (A.18) is positive if the term in brackets following e_1 is positive. If it is negative, expression (A.18) would decrease in e_1. Thus, at its minimum for $e_1 = e_G$, that value would be

$$
\delta^T\left(\frac{\left(1-\delta+\left(e_G\right)^T\delta^{T+1}\left(1-e_G\right)\right)}{1-\delta e_G}-\left(e_G\right)^T\right).
$$

The sign of this expression is the same as the sign of the term in brackets. The minimum of that term is zero, which is obtained for $\delta = 1$.[31] Hence, for any T, $\delta \in [0,1]$, $e_G \in [0,1]$, $e_1 \in [0, e_G]$, (A.18) is (weakly) positive, and it is strictly positive for $e_G < 1$ and $\delta < 1$, implying that we have strictly increasing differences in (A.17). By Lemma A.1, this proves the claim.

Next, we argue that

$$
\frac{\partial}{\partial\Delta x}\left(V^*\left(\mathbf{w}'\right)-V^*\left(\mathbf{w}\right)\right)=\frac{\frac{\partial}{\partial\Delta x}s\left(\mathbf{w}'\right)\left(1-p\left(\mathbf{w}\right)\right)-\frac{\partial}{\partial\Delta x}s\left(\mathbf{w}\right)\left(1-p\left(\mathbf{w}'\right)\right)}{\left(1-p\left(\mathbf{w}\right)\right)\left(1-p\left(\mathbf{w}'\right)\right)}>0. \tag{A.19}
$$

[31] We have

$$
\frac{\partial}{\partial\delta}\left(\frac{\left(1-\delta+\left(\delta e_G\right)^T\delta\left(1-e_G\right)\right)}{1-\delta e_G}-\left(e_G\right)^T\right)=\left(1-e_G\right)\frac{T\left(\delta e_G\right)^T\left(1-\delta e_G\right)+\left(\delta e_G\right)^T-1}{\left(1-\delta e_G\right)^2}
$$

This term is nonpositive, as the maximum value (of zero) of the numerator is obtained for $\delta e_G = 1$.

Observe, first, that from (A.6) we have

$$\frac{\partial}{\partial \Delta x} s(\mathbf{w}) = \left(\overline{\theta} + e_1\left(\theta_G - \overline{\theta}\right)\right) + \delta e_1 \frac{\left(1 - e\left(\theta_G\right)^{T-1} \delta^{T-1}\right)}{1 - e\left(\theta_G\right)\delta} \left(\overline{\theta} + e\left(\theta_G\right)\left(\theta_G - \overline{\theta}\right)\right).$$

Plugging in for $p(\mathbf{w})$, (A.19) becomes

$$\begin{aligned}
& \left(\left(\overline{\theta} + e_1\left(\theta_G - \overline{\theta}\right)\right) + \delta e_1 \frac{\left(1 - e_G^T \delta^T\right)}{1 - e_G\delta} \left(\overline{\theta} + e_G\left(\theta_G - \overline{\theta}\right)\right) \right) (1-\delta) \left(1 + \frac{\delta e_1 \left(1 - e_G^{T-1}\delta^{T-1}\right)}{1 - e_G\delta} \right) \\
& - \left(\left(\overline{\theta} + e_1\left(\theta_G - \overline{\theta}\right)\right) + \delta e_1 \frac{\left(1 - e_G^{T-1} \delta^{T-1}\right)}{1 - e_G\delta} \left(\overline{\theta} + e_G\left(\theta_G - \overline{\theta}\right)\right) \right) (1-\delta) \left(1 + \frac{\delta e_1 \left(1 - e_G^{T}\delta^{T}\right)}{1 - e_G\delta} \right) \\
= \; & \delta^T e_1 \left(e_G - e_1\right) e^{T-1} (1-\delta) \left(\theta_G - \overline{\theta}\right) > 0
\end{aligned}$$

proving the claim. **Q.E.D.**

Proof of Proposition 6. Recall that from expressions (A.3) and (A.4), we can express

$$V^*(\mathbf{w}) = \frac{s(\mathbf{w}) + \frac{\overline{U}}{1-\delta}\left(p(\mathbf{w}) - \delta^T\right) - U_1(\mathbf{w})}{1 - p(\mathbf{w})}.$$

Hence, $\frac{\partial}{\partial \overline{U}} V^*(\mathbf{w}) > 0$ as long as $\left(p(\mathbf{w}) - \delta^T\right) > (1-\delta) \frac{\partial}{\partial \overline{U}} U_1(\mathbf{w})$. This is trivially satisfied if the board stimulates truthful reporting in all periods and the first-best condition (21) is not satisfied. In this case, $U_1(\mathbf{w}) = \frac{\theta_N c}{e_1 \Delta\theta} + \sum_{j=2}^{T} \delta^{j-1} \frac{\theta_N c}{e_j(\theta_G)\Delta\theta}$ (cf. (20)), which is independent of $\overline{U}$. **Q.E.D.**

Appendix B For Online Publication: Supplementary Material

Binding constraints in Proposition 1. In what follows, we take the board's truthful reporting policy as given and analyze which of the conditions (7)–(10), $w_t, \Delta w_t, w_{s,t} \geq 0$ are binding when minimizing the manager's period one payoff $U_1(\mathbf{w})$. The proof proceeds by initially assuming that, to minimize $U_1(\mathbf{w})$, the board minimizes U_t in all periods t. Proposition 3 shows that minimizing U_t helps to minimize U_{t-1} and, thus, preceding recursively, one minimizes the manager's expected payoff all the way to period one.

