Abstract

Firm executives are often hired with renewable fixed-term contracts. This paper asks why and what determines the length of such contracts. The model’s rationale derives from a setting in which the fit between a firm and its managers changes over time. Though compensating managers with severance pay for premature termination mitigates their incentives to conceal a deteriorating fit today, such pay increases these incentives in preceding periods. This trade-off can be managed by optimally choosing the length of renewable fixed-term contracts. Analyzing the determinants of contract length, such as 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
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.” 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).

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 replaces managers that it believes ex post to be a bad fit.2

In a world in which the managers’ fit is known to all, the board should replace managers who are a bad fit and retain those who are a good fit. This strategy improves not only current, but also future, firm performance, as a manager who is a good (bad) fit today is more likely to be a good (bad) fit tomorrow. Trying to follow this strategy when boards are less-informed than managers about their fit requires offering managers severance pay to compensate those with deteriorating fit for not concealing it and forgoing future wages. Since wages are typically above managers’ outside options to incentivize investment in firm-specific human capital, the cost of pursuing this strategy can be high. In particular, offering severance pay has two crucial disadvantages in a dynamic setting. First, severance pay reduces the benefit of deferring compensation, which is otherwise key in dynamic models.

---

2 Similar to papers with limited commitment (Hart and Tirole, 1988; Laffont and Tirole, 1990; Malcomson, 2016), the inability to commit not to fire a bad fit makes it suboptimal to always screen agents’ private information. Note that none of the following arguments requires that the manager agrees that someone else might be a better fit for the job.
Since the board may retain a manager that it believes is a good fit even after low cash flow realizations (attributing them to bad luck), the cost of deferring pay is that it makes managers more willing to conceal a deteriorating fit. Second, offering severance pay in a given period increases a 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, one way of reducing the cost of severance pay when prematurely replacing a manager is to offer shorter contracts. A key factor determining contract length is the manager’s outside option conditional on leaving the firm. If a manager’s outside employment opportunities pay little, she would be reluctant to reveal information that may lead to her dismissal unless she is offered higher severance pay. Shorter contracts help by reducing the dynamic problem that high severance pay tomorrow requires even higher severance pay today. Naturally, firms that stand to gain more from having the right manager in charge, such as firms with better growth options, would tolerate also higher severance pay. Thus, such firms will offer longer contracts.

The intuition about the factors affecting contract length extends to the general case in which the board may optimally decide not to offer severance pay in all periods. Figure 1 illustrates the board’s optimal choice of a contract that gives it the right to dismiss the manager in some periods without paying severance (solid line). 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.\(^3\) The length of the fixed-term contracts corresponds then to the time between renewal dates at which the board has the right to dismiss the manager without severance pay.

The advantage of having the right to fire the manager without severance pay 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 (take actions to) conceal her fit. This forces the board to rely more heavily on the firm’s performance as a noisy indicator.

\(^3\)Instead, above, a shorter contract meant hiring a new manager after the incumbent’s contract is over.
of the manager’s fit and increases the risk of incorrect replacement and retention decisions.\(^4\) 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 without severance pay following underperformance. The factors are the same as discussed above. Specifically, since offering severance pay is more expensive when managers’ outside options are low, boards would rely more heavily on firm performance to evaluate managers. This would go 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. Since age can serve as a natural commitment device to shorter contracts, the paper discusses several implications for when firms will choose older managers.

The paper’s main contribution is that it rationalizes the use of renewable fixed-term contracts and analyzes key determinants of contract length. Building on prior static models (Levitt and Snyder, 1997; Inderst and Mueller, 2010; Almazan and Suarez, 2003; Van Wesep, 2010), its results add to our understanding of the dynamic trade-offs and implications of offering severance pay. 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. Further developing this insight is important, as it helps derive new implications for contract length and helps explain seemingly puzzling hiring practices. Specifically, the paper shows 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.\(^5\) This would imply that the board may appear to choose from a “select club” of managers with better outside options (even if they are not better), as they would be subsequently less desperate to conceal a deteriorating fit. For the same reason, the board may tolerate investment in general human

\(^4\)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, 2018).

\(^5\)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.
capital, even if it comes 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 contract on the manager’s fit. The resulting suboptimality of full revelation has been analyzed also in the literature on relational contracts (Halac, 2012; Malcomson, 2016). However, by assuming an extreme inability to write explicit contracts, this literature better describes “at will” employment, while in the current paper contract horizon is the main focus.6

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.7

---

6A 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.