Lemma B.1 *(i) If the board does not seek truthful reporting in period $t < T$, it is optimal to set $w_{s,t} = 0$ and $\Delta w_t = 0$; w_t is determined by having the more stringent of conditions (10) and $w_t \geq 0$. (ii) If the board seeks truthful reporting in period t, $w_{s,t}$ is determined by condition (7) subject to $w_{s,t} \geq 0$; Δw_t is determined by condition (8) if the board seeks truthful reporting in $t-1$. Without truthful reporting from $t-n$ to $t-1$ ($n \geq 1$), Δw_t is determined by the more stringent of condition (8) and the analogue of condition (9) for the preceding n periods, subject to $\Delta w_t \geq 0$.*

Proof of Lemma B.1. Assume initially that the board offers incentives for truthful reporting in $t-1$ (if there is such a period). This leads to four main cases depending on whether the board offers incentives for truthful reporting in period t, and $t+1$. Towards the end, the proof considers the case in which the board does not offer incentives for truthful reporting in $t-1$. The Lemma is shown by induction by arguing first that it is always satisfied in period 1. It is then argued that if the conditions stated in the Lemma are satisfied in t, then the $t+1$ analogue of these conditions must also be satisfied.

Truthful reporting in period t**.** The induction hypothesis for this case is that Δw_t, w_t, and $w_{s,t}$ are given by (14)–(16). Recall that $e_1(\theta_{t-1}) = e_1$ in period $t = 1$. Setting $\{w_1, \Delta w_1, w_{s,1}\}$ to their minimal values maximizes the board's expected payoff, as it minimizes the manager's payoff, without affecting her incentives in the following periods. Thus, $w_{s,1}$ is determined by making (7) binding. Using this, we see that setting $w_1 = 0$ relaxes (7), while not affecting (8). Finally, Δw_1 is determined by the more stringent condition of (8) and $\Delta w_1 \geq 0$.

Case 1: Truthful reporting in periods t ***and*** $t+1$**.** Consistent with the notation in Proposition 4, define $\widetilde{e}_t(\theta_{t-1}) = e_t(\theta_{t-1})$ ($\widetilde{c} = c$) if the manager, whose fit in $t-1$ is θ_{t-1}, invests in firm-specific human capital in t, and $\widetilde{e}_t(\theta_{t-1}) = 0$ ($\widetilde{c} = 0$) otherwise. This notation takes into account that the contract may not provide sufficient incentives for such investment to a manager whose fit in the preceding period is θ_N. The manager's expected payoff in period t is then

$$U_t(\theta_{t-1},\mathbf{w}) = \widetilde{e}_t(\theta_{t-1})(w_t + \theta_G \Delta w_t + \delta U_{t+1}(\theta_G,\mathbf{w})) + (1-\widetilde{e}(\theta_{t-1}))\left(w_{s,t} + \sum_{j=t}^{T} \delta^{j-t}\overline{U}_j\right) - \widetilde{c}.$$

Using the induction hypothesis (14)–(16) to plug into w_t, Δw_t, and $w_{s,t}$, this payoff becomes

$$\begin{aligned}&\widetilde{e}_t(\theta_{t-1})\left(w_t + \frac{\theta_G c}{e_t(\theta_G)\Delta\theta} + \frac{\theta_G \delta U_{t+1}(\theta_N,\mathbf{w}) - \theta_N \delta U_{t+1}(\theta_G,\mathbf{w})}{\Delta\theta}\right)\\ &+ (1-\widetilde{e}_t(\theta_{t-1}))\max\left\{\sum_{j=t}^{T}\delta^{j-t}\overline{U}_j, \frac{\theta_N c}{e_t(\theta_G)\Delta\theta} + \delta\frac{\theta_G U_{t+1}(\theta_N,\mathbf{w}) - \theta_N U_{t+1}(\theta_G,\mathbf{w})}{\Delta\theta}\right\} - \widetilde{c}.\end{aligned}$$

if the zero lower bounds of Δw_t and $w_{s,t}$ are not binding (we consider these cases at the end). In what follows, it is shown that choosing $\{w_{s,t+1}, w_{t+1}, \Delta w_{t+1}\}$ as dictated in Proposition 1 minimizes

$$\theta_G U_{t+1}(\theta_N,\mathbf{w}) - \theta_N U_{t+1}(\theta_G,\mathbf{w}), \tag{B.1}$$

and, thus, minimizes $U_t(\theta_{t-1},\mathbf{w})$.

The condition that the manager invests in firm-specific human capital in $t+1$ if her fit in t is θ_t^* is (this is the $t+1$ analogue of (8))

$$w_{t+1} + \theta_t \Delta w_{t+1} + \delta U_{t+2}(\theta_G,\mathbf{w}) - \frac{c}{e_{t+1}(\theta_t^*)} \geq w_{s,t+1} + \sum_{j=t+1}^{T}\delta^{j-t-1}\overline{U}_j. \tag{B.2}$$

Truthful reporting in period $t+1$ would further require that

$$w_{s,t+1} + \sum_{j=t+1}^{T}\delta^{j-t-1}\overline{U}_j \geq w_{t+1} + \theta_N \Delta w_{t+1} + \delta U_{t+2}(\theta_N,\mathbf{w}). \tag{B.3}$$

In what follows, it is shown that (B.2), (B.3), and $w_{t+1} \geq 0$ will be binding.

To find the contract parameters $\{w_{s,t+1}, w_{t+1}, \Delta w_{t+1}\}$ that minimize (B.1), subject to (B.2), (B.3), and $w_{s,t+1}, w_{t+1}, \Delta w_{t+1} \geq 0$, we apply Kuhn Tucker's Theorem. Define the function

$$
\begin{aligned}
& L_1(\mathbf{w},\Lambda) \\
= & -\left(\theta_G \widetilde{e}_{t+1}(\theta_N) - \theta_N e_{t+1}(\theta_G)\right)\left(w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2}(\theta_G,\mathbf{w})\right) \\
& -\left(\theta_G\left(1-\widetilde{e}_{t+1}(\theta_N)\right) - \theta_N\left(1-e_{t+1}(\theta_G)\right)\right)\left(w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1}\overline{U}_j\right) + \theta_G \widetilde{c} - \theta_N c \\
& +\lambda\left(w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2}(\theta_G,\mathbf{w}) - \frac{c}{e_{t+1}(\theta_t^*)} - w_{s,t+1} - \sum_{j=t+1}^{T} \delta^{j-t-1}\overline{U}_j\right) \\
& +\mu\left(w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1}\overline{U}_j - w_{t+1} - \theta_N \Delta w_{t+1} - \delta U_{t+2}(\theta_N,\mathbf{w})\right) \\
& +\kappa w_{s,t+1} + \rho w_{t+1} + \chi \Delta w_{t+1}
\end{aligned}
$$

where the first two lines correspond to the negative of (B.1) (as the objective is to minimize (B.1)), and $\Lambda = \{\lambda, \mu, \kappa, \rho, \chi\}$ is the set of weakly positive Kuhn Tucker multipliers. Taking the first order conditions

$$
\begin{aligned}
\frac{\partial L_1(\mathbf{w},\Lambda\theta)}{\partial w_{s,t+1}} &= 0 = -\left(\theta_G\left(1-\widetilde{e}_{t+1}(\theta_N)\right) - \theta_N\left(1-e_{t+1}(\theta_G)\right)\right) - \lambda + \mu + \kappa \\
\frac{\partial L_1(\mathbf{w},\Lambda)}{\partial \Delta w_{t+1}} &= 0 = -\left(\theta_G \widetilde{e}_{t+1}(\theta_N) - \theta_N e_{t+1}(\theta_G)\right)\theta_G + \lambda\theta_G - \mu\theta_N + \chi \\
\frac{\partial L_1(\mathbf{w},\Lambda)}{\partial w_{t+1}} &= 0 = -\left(\theta_G \widetilde{e}_{t+1}(\theta_N) - \theta_N e_{t+1}(\theta_G)\right) + \lambda - \mu + \rho,
\end{aligned}
$$

we obtain from the second and third conditions that $\theta_G \rho = \mu \Delta\theta + \chi$. From the first and third conditions, we further have $\kappa + \rho = \Delta\theta$. Assuming now that $\Delta w_{t+1} \geq 0$ and $w_{s,t+1} \geq 0$ are not binding, i.e., $\chi = 0$ and $\kappa = 0$, we have: $\rho = \Delta\theta$, $\mu = \theta_G$, and $\lambda = \theta_G \widetilde{e}_{t+1}(\theta_N) + \theta_N\left(1 - e_{t+1}(\theta_G)\right) > 0$. Thus, $\rho, \mu, \lambda > 0$, imply that the board minimizes $w_{s,t+1}$, Δw_{t+1}, and w_{t+1}, subject to the binding constraints (B.2), (B.3), and $w_{t+1} \geq 0$, as was to be shown. Since the board knows that the manager's fit in t is θ_G, this implies that the board will choose to satisfy (B.2) for $\theta_t^* = \theta_G$.[32] For completeness, using that (B.3) is binding, note that we obtain

$$
\begin{aligned}
w_{s,t+1} &= \theta_N \Delta w_{t+1} + \delta U_{t+2}(\theta_N,\mathbf{w}) - \sum_{j=t+1}^{T} \delta^{j-t-1}\overline{U}_j \\
\Delta w_{t+1} &= \frac{c}{e_{t+1}(\theta_G)\,\Delta\theta} + \frac{\delta U_{t+2}(\theta_N,\mathbf{w}) - \delta U_{t+2}(\theta_G,\mathbf{w})}{\Delta\theta},
\end{aligned}
$$

[32] Note that this implies that the manager has no incentives to invest in firm-specific human capital if $\theta_t = \theta_N$, i.e., $\widetilde{e}_{t+1}(\theta_N) = 0$ and $\widetilde{c} = 0$.

which are the $t+1$ analogues of (14)–(16).

Consider, now, the case in which the zero lower bound of Δw_t is binding. Then the manager's expected payoff is simply

$$\widetilde{e}_t\left(\theta_{t-1}\right)\delta U_{t+1}\left(\theta_G,\mathbf{w}\right)+\left(1-\widetilde{e}\left(\theta_{t-1}\right)\right)U_{t+1}\left(\theta_N,\mathbf{w}\right)-\widetilde{c},$$

which is clearly minimized when w_{t+1}, Δw_{t+1}, and $\Delta w_{s,t+1}$ are minimized subject to the more stringent of (B.2), (B.3), and w_{t+1},$\Delta w_{t+1},\Delta w_{s,t+1}\geq 0$ as was to be shown. Finally, if $w_{s,t}=0$ is binding (or respectively $w_{s,t+1}\geq 0$ is binding) the manager extracts no rent in the respective period. For this case, we argue in the main text that the manager extracts no rent in all preceding periods until period $t=1$, making the proposed contract optimal. The same arguments will apply to Case 3 below.[33]