7The paper also contributes to prior work on human capital investments (Jovanovic, 1979 a,b; and Felli and 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).
2 Model

Consider an infinitely-lived firm in which a board is in charge of hiring and replacing the firm’s managers and designing the employment contracts governing the relationship with the firm. 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 (“she”) 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 $\theta_G$ with probability $e_t(\theta_{t-1})$. With probability $1 - e_t(\theta_{t-1})$ or, respectively, if she does not invest in firm-specific human capital, her fit is $\theta_N < \theta_G$. A manager investing in firm-specific human capital is more likely to be a good fit in $t$ if she was a good fit in $t - 1$ and vice versa. That is, there is a positive correlation with $e_t(\theta_G) > e_1 > e_t(\theta_N)$, where $e_1$ is the likelihood of $\theta_G$ in the manager’s first (complete) period after being hired. 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 $\bar{U}$ per period, are fixed, but Section 3.3 relaxes these assumptions.

At the interim date $\tau_t = 1$ of a period, the manager privately learns her fit and can report it, and the board can decide whether to replace her with a new manager. While the paper uses throughout the term “reporting a bad fit,” in practice, this would correspond to not concealing information or not taking actions that would prevent outsiders from realizing that the manager’s fit has deteriorated. 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 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 finitely-lived and leave the labor market once they reach their retirement age. The key information frictions are that a manager’s investments in firm-specific human capital, as well as the realizations of $\theta_t$ at the interim date $\tau_t = 1$ of every period, are known only to the manager.
Contracting If the board replaces the incumbent at the interim date of a period, then at a cost of $\mathcal{U}$ (e.g., for arranging an interim manager), the board completes the period and has a probability $\mathcal{B}$ of achieving the high cash flow, where $\theta_N < \mathcal{B} < \theta_C$. Then, at the beginning of the following period, the board makes a new manager an offer covering the whole potential relationship. If a manager is replaced, she is not rehired.

**Assumption 1:** The board maximizes ex ante shareholder wealth subject to the restriction that it must replace a manager whenever its posterior beliefs indicate that a new manager is more likely to generate high cash flows or be a good fit next period.

Assumption 1 can be motivated with career-concerned boards that want to avoid the risk of offering contracts that force them to make replacement or retention decisions that may appear wrong to outsiders, given the information revealed until this point.\(^8\) Note that Assumption 1 stipulates only how the board reacts to information about the manager’s fit, but says nothing about whether the board should try to elicit the manager’s private information or rely on the firm’s cash flows as noisy signals of the manager’s fit. Indeed, as will become clear already in Section 3.1, the main novel economic forces are not an artefact of Assumption 1, but the assumption helps to make the analysis more tractable. Section 3.3 dispenses with Assumption 1.

Let $w = (w_t, \Delta w_t, w_{s,t}, \psi^1_t, \psi^2_t)_{t=1}^T$ be the contract that the board offers the manager at the beginning of the manager’s first period. Assumption 1 simplifies the contracting space by making conditioning on the history of reports about the manager’s fit until $t-1$ immaterial since managers that report being a bad fit are dismissed. Note that the prospect of such dismissal means that a manager may not have incentives to report truthfully a deteriorating fit (unless she is offered severance pay for premature termination). If this is the case, the manager’s report is not informative and can be ignored. Thus, the relevant history on which the contract can be conditioned comprises of the sequence of periods in which the contract has offered the manager adequate incentives to report her fit truthfully and the history of cash flow realizations $(x_t)_{t=1}^{t-1}$. In any given period $t$, $w_t$ stands for the manager’s wage in the low cash flow state, and $\Delta w_t$ stands for how much she receives in addition (i.e., her “bonus”) in 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; $\psi^1_t = \{0, 1\}$ is an indicator function describing whether the board terminates the manager’s contract at the interim date depending on the manager’s report about her fit in $t$; $\psi^2_t = \{0, 1\}$ stands for whether the board terminates the manager’s contract at the end of the period, i.e., in $\tau_t = 2$, depending in addition on the cash flow realization $x_t$.\(^9\)

---

8In a setting in which incompetent boards make random replacement decisions, a competent board may want to avoid the risk of appearing incompetent.

9The payment to the manager at the end of the period could also be reinterpreted as the manager’s
If new manager: board offers contract & manager accepts/rejects. Manager privately learns \( \theta_t \in \{ \theta_G, \theta_N \} \), can report \( \hat{\theta} \). \( x_t \in \{ x, x + \Delta x \} \) realized. Probability of \( x + \Delta x \) is \( \theta_t \). Contract terminated or continued. Manager paid \( w_t \) or \( w_t + \Delta w_t \). If new manager hired at \( \tau_t = 1 \), the prob. of \( x + \Delta x \) is \( \overline{\theta} \) and new manager paid \( \overline{U} \) in \( t \). If board replaces manager, manager is paid \( w_{s,t} \). If no investment: fit is \( \theta_N \). 