Case 2: Truthful reporting in period t and no truthful reporting in period $t+1$. Similar to case 1, we can show that the board would like to minimize (B.1). The difference is that, absent truthful reporting in $t+1$, the manager's payoff in that period is

$$U_{t+1}\left(\theta_t,\mathbf{w}\right)=\max\left\{\begin{array}{c} w_{t+1}+\left(\theta_N+e_{t+1}\left(\theta_t\right)\Delta\theta\right)\Delta w_{t+1}-c+\delta\mathbb{E}_{\theta_t}\left[U^e_{t+2}\left(\theta_{t+1},\mathbf{w}\right)\right] \\ w_{t+1}+\theta_N\Delta w_{t+1}+\delta U^e_{t+2}\left(\theta_N,\mathbf{w}\right)\end{array}\right\} \tag{B.4}$$

The manager invests in firm-specific human capital in $t+1$ when her fit in t is θ^*_t if

$$w_{t+1}+\left(\theta_N+e_{t+1}\left(\theta^*_t\right)\Delta\theta\right)\Delta w_{t+1}+\delta\mathbb{E}_{\theta_t}\left[U^e_{t+2}\left(\theta_{t+1},\mathbf{w}\right)\right]-c\geq w_{t+1}+\theta_N\Delta w_{t+1}+\delta U^e_{t+2}\left(\theta_N,\mathbf{w}\right).$$

which can be restated as

$$\Delta w_{t+1}\geq\frac{c}{e_{t+1}\left(\theta^*_t\right)\Delta\theta}+\delta\frac{U^e_{t+2}\left(\theta_N,\mathbf{w}\right)-U^e_{t+2}\left(\theta_G,\mathbf{w}\right)}{\Delta\theta}. \tag{B.5}$$

In analogy to (10), we further need to satisfy the interim participation constraint in $t+1$

$$w_{t+1}+\theta_N\Delta w_{t+1}+\delta U^e_{t+2}\left(\theta_N,\mathbf{w}\right)\geq w_{s,t+1}+\sum_{j=t+1}^{T}\delta^{j-t-1}\overline{U}_j. \tag{B.6}$$

[33] It is straightforward to verify that $\Delta w_t,\Delta w_{t+1}>0$ when the constrain that the manager invests in firm-specific human capital in $t+1$ is binding.

Hence, when the zero lower bound of Δw_t is not binding, we need to minimize (B.1), subject to (B.5), (B.6), and feasibility. Define[34]

$$\begin{aligned} L_2 \;=\;& -\theta_G\left(w_{t+1} + \left(\theta_N + \widetilde{e}_{t+1}\left(\theta_N\right)\Delta\theta\right)\Delta w_{t+1} + \delta\widetilde{\mathbb{E}}_{\theta_N}\left[U^e_{t+2}\left(\theta_{t+1},\mathbf{w}\right)\right] - \widetilde{c}\right) \\ & +\theta_N\left(w_{t+1} + \left(\theta_N + e_{t+1}\left(\theta_G\right)\Delta\theta\right)\Delta w_{t+1} + \delta\mathbb{E}_{\theta_G}\left[U^e_{t+2}\left(\theta_{t+1},\mathbf{w}\right)\right] - c\right) \\ & +\lambda\left(\Delta w_{t+1} - \frac{c}{e_{t+1}\left(\theta^*_t\right)\Delta\theta} - \delta\frac{U^e_{t+2}\left(\theta_N,\mathbf{w}\right) - U^e_{t+2}\left(\theta_G,\mathbf{w}\right)}{\Delta\theta}\right) \\ & +\mu\left(w_{t+1} + \theta_N\Delta w_{t+1} + \delta U^e_{t+2}\left(\theta_N,\mathbf{w}\right) - w_{s,t+1} - \sum_{j=t+1}^{T}\delta^{j-t-1}\overline{U}_j\right) \\ & +\kappa w_{s,t+1} + \rho w_{t+1} + \chi\Delta w_{t+1} \end{aligned}$$

giving the first-order conditions

$$\begin{aligned} \frac{\partial L_2}{\partial w_{t+1}} = 0 \;&=\; -\Delta\theta + \mu + \rho \\ \frac{\partial L_2}{\partial w_{s,t+1}} = 0 \;&=\; -\mu + \kappa \\ \frac{\partial L_2}{\partial \Delta w_{t+1}} = 0 \;&=\; -\left(\theta_G\left(\theta_N + \widetilde{e}_{t+1}\left(\theta_N\right)\Delta\theta\right) - \theta_N\left(\theta_N + e_{t+1}\left(\theta_G\right)\Delta\theta\right)\right) + \lambda + \theta_N\mu + \chi \\ &=\; -\left(\theta_N - e_{t+1}\left(\theta_G\right)\theta_N + \widetilde{e}_{t+1}\left(\theta_N\right)\theta_G\right)\Delta\theta + \lambda + \theta_N\mu + \chi \end{aligned}$$

The first FOC implies that μ or/and ρ are positive, implying that w_{t+1} is either zero or determined by (B.6). As discussed in the main text, if $\mu > 0$, the board would be able to achieve truthful reporting without paying severance pay in $t+1$, in which case truthful reporting would be preferable in $t+1$ and first-best would be achievable. The second FOC implies that if $\mu \geq 0$, then $\kappa \geq 0$. Thus, it is (weakly) optimal to set $w_{s,t} = 0$, as it relaxes (B.6). Finally, the first term in the third FOC is negative. This implies that λ, μ, and/or χ are positive. If we don't have first-best ($\mu = 0$), this means that Δw_{t+1} is determined by the more stringent of conditions (B.5) and $\Delta w_{t+1} \geq 0$. We verify below that the bonus in a period without truthful reporting will be zero.[35] Similar to Case 1, the argument for the case when $\Delta w_t = 0$ is immediate.

No truthful reporting in period t**.** We continue in the steps laid out above. Suppose that we have no truthful reporting in t. The aim is to minimize the manager's payoff given by (18), subject to (9), (10), $w_t, \Delta w_t, w_{s,t} \geq 0$. The constraint that the manager invests in firm-specific human capital (9), together with the feasibility restriction $\Delta w_t \geq 0$ can be

[34] $\widetilde{E}_{\theta_N}$ is defined as E_{θ_N} but for $\widetilde{e}_{t+1}\left(\theta_N\right)$.

[35] As in case 1, the board will satisfy (B.5) only for $\theta^*_t = \theta_G$.

stated as

$$\Delta w_t \geq \max\left\{0, \frac{\frac{c}{e_t\left(\theta_{t-1}^*\right)}+\delta\left(U_{t+1}^e\left(\theta_N,\mathbf{w}\right)-U_{t+1}^e\left(\theta_G,\mathbf{w}\right)\right)}{\Delta\theta}\right\}. \tag{B.7}$$

Using that $U_{t+1}^e\left(\theta_t,\mathbf{w}\right)=\theta_t U_{t+1}\left(\theta_t,\mathbf{w}\right)+\left(1-\theta_t\right)\sum_{j=t+1}^T\delta^{j-t-1}\overline{U}_j$, the zero lower bound of Δw_t in (B.7) is binding if

$$\frac{\frac{c}{e_t\left(\theta_{t-1}^*\right)}+\delta\left(\theta_N U_{t+1}\left(\theta_N,\mathbf{w}\right)-\theta_G U_{t+1}\left(\theta_G,\mathbf{w}\right)+\Delta\theta\sum_{j=t+1}^T\delta^{j-t-1}\overline{U}_j\right)}{\Delta\theta}\leq 0. \tag{B.8}$$

Given the assumption of truthful reporting in $t-1$, the induction hypothesis is that $w_{s,t}=0$, Δw_t is given by (B.7), and $w_t=0$. Clearly, in $t=1$, the choice of w_1, Δw_1, and $w_{s,1}$ has no effect on the payoffs in neither previous (as there are none) nor following periods. Thus, the aim is to minimize Δw_1, $w_{s,1}$ and w_1, subject to (9) and (10) proving the first induction step.

Using the induction hypothesis to plug into the manager's payoff (18) in period t, we have

$$\begin{aligned}
&\left(\theta_N+\widetilde{e}_t\left(\theta_{t-1}\right)\Delta\theta\right)\max\left\{0, \frac{\frac{c}{e_t\left(\theta_{t-1}^*\right)}+\delta\left(U_{t+1}^e\left(\theta_N,\mathbf{w}\right)-U_{t+1}^e\left(\theta_G,\mathbf{w}\right)\right)}{\Delta\theta}\right\}-\widetilde{c}\\
&+\widetilde{e}_t\left(\theta_{t-1}\right)\delta U_{t+1}^e\left(\theta_G,\mathbf{w}\right)+\left(1-\widetilde{e}_t\left(\theta_{t-1}\right)\right)\delta U_{t+1}^e\left(\theta_N,\mathbf{w}\right)\\
=&\left\{\begin{array}{ll}\delta\frac{\theta_G U_{t+1}^e\left(\theta_N,\mathbf{w}\right)-\theta_N U_{t+1}^e\left(\theta_G,\mathbf{w}\right)}{\Delta\theta}+\frac{\left(\theta_N+\widetilde{e}_t\left(\theta_{t-1}\right)\Delta\theta\right)c}{e_t\left(\theta_{t-1}^*\right)}-\widetilde{c} & \text{if } \Delta w_t\geq 0\\ \widetilde{e}_t\left(\theta_{t-1}\right)\delta U_{t+1}^e\left(\theta_G,\mathbf{w}\right)+\left(1-\widetilde{e}_t\left(\theta_{t-1}\right)\right)\delta U_{t+1}^e\left(\theta_N,\mathbf{w}\right)-\widetilde{c} & \text{if } \Delta w_t=0\end{array}\right..
\end{aligned} \tag{B.9}$$

Case 3: No truthful reporting in period t ***and truthful reporting in period*** $t+1$**.** Clearly, if (B.8) is lax (and so $\Delta w_t=0$), the objective would be to minimize $U_{t+1}^e\left(\theta_t,\mathbf{w}\right)$ and, thus, all of $\{w_{s,t+1}w_{t+1},\Delta w_{t+1}\}$; with $w_{s,t+1}$ defined by (B.3), $w_{t+1}=0$, and Δw_{t+1} defined by the more stringent of (B.8) and (B.2). We now show that if the weak inequality in (B.8) were reversed, it would always be binding, implying that $\Delta w_t=0$. To minimize the first line of (B.9), we need to minimize

$$\begin{aligned}
&\theta_G U_{t+1}^e\left(\theta_N,\mathbf{w}\right)-\theta_N U_{t+1}^e\left(\theta_G,\mathbf{w}\right)\\
=\;&\theta_G\theta_N\left(U_{t+1}\left(\theta_N,\mathbf{w}\right)-U_{t+1}\left(\theta_G,\mathbf{w}\right)\right)+\Delta\theta\sum_{j=t+1}^T\delta^{j-t-1}\overline{U}_j
\end{aligned} \tag{B.10}$$

or, thus, equivalently minimize $U_{t+1}(\theta_N, \mathbf{w}) - U_{t+1}(\theta_G, \mathbf{w})$, subject to (B.2), (B.3), the reverse inequality in (B.8) and $w_{s,t+1}, w_{t+1}, \Delta w_{t+1} \geq 0$. Hence, define