**Figure 2:** Timing of events in period \( t \).
realization is enough to make the board believe that the manager’s fit has deteriorated even if her fit in the last period was still good; the opposite is assumed (given incentives for investment in firm-specific human capital) if the current cash flow is high:

\[
\sum_{\theta_t \in \{\theta_N, \theta_G\}} \Pr (\theta_t| x, \theta_{t-1} = \theta_G) e_{t+1}(\theta_t) < e_1 < \sum_{\theta_t \in \{\theta_N, \theta_G\}} \Pr (\theta_t| x + \Delta x, \theta_{t-1} = \theta_N) e_{t+1}(\theta_t). \tag{1}
\]

Condition (1) is a sufficient condition independent of the history of contracting; it is given in terms of primitives in Appendix B. Together with Assumption 1, condition (1) implies that in periods without incentives for truthful reporting, the board will fire managers if and only if the firm’s cash flows are low. Just as Assumption 1, condition (1) helps to focus on the main novel effects, but can be relaxed. Section 3.3.4 discusses several such extensions, including one in which the manager’s fit is distributed continuously.\(^{12}\)

3 A Multi-Period Employment Relationship

Given a contract offer \(w = (w_t, \Delta w_t, w_{s,t}, \psi_1^{-1}, \psi_2^{-1})^T_{t=1}\), let \(\omega_t \in \{w_t, w_t + \Delta w_t, w_{s,t} + U\}\) denote the firm’s wage bill in period \(t\).\(^{13}\) The board’s expected payoff in the first period is

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

where \(E\) is the expectation over the future \(\theta_t\) and \(x_t\) realizations; and \(q_i\) and \(\bar{q}_i\) are the endogenous probabilities (defined in the Appendix) 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.

\(^{12}\)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. For example, one could assume that the manager believes that a replacement’s success likelihood would be only \(\bar{\theta} < \theta_N\). 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).

\(^{13}\)Recall that if the board replaces a manager at \(\tau_t = 1\), it pays \(w_{s,t}\) to the departing manager and \(U\) to the replacement manager to complete the period.
The board’s promise-keeping constraint implies that the manager’s expected payoff in any given period \( t \) during her tenure is

\[
U_t (\theta_{t-1}, w) = E \left[ \sum_{i=t}^{T} \delta^{i-t} (\bar{U} + q_i (\omega_i - \tilde{c}_i - \bar{U})) \right] | \theta_{t-1},
\]

(3)

where \( \tilde{c}_i = c \) if the manager invests in firm-specific human capital and \( \tilde{c}_i = 0 \) otherwise. Expression (3) states that the manager can obtain \( \bar{U} \) in every period until she leaves the labor market in \( T \), but she might receive something different from \( \bar{U} \) 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 \( t \), \( U_t (\theta_{t-1}, w) \), depends on her fit realization in \( t - 1 \) (because \( e_t \) depends on \( \theta_{t-1} \)). There is no such prior realization when she is hired, so for period one, we write \( U_1 (w) \).

Using (3) to plug into (2), the board’s objective when hiring a manager is to choose \( w \) to maximize

\[
\max_w E \left[ \sum_{i=1}^{T} \delta^{i-1} (q_i (x_i - \tilde{c}_i - \bar{U}) + \tilde{q}_i \delta V^*) \right] - U_1 (w) + \sum_{i=1}^{T} \delta^{i-1} \bar{U},
\]

(4)

subject to the constraints that the contract \( w \) is feasible, incentive-compatible, and individually 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

\[
\nu_t (\theta_{t-1}, w) := U_t (\theta_{t-1}, w) - \sum_{j=t}^{T} \delta^{j-t} \bar{U}.
\]

(5)

in the first period. Note that the offered wage contract implicitly reflects also the board’s decision whether to stimulate the manager to invest in firm-specific human capital. We now state the relevant constraints.