$$\begin{aligned}
& L_3(\mathbf{w}, \Lambda) \\
= \; & -(\widetilde{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G))(w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2}(\theta_G, \mathbf{w})) \\
& -((1 - \widetilde{e}_{t+1}(\theta_N)) - (1 - e_{t+1}(\theta_G)))\left(w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} \overline{U}_j\right) \\
& +\lambda\left(w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2}(\theta_G, \mathbf{w}) - \frac{c}{e_{t+1}(\theta_t^*)} - w_{s,t+1} - \sum_{j=t+1}^{T} \delta^{j-t-1} \overline{U}_j\right) \\
& +\mu\left(w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} \overline{U}_j - w_{t+1} - \theta_N \Delta w_{t+1} - \delta U_{t+2}(\theta_N, \mathbf{w})\right) \\
& +\kappa w_{s,t+1} + \rho w_{t+1} + \chi \Delta w_{t+1} \\
& +\sigma\left(\theta_N U_{t+1}(\theta_N, \mathbf{w}) - \theta_G U_{t+1}(\theta_G, \mathbf{w}) + \Delta\theta \sum_{j=t+1}^{T} \delta^{j-t-1} \overline{U}_j + \frac{c}{\delta e_t(\theta_{t-1}^*)}\right),
\end{aligned}$$

where $\Lambda = \{\lambda, \mu, \kappa, \rho, \chi, \sigma\}$ is the set of weakly positive Kuhn Tucker multipliers. Taking the first-order conditions

$$\begin{aligned}
\frac{\partial L_3(\mathbf{w}, \Lambda)}{\partial w_{s,t+1}} = 0 = \quad & -((1 - \widetilde{e}_{t+1}(\theta_N)) - (1 - e_{t+1}(\theta_G))) \\
& +\sigma(\theta_N(1 - \widetilde{e}_{t+1}(\theta_N)) - \theta_G(1 - e_{t+1}(\theta_G))) - \lambda + \mu + \kappa \\
\frac{\partial L_3(\mathbf{w}, \Lambda)}{\partial \Delta w_{t+1}} = 0 = \quad & -(\widetilde{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G))\theta_G \\
& +\sigma(\theta_N \widetilde{e}_{t+1}(\theta_N) - \theta_G e_{t+1}(\theta_G))\theta_G + \theta_G \lambda - \theta_N \mu + \chi \\
\frac{\partial L_3(\mathbf{w}, \Lambda)}{\partial w_{t+1}} = 0 = \quad & -(\widetilde{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G)) \\
& +\sigma(\theta_N \widetilde{e}_{t+1}(\theta_N) - \theta_G e_{t+1}(\theta_G)) + \lambda - \mu + \rho
\end{aligned}$$

we obtain from the second and third condition that $\mu\Delta\theta = \theta_G \rho - \chi$ (observe that if $\rho = 0$, then we must have $\mu = \chi = 0$). From the first and third condition, we have $\sigma = \frac{\kappa+\rho}{\Delta\theta}$. To see that this implies $\sigma > 0$, suppose to the contrary that $\kappa = \rho = 0$. Since this would imply $\mu = \chi = 0$, we obtain that the RHS of the second first-order condition would be strictly positive, leading to a contradiction. Hence, we must have $\sigma > 0$, i.e., the continuation payoff is high enough that $\Delta w_t = 0$, as was to be shown. That is, the bonus in t is deferred, and Δw_{t+1} (and the continuation payoffs $U_{t+2}(\theta_{t+1}, \mathbf{w})$) must be chosen such that the more stringent of (B.8) and (B.2) is binding.

Case 4: No truthful reporting in periods t **and** $t+1$**.** In analogy to Case 3, the objective is to minimize (B.9), where $U_{t+1}(\theta_t, \mathbf{w})$ is given by (B.4). We show again that $\Delta w_t \geq 0$ is binding by arguing to a contradiction. To minimize $U_{t+1}(\theta_N, \mathbf{w}) - U_{t+1}(\theta_G, \mathbf{w})$, define

$$
\begin{aligned}
& L_4(\mathbf{w}, \Lambda) \\
= \; & -(\widetilde{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G)) \Delta\theta \Delta w_{t+1} - \delta \widetilde{\mathbb{E}}_{\theta_N}\left[U^e_{t+2}(\theta_{t+1}, \mathbf{w})\right] + \delta E_{\theta_G}\left[U^e_{t+2}(\theta_{t+1}, \mathbf{w})\right] \\
& + \lambda\left(\Delta w_{t+1} - \frac{c}{e_{t+1}(\theta^*_t)\Delta\theta} - \delta \frac{U^e_{t+2}(\theta_N, \mathbf{w}) - U^e_{t+2}(\theta_G, \mathbf{w})}{\Delta\theta}\right) \\
& + \mu\left(w_{t+1} + \theta_N \Delta w_{t+1} + \delta U^e_{t+2}(\theta_N, \mathbf{w}) - w_{s,t+1} - \sum_{j=t+1}^{T} \delta^{j-t-1} \overline{U}_j\right) \\
& + \kappa w_{s,t+1} + \rho w_{t+1} + \chi \Delta w_{t+1} \\
& + \sigma\left(\theta_N U_{t+1}(\theta_G, \mathbf{w}) - \theta_G U_{t+1}(\theta_N, \mathbf{w}) + \Delta\theta \sum_{j=t+1}^{T} \delta^{j-t-1} \overline{U}_j + \frac{c}{\delta e\left(\theta^*_{t-1}\right)}\right).
\end{aligned}
$$

Taking the first order condition with respect to Δw_{t+1}, we have

$$
\begin{aligned}
\frac{\partial L_4(\mathbf{w}, \Lambda)}{\partial \Delta w_{t+1}} = 0 = \quad & -(\widetilde{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G)) \Delta\theta + \lambda + \mu\theta_N + \chi \\
& + \sigma\left(\theta_N(\theta_N + e_{t+1}(\theta_G)\Delta\theta) - \theta_G(\theta_N + \widetilde{e}_{t+1}(\theta_N)\Delta\theta)\right).
\end{aligned} \tag{B.11}
$$

Since the first line in (B.11) is positive, it must be that $\sigma > 0$ as the term following σ can be rewritten as $-(\theta_N(1 - e_{t+1}(\theta_G)) + \theta_G \widetilde{e}_{t+1}(\theta_N))\Delta\theta$, which is strictly negative. Hence, both in Case 3 and 4, we have that $\sigma > 0$, implying that the continuation payoff must be high enough that $\Delta w_t = 0$. For completeness, note that $\frac{\partial L_4(\mathbf{w}, \Lambda)}{\partial \Delta w_{t+1}} = \mu + \rho - \Delta\theta\sigma$ and $\frac{\partial L_4(\mathbf{w}, \Lambda)}{\partial w_{s,t+1}} = -\mu + \kappa$. As before, if $\mu > 0$, the board would be able to achieve truthful reporting in $t+1$ without paying severance (which would be preferred). If, instead, $\mu = 0$, then $\sigma > 0$ implies that $\rho > 0$, and thus $w_{t+1} = 0$; setting $w_{s,t+1} = 0$ is then also weakly optimal.

Case: No truthful reporting in $t-1$ Suppose that the board does not seek truthful reporting in period $t-1$. If the board seeks truthful reporting in t, the condition that the manager invests in firm-specific human capital in period $t-1$ even if $\Delta w_{t-1}=0$ is

$$\begin{aligned}
0 &\geq \frac{c}{e_{t-1}\left(\theta_{t-2}^{*}\right)}+\delta\left(U_{t}^{e}\left(\theta_{N}, \mathbf{w}\right)-U_{t}^{e}\left(\theta_{G}, \mathbf{w}\right)\right) \\
&= \frac{c}{e_{t-1}\left(\theta_{t-2}^{*}\right)}+\delta\left(\begin{array}{c}
\theta_{N}\left(\begin{array}{c}
-\widetilde{c}+\left(\widetilde{e}_{t}\left(\theta_{N}\right) \theta_{G}+\left(1-\widetilde{e}_{t}\left(\theta_{N}\right)\right) \theta_{N}\right) \Delta w_{t} \\
+\delta\left(\widetilde{e}_{t}\left(\theta_{N}\right) U_{t+1}\left(\theta_{G}, \mathbf{w}\right)+\left(1-\widetilde{e}_{t}\left(\theta_{N}\right)\right) U_{t+1}\left(\theta_{N}, \mathbf{w}\right)\right)
\end{array}\right) \\
-\theta_{G}\left(\begin{array}{c}
-c+\left(e_{t}\left(\theta_{G}\right) \theta_{G}+\left(1-e_{t}\left(\theta_{G}\right)\right) \theta_{N}\right) \Delta w_{t} \\
+\delta\left(e_{t}\left(\theta_{G}\right) U_{t+1}\left(\theta_{G}, \mathbf{w}\right)+\left(1-e_{t}\left(\theta_{G}\right)\right) U_{t+1}\left(\theta_{N}, \mathbf{w}\right)\right)
\end{array}\right)
\end{array}\right) \\
& +\delta \Delta \theta \sum_{j=t}^{T} \delta^{j-t} \overline{U}_{j}
\end{aligned}$$

Defining

$$\Delta w_{t-1}^{-1}(0):=\frac{\begin{array}{c}\frac{c}{e_{t-1}\left(\theta_{t-2}^{*}\right)}+\delta\left(\theta_{G} c-\theta_{N} \widetilde{c}\right)+\delta \Delta \theta \sum_{j=t}^{T} \delta^{j-t} \overline{U} \\
+\delta^{2}\left(\begin{array}{c}\left(\theta_{N} \widetilde{e}_{t}\left(\theta_{N}\right)-\theta_{G} e_{t}\left(\theta_{G}\right)\right) U_{t+1}\left(\theta_{G}, \mathbf{w}\right) \\
-\left(\theta_{G}-\theta_{N}-\theta_{G} e_{t}\left(\theta_{G}\right)+\theta_{N} \widetilde{e}_{t}\left(\theta_{N}\right)\right) U_{t+1}\left(\theta_{N}, \mathbf{w}\right)\end{array}\right)\end{array}}{\delta \Delta \theta\left(\theta_{N}+\theta_{G} e_{t}\left(\theta_{G}\right)-\theta_{N} \widetilde{e}_{t}\left(\theta_{N}\right)\right)}, \tag{B.12}$$

we, therefore, need that $\Delta w_{t} \geq \Delta w_{t-1}^{-1}(0)$. We can now proceed in analogy to Cases 1 and 2 by augmenting the induction hypothesis with

$$\Delta w_{t}=\max \left\{\begin{array}{c}\frac{\frac{c}{e_{t}\left(\theta_{t}^{*}\right) \Delta \theta}+\delta \frac{U_{t+1}\left(\theta_{N}, \mathbf{w}\right)-U_{t+1}\left(\theta_{G}, \mathbf{w}\right)}{\Delta \theta}}{}, \\
\frac{c}{e_{t}\left(\theta_{t}^{*}\right) \theta_{G}}+\frac{\sum_{j=t}^{T} \delta^{j-t} \overline{U}-\delta U_{t+1}\left(\theta_{G}, \mathbf{w}\right)}{\theta_{G}}, \Delta w_{t-1}^{-1}(0), 0\end{array}\right\}. \tag{B.13}$$

It is now straightforward to show that, regardless of θ_{t-1}^{*} and whether the first or third term is larger, the objective to minimize U_t again boils down to minimizing (B.1). When $\Delta w_t = 0$, the result is immediate as in Case 1. Finally, when the second term is largest (i.e., all incentive constraints are satisfied for $w_{s,t}=0$), the manager extracts no rent, and the contract is again optimal.