### 3.1 The Cost-Benefit Mismatch of Offering Severance Pay

The incentive constraints that the manager truthfully reports her fit at date \( \tau_t = 1 \) of period \( t \) in which case she stays if her fit is \( \theta_G \) or is dismissed with a severance package if her fit is
\[ w_t + \theta_G \Delta w_t + \delta U_{t+1}(\theta_G, w) \geq w_{s,t} + \sum_{j=t}^{T} \delta^{j-t}U \]  

\[ w_{s,t} + \sum_{j=t}^{T} \delta^{j-t}U \geq w_t + \theta_N \Delta w_t + \delta U_{t+1}(\theta_N, w). \]

Note that the manager’s continuation payoff \( U_{t+1}(\theta_t, w) \) can take on two values depending on her fit realization (\( \theta_G \) or \( \theta_N \)) in \( t \), which means that both payoffs will play the role of state variables for characterizing the manager’s contract.\(^{14}\)

To induce a manager whose fit in \( t - 1 \) is \( \theta_{t-1}^* \) to invest in firm-specific human capital in period \( t \), the contract must further satisfy

\[
\left( e_t(\theta_{t-1}^*) (w_t + \theta_G \Delta w_t + \delta U_{t+1}(\theta_G, w)) + (1 - e_t(\theta_{t-1}^*)) (w_{s,t} + \sum_{j=t}^{T} \delta^{j-t}U) - c \right) \geq w_{s,t} + \sum_{j=t}^{T} \delta^{j-t}U,
\]

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 \( \tau_t = 1 \) of the period. If (8) is satisfied, the first incentive constraint (6) is lax. Based on (8), it is convenient to express the manager’s expected payoff in \( t \) as

\[ U_t(\theta_{t-1}, w) = w_{s,t} + \sum_{j=t}^{T} \delta^{j-t}U + \max \left\{ e_t(\theta_{t-1}) \left( w_t + \theta_G \Delta w_t + \delta U_{t+1}(\theta_G, w) \right) - w_{s,t} - \sum_{j=t}^{T} \delta^{j-t}U - c, 0 \right\}. \]

The max-operators in (9) takes into account that the manager invests in firm-specific human capital only if it is optimal to do so. Whether this is the case may depend on her type.

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 \( w_t = U \), \( w_{s,t} = \Delta w_t = 0 \). However, the incumbent would then add no value to the firm, as \( \theta_N < \bar{\theta} \) 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).

\(^{14}\)Eliciting the manager’s fit at the end of the period is suboptimal. It requires offering the manager the same information rent, 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 \( \theta_{t+1} \) on the equilibrium path in \( t + 1 \), she does the same off the equilibrium path (i.e., after misreporting in \( t \)), and restricting attention to one-shot deviations is sufficient (Golosov et al., 2016). Finally, note that, since the cash flow realization \( x_t \) carries no additional information about \( \theta_t \), it is without loss of generality not to condition the continuation payoff \( U_{t+1} \) on \( x_t \). Such history dependence becomes relevant in some of the extensions presented in Section 3.3.
Offering Severance Pay in All Periods  In a first-best world, the positive correlation of fit between periods implies that a manager who stays a good fit should be kept as long as possible. However, in this setting, the board does not know whether the manager is a good fit. This section presents the case in which the board tries to extract this information by offering severance pay in all periods. Though this policy is not optimal in general, this case will illustrate in a simple way the main novel insights regarding contract horizon and the factors affecting this horizon.

Incentivizing the manager to keep investing in firm-specific human capital and reporting truthfully her fit involves offering a contract that satisfies constraints (7) and (8) with equality in every period. Intuitively, if (7) were lax, the board would increase its payoff by reducing the manager’s severance pay until (7) binds. This would also make it easier to satisfy (8). The latter constraint would also be binding, since otherwise, the board would benefit from decreasing $w_t$ or the manager’s continuation payoff (all technical details are in the Appendix). Note that, since a manager whose fit is $\theta_N$ will be replaced in $t - 1$, condition (8) needs to be satisfied only for managers whose fit was $\theta_G$ in $t - 1$. Thus, the first term in the max-operator of (9) is zero, and that expression reduces then simply to

$$U_t (\theta_{t-1}, w) = w_{s,t} + \sum_{j=t}^{T} \delta^{j-t} U.$$

Hence, offering the manager severance pay not to withhold information revealing her as a bad fit leads to information rent of proportionate size.\(^{15}\) Express now $w_{s,t} = \theta_N \Delta w_t + \delta U_{t+1} (\theta_N, w) - \sum_{j=t}^{T} \delta^{j-t} U$ from (7) and assume for the moment that this expression is positive. Plugging $w_{s,t}$ into (8), we obtain that $\Delta w_t = \frac{c_{\theta_G}}{s_t(\theta_G) \Delta \theta}$. Thus, contrary to models in which seeking truthful revelation is not an objective, the manager’s compensation cannot be fully deferred. The reason is that rewarding the manager with severance when she is dismissed for being a bad fit makes it hard to stimulate investment in firm-specific human capital. To generate such incentives, there must be a sufficiently large difference between the manager’s payoff when she is a good fit, $w_t + \theta_G \Delta w_t + \delta U_{t+1} (\theta_G, w)$, and when she is a bad fit, $w_{s,t} + \sum_{j=t}^{T} \delta^{j-t} U$. However, with severance pay making the payoff for being a bad fit equal to $\theta_N \Delta w_t + \delta U_{t+1} (\theta_N, w)$ and with continuation payoffs in the present special case being the same, $U_{t+1} (\theta_N, w) = U_{t+1} (\theta_G, w) = w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t} U$ (see (10)), incentives for investment in firm-specific human capital can only be achieved by offering a positive

\(^{15}\) Note that the manager’s rent in $t$ does not depend on her fit 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 $t$. 

13
bonus in the current period.

Plugging in for the severance pay \( w_{s,t} \), the manager’s expected rent (5) when she is hired in period one is

\[
\nu_1(w) = w_{s,1} = \frac{\theta_N}{e_1 \Delta \theta} c + \delta \left( w_{s,2} + \sum_{j=2}^{T} \delta^{j-2} \bar{U} \right) - \sum_{j=1}^{T} \delta^{j-1} \bar{U}
\]

\[
= \sum_{j=1}^{T} \delta^{j-1} \left( \frac{\theta_{NC}}{e_j (\theta_G) \Delta \theta} - \bar{U} \right), \tag{11}
\]

**Proposition 1** (i) Following the first-best policy of hiring the manager with the longest time until retirement and incentivizing truthful reporting in all periods is optimal if

\[
\frac{\theta_{NC}}{e_1 (\theta_G) \Delta \theta} \leq \bar{U}, \tag{12}
\]

in which case the manager is truthful also without severance pay \( w_{s,t} = 0 \). (ii) If (12) is not satisfied and the board seeks truthful reporting in all periods, the severance pay increases in the number of remaining periods \((T - t)\), \( w_{s,t} = \sum_{j=t}^{T} \delta^{j-t} \left( \frac{\theta_{NC}}{e_j (\theta_G) \Delta \theta} - \bar{U} \right) \). Furthermore, the board offers a positive bonus \( \Delta w_t = \frac{c}{e_i (\theta_G) \Delta \theta} \) in every period. There is a threshold \( \hat{\delta} \), such that for \( \delta > \hat{\delta} \) the board will set a finite contract length (hire a manager with a shorter time to retirement) if seeking truthful reporting in all periods.

If condition (12) 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 the manager’s incentive to conceal a deteriorating fit. First, because the board retains the manager even following low cash flow realizations when she reports \( \theta_G \), the manager can enjoy a positive continuation payoff regardless of the cash flows in the present period. As noted above, this dramatically reduces the board’s ability to defer compensation.\(^{16}\) 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 (dashed line in Figure 1).

\(^{16}\)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.
Observe that 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, part (ii) of Proposition 1 concludes that 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 offer shorter contracts (by hiring older managers).

In all that follows, suppose that the first-best condition (12) is not satisfied. As noted, to gain some intuition about the factors affecting contract length, it is useful to analyze the case in which the board still seeks truthful reporting in all periods. Condition (12) suggests that a higher per-period outside option $U$ 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 report her fit truthfully. 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. Furthermore, longer contracts are optimal if $\Delta x$ is higher, as there is more to gain from holding on longer to a manager who is a good fit.

Proposition 2 Suppose that the first-best condition (12) is not satisfied and that the board chooses to stimulate truthful reporting by offering severance pay in all periods. Then, the board offers a longer contract if the manager’s per-period outside option $U$ and the cash flow upside $\Delta x$ are higher.

If the contract’s length $T$ is chosen to coincide with the manager’s time to retirement, Proposition 2 implies that younger managers would be offered higher severance packages, which is in line with Rau and Xu (2013). Furthermore, to the extent that outside options are lower in industry downturns, which calls for shorter $T$, older managers would be preferred in such downturns.

Implication 1 If the board offers severance pay in all periods, ex ante severance agreements must be higher for younger managers. Furthermore, the attractiveness of hiring older manager increases (i.e., a lower $T$ is preferable) in industry downturns and when firms have low growth potential.

The remainder of the paper shows that the qualitative insights of Proposition 2 regarding contract length and the factors affecting this length extend to the case in which the board optimally chooses whether to offer severance pay. Readers mainly interested in the implications of the model can skip to Implications 2–6.
3.2 The Board’s Choice of Severance Pay, Contract Horizon, and Turnover

The preceding analysis shows that whenever the first-best condition (12) is not satisfied, the board will be forced to introduce some inefficiency in its replacement policy to reduce the manager’s rent. In the present setting, one such inefficiency amounts to abstaining from offering incentives for truthful reporting and relying on the firm’s cash flow performance to inform replacement and retention decisions.

Evaluating the Manager Based on Noisy Firm Performance  In a period in which the contract does not offer incentives for truthful reporting, the constraint that the manager prefers investing in firm-specific human capital when her fit in $t - 1$ is $\theta_{t-1}^*$ to not investing and forgoing the chance of being a good fit is

$$w_t + (\theta_N + e_t(\theta_{t-1}^*) \Delta \theta) \Delta w_t + \delta E_{\theta_{t-1}}^e [U_{t+1}^e (\theta_t, w)] - c \geq w_t + \theta_N \Delta w_t + \delta U_{t+1}^e (\theta_N, w),$$

where $\Delta \theta \equiv \theta_G - \theta_N$ and where the expected continuation payoffs are defined as

$$E_{\theta_{t-1}} [U_{t+1}^e (\theta_t, w)] \equiv e_t(\theta_{t-1}) U_{t+1}^e (\theta_G, w) + (1 - e_t(\theta_{t-1})) U_{t+1}^e (\theta_N, w)$$

$$U_{t+1}^e (\theta_t, w) \equiv \theta_t U_{t+1} (\theta_t, w) + (1 - \theta_t) \sum_{j=t+1}^T \delta^{j-t-1} U.$$

Based on (13), it is convenient to express the manager’s expected payoff in $t$ as

$$U_t (\theta_{t-1}, w) = \max \left\{ w_t + (\theta_N + e_t(\theta_{t-1}) \Delta \theta) \Delta w_t - c + \delta E_{\theta_{t-1}} [U_{t+1}^e (\theta_t, w)], \right\}.$$

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

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

(15)

Condition (15) 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 $\theta_N$. 

16
3.2.1 Dynamics of the Manager’s Contract

We can use now conditions (6)–(8), (13), and (15) to derive the dynamics of the manager’s contract for any given reporting policy that the board may want to implement. 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 3** If the first-best condition (12) is not satisfied, the board pursues one of two strategies in any given period \( t \). (i) The first abandons from offering severance pay at \( t = 1 \) in which case the manager is replaced in this period if and only if the realized cash flow is low (i.e., \( \psi^1_t = 0; \psi^2_t = 1 \) if and only if \( x_t = x \)). Implementing this strategy in period \( t < T \) goes hand in hand with deferring compensation. We have

\[
w_t = \Delta w_t = w_{s,t} = 0. \tag{16}
\]

(ii) The second strategy involves incentivizing truthful reporting (tr.rep.) by offering severance pay in which case the manager is replaced in this period if and only if she reports a bad fit (i.e., \( \psi^1_t = 1 \) if and only if \( \theta_t = \theta_N; \psi^2_t = 0 \) regardless of \( x_t \)). Implementing this alternative requires that

\[
\Delta w_t = \begin{cases} 
\max \left\{ \frac{c}{\varepsilon_t(\theta_G)} + \frac{\delta U_{t+1}(\theta_N,w) - U_{t+1}(\theta_G,w)}{\delta \theta}, 0 \right\} & \text{if } \text{tr.rep. in } t - 1 \\
\max \left\{ \frac{c}{\varepsilon_t(\theta_G)} + \frac{\delta U_{t+1}(\theta_N,w) - U_{t+1}(\theta_G,w)}{\delta \theta}, \Delta w_{t-n}^{-1}(0), 0 \right\} & \text{otherwise}
\end{cases} \tag{17}
\]

\[
w_{s,t} = \theta_N \Delta w_t + \delta U_{t+1}(\theta_N,w) - \sum_{j=t}^{T} \delta^{j-t} U, \tag{18}
\]

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 (13) in these periods. The board always pursues truthful reporting in the manager’s retirement period \( T \). In the first period, \( \varepsilon_t(\theta_G) \) must be replaced by \( \varepsilon_1 \) in (17).

Part (i) of Proposition 3 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., \( w_t = 0 \)), 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, $\theta^*_t$ in expression (17) should be set to $\theta_N$.\footnote{The second expression in (17) 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.}

In analogy to Section 3.1, there are two major changes in the contract if the board tries to elicit truthful reporting from the manager (part (ii) of Proposition 3). The first one is that the manager needs to be offered adequate severance pay. The second one is that deferring bonus payments may become infeasible. A virtually identical intuition to that in Section 3.1 explains the conceptual differences to part (i) of Proposition 3. Specifically, when offering severance pay to stimulate truthful reporting, 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 $\theta_G$. The problem is that this is made difficult by the fact that the manager’s severance $w_{s,t}$ compensates her for the expected wage she would forgo upon dismissal when not concealing that her fit is no longer $\theta_G$. This makes it necessary to front load compensation, since that makes the manager’s payoff more sensitive to her true current fit in $t$. The necessary bonus might be higher if the payoff in $t$ needs to create incentives for investing in firm-specific human capital in preceding periods without severance pay. 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.

### 3.2.2 The Optimal Contract With $T = 2$

The objective of maximizing the board’s payoff (4) can be stated now as choosing the optimal reporting policy for every period, subject to (16)–(19). In this problem, the two continuation payoffs $U_{t+1} (\theta_G, w)$ and $U_{t+1} (\theta_N, w)$, 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). Given the finite dimensionality of the problem, a maximum is always achieved.

For an illustration, consider the case in which managers retire after two periods, i.e., \(T = 2\), and assume that the primitives take the values from the description of Table 1. Suppose, first, that the board offers severance pay to incentivize truthful reporting in both periods. By Proposition 3, if the manager’s outside option is \(U = 1\), the board needs to offer \(\{w_1, w_2\} = \{0, 0\}\), \(\{\Delta w_1, \Delta w_2\} = \{6.1, 5.1\}\), and \(\{w_{s,1}, w_{s,2}\}\). From (9), the manager’s payoffs in \(t = 2\) and \(t = 1\) are \(U_2(\theta_G, w) = 2.1\), \(U_2(\theta_N, w) = 2.1\), and \(U_1(w) = 4.4\) (note that \(e_1(\theta_{t-1}) = e_1\) in \(t = 1\) and that \(U_{T+1} = 0\)). We can now verify that conditions (7) and (8) are satisfied with equality, while (6) is lax. Note that, if the manager’s outside option would be lower, \(U = 0\), the manager’s severance pay would need to be increased to \(\{w_{s,1}, w_{s,2}\} = \{4.4, 2.1\}\) to satisfy the constraint that she truthfully reports \(\theta_N\). Two insights follow immediately in line with Propositions 1 and 2. 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 for forgoing 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 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)–(8) and (13) are satisfied. Furthermore, the manager’s payoff in \(t = 2\) can be calculated again from (9) as \(U_2(\theta_G, w) = 6.4\) and \(U_2(\theta_N, w) = 5.6\). These second-period payoffs are higher than with truthful reporting in both periods, as the manager’s period-one bonus is deferred until \(t = 2\). However, from (14), the manager’s payoff in \(t = 1\) (for an outside option of \(U = 1\)) is \(U_1(w) = 2.7\), 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 more reliant on the firm’s
Table 1: Board’s expected payoff. The table compares the board’s expected payoff \( V^*_{n,r} \) from offering incentives for truthful reporting in both periods and from abstaining from offering such incentives in the first period, \( V^*_{r,r} \), for different values of \( \Delta x \) and \( \overline{U} \). The examples are calculated with \( e_1 = 0.55, e(\theta_G) = 0.65, e(\theta_N) = 0.45, c = 1, \theta_G = 0.7, \theta_N = 0.4, \overline{\theta} = 0.48, \) and \( \delta = 0.95 \).

<table>
<thead>
<tr>
<th>( {x, \Delta x} )</th>
<th>( \overline{U} )</th>
<th>( V^*_{n,r} )</th>
<th>( V^*_{r,r} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( {500, 25} )</td>
<td>0</td>
<td>10,246</td>
<td>10,227</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>10,240</td>
<td>10,232</td>
</tr>
<tr>
<td>( {500, 150} )</td>
<td>0</td>
<td>11,697</td>
<td>11,748</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>11,691</td>
<td>11,754</td>
</tr>
</tbody>
</table>

*noisy performance*), especially if the cash flow upside is small (\( V^*_{n,r} > V^*_{r,r} \) if \( \Delta x = 25 \)).\(^{18}\)

These insights hold generally:

**Proposition 4** Suppose that managers retire after two periods (\( T = 2 \)). Offering severance pay to incentivize truthful reporting becomes less attractive for the board as \( \overline{U} \) and \( \Delta x \) decrease.

The determinants of whether to offer severance pay in both periods (Proposition 4) are the same as those determining the manager’s contract length when the board offers adequate severance pay in all periods (Proposition 2). This should not be surprising, as the more periods with severance pay the manager can look forward to, the higher the severance pay the board needs to offer today to achieve truthful reporting. In light of this, the following implementation of the optimal contract should be very natural.

### 3.2.3 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) 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. In the above simple example with \( T = 2 \) in which the board does not seek truthful reporting in \( t = 1 \), one can interpret the optimal contract as a renewable contract that goes for one year and can be renewed at the end of that year for one more year. Note that, in practice, an executive cannot claim severance pay without a “good

\(^{18}\)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 \( U_1 (w) = 1.8 \).
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 terminate the manager’s contract, triggering such payment. As will be discussed in Section 3.3, this implementation extends to more general formulations of the model.

In the U.S., 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 3, 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, 2018). Thus, the paper provides a simple intuition for the widespread use of such renewable fixed-term contracts.

### 3.2.4 More Than Two Periods

Solving for the optimal truthful reporting policy for \( T > 2 \) is less tractable. Numerically, this can easily be done by recursively deriving the manager’s payoff from (9)–(14) in every period for any truthful reporting policies that the board can choose from, and then selecting the policy that maximizes (4). When \( T > 2 \), introducing a period in which the board does not offer severance pay helps reduce the manager’s rent not only in that period, but also in all preceding periods. Intuitively, the prospect of lower future rent implies that the manager is truthful about her fit even when offered lower severance pay in all preceding periods. Thus, by choosing whether to seek truthful reporting in future periods, the board faces a trade-off between making more efficient replacement decisions in those periods and minimizing the manager’s rent today. 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 (solid line) and compares it to a contract that always offers incentives for truthful reporting (dashed line).

Figure 3 illustrates that the manager’s severance pay increases in the remaining periods stipulating severance pay for premature dismissal (Proposition 1). 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. In the example of Figure 3, it is optimal for the board to abstain.

---

\(^{19}\)Evergreen contracts that automatically renew every day and are, thus, equivalent to at-will contracts are signed only 2% of the time in the the U.S. (Cziraki and Groen-Xu, 2018). However, in other countries, such as the UK, they are more frequent.
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 $T = 10$, $e_1 = 0.55$, $e(\theta_G) = 0.65$, $e(\theta_N) = 0.45$, $c = 1$, $\theta_G = 0.7$, $\theta_N = 0.4$, $\bar{\theta} = 0.48$, $\delta = 0.95$, $x = 500$ and $\Delta x = 150$, $\bar{U} = 0$. 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.

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, but including periods without truthful reporting helps to reduce these bonuses (Proposition 3).

As can be expected from Proposition 4, Figure 4 illustrates that the board offers more often severance pay to stimulate truthful reporting when the manager’s outside option $\bar{U}$ is higher. The intuition is that a higher outside option makes the manager less reluctant to report that her fit is $\theta_N$ 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 $\Delta x$ increases. Intuitively, firms with higher growth prospects have more to lose from not having the right manager in charge.

The following implications follows straightforwardly from the preceding analysis:

---

20 In the example of Figure 3, which is plotted for $\bar{U} = 0$, the bonus needed to stimulate investment in firm-specific human capital in periods four and seven (which follow periods without truthful reporting) is high enough to stimulate such investments also in periods three and six. Higher outside options, however, make it more difficult to satisfy (13) and, thus, would call for a higher bonuses following periods without truthful reporting.
Figure 4: Comparative Statics in $\bar{U}$ and $\Delta x$. The figure plots how the number of periods in which the board incentivizes truthful reporting through severance pay changes in $\bar{U}$ and $\Delta x$. The primitives are the same as in Figure 3, with $T = 10$ and $\bar{U}$ taking values from zero to two and \{x, $\Delta x$\} taking values \{500, 25\} and \{500, 150\}, respectively. The figure illustrates that the board seeks more truthful reporting when the manager’s outside option $\bar{U}$ and the cash flow upside $\Delta x$ are higher.

Implication 2 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 3 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 empirically-documented 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.
3.3 Extensions and Discussion

The following section discusses several extensions of the model, including a generalization dispensing with Assumption 1.

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 $\bar{U}$. Assume further that condition (12) is not satisfied for any $\bar{U}$ 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 $\bar{U}$. However, this prediction may revert if managers are privately informed about their fit and their investments in firm-specific human capital. On the one hand, a lower outside option makes it easier to stimulate incentives for investment in firm-specific human capital in periods in which the manager’s contract offers no incentives for truthful reporting. On the other hand, in periods in which the board seeks truthful reporting, hiring a manager with a higher outside option lowers the board’s expected wage bill since it reduces the need for generous severance pay $w_{s,t}$. The second effect dominates for policies relying more on severance pay in which case a manager with a higher outside option becomes preferable.

**Proposition 5** Take any given policy of stimulating truthful reporting through severance pay. If $\left( p(w) - \delta^T \right) > (1 - \delta) \frac{\partial}{\partial U} U_1(w)$, the board’s payoff increases in the manager’s outside option $\bar{U}$ for that policy ($p(w)$ is defined in the Appendix).

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 obtain 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, given that managers often seem