Analogously, if the board does not seek truthful reporting in t, the condition that the manager invests in firm-specific human capital in period $t-1$ even if $\Delta w_{t-1}=0$ requires that $\Delta w_t \geq \Delta w_{t-1}^{-1}(0)$, where

$$\Delta w_{t-1}^{-1}(0):=\frac{\begin{array}{c}\frac{c}{e_{t-1}\left(\theta_{t-2}^{*}\right)}+\delta\left(\theta_{G} c-\theta_{N} \widetilde{c}\right)+\delta \Delta \theta \sum_{j=t}^{T} \delta^{j-t} \overline{U} \\
+\delta^{2}\left(\begin{array}{c}\left(\theta_{N} \widetilde{e}_{t}\left(\theta_{N}\right)-\theta_{G} e_{t}\left(\theta_{G}\right)\right) U_{t+1}^{e}\left(\theta_{G}, \mathbf{w}\right) \\
-\left(\theta_{G}-\theta_{N}-\theta_{G} e_{t}\left(\theta_{G}\right)+\theta_{N} \widetilde{e}_{t}\left(\theta_{N}\right)\right) U_{t+1}^{e}\left(\theta_{N}, \mathbf{w}\right)\end{array}\right)\end{array}}{\delta \Delta \theta\left(\theta_{N}+\theta_{G} e_{t}\left(\theta_{G}\right)-\theta_{N} \widetilde{e}_{t}\left(\theta_{N}\right)\right)},$$

We can now augment the induction hypothesis with

$$\Delta w_{t}=\max \left\{0, \frac{\frac{c}{e_{t}\left(\theta_{t-1}\right)}+\delta\left(U_{t+1}^{e}\left(\theta_{N}, \mathbf{w}\right)-U_{t+1}^{e}\left(\theta_{G}, \mathbf{w}\right)\right)}{\Delta \theta}, \Delta w_{t-1}^{-1}(0)\right\}. \tag{B.14}$$

Regardless of whether the second or third term is larger, the objective to minimize U_t boils down to minimizing (B.10). Together with the arguments above, this implies that we can follow again the steps in Cases 1-4 to minimize (B.1) and (B.10), respectively. With more than two periods in which the board does not seek truthful reporting, we proceed analogously. **Q.E.D.**

Omitted Derivations in Main Text

Lemma B.2 *A sufficient condition for (1) to hold is*

$$\begin{aligned} & \frac{e_N \theta_G}{e_N \theta_G + (1-e_N)\theta_N} e_G + \frac{(1-e_N)\theta_N}{e_N \theta_G + (1-e_N)\theta_N} e_N \\ > \quad & e_1 > \frac{e_G(1-\theta_G)}{e_G(1-\theta_G) + (1-e_G)(1-\theta_N)} e_G + \frac{(1-e_G)(1-\theta_N)}{e_G(1-\theta_G) + (1-e_G)(1-\theta_N)} e_N . \end{aligned} \tag{B.15}$$

Proof of Lemma B.2 The second inequality in (1) is most difficult to satisfy if underperformance in period t should (ex post) make it optimal to replace a manager even if her fit in $t-1$ was θ_G. The first inequality is most difficult to satisfy if it should be ex post optimal to keep the manager after realizing the high cash flow in t even if her fit in $t-1$ was θ_N. These conditions are captured by (B.15). **Q.E.D.**

Calculating the Board's Expected Payoff in Section 3.2.1 In equilibrium, the board's expected payoff is V^*. Hence, we can rewrite (4) as

$$V^* = \frac{\sum_{i=1}^{2} \delta^{i-1} \overline{U} - U_1(\mathbf{w}) + \mathbb{E}\left[\sum_{i=1}^{2} \delta^{i-1} q_i \left(x_i - c - \overline{U}\right)\right]}{1 - \mathbb{E}\left[\sum_{i=1}^{2} \delta^{i-1} \widetilde{q}_i\right] \delta}.$$

If the board seeks truthful reporting in both periods, this expression becomes

$$V^* = \frac{-U_1(\mathbf{w}) + \begin{pmatrix} x + \left(\overline{\theta} + e_1\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c \\ +\delta e_1 \left(x + \left(\overline{\theta} + e_2(\theta_G)\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c\right) \end{pmatrix} + \delta(1-e_1)\overline{U}}{1 - \delta(1-e_1) - \delta^2 e_1}.$$

If instead, the board seeks truthful reporting only in the second period, we have

$$V^* = \frac{\begin{matrix} -U_1(\mathbf{w}) + \begin{pmatrix} (x + (\theta_N + e_1(\theta_G - \theta_N))\Delta x - c) \\ +\delta e_1 \theta_G \left(x + \left(\overline{\theta} + e_2(\theta_G)\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c\right) \\ +\delta(1-e_1)\theta_N \left(x + \left(\overline{\theta} + e_2(\theta_N)\left(\theta_G - \overline{\theta}\right)\right)\Delta x - c\right) \end{pmatrix} \\ +\delta\left(e_1(1-\theta_G) + (1-e_1)(1-\theta_N)\right)\overline{U} \end{matrix}}{1 - \left(\delta\left(e_1(1-\theta_G) + (1-e_1)(1-\theta_N)\right) + \delta^2\left(e_1\theta_G + (1-e_1)\theta_N\right)\right)}.$$

By plugging in for $\{x, \Delta x\}$, $\overline{U}$, and $U_1(\mathbf{w})$, we obtain the values in Table 1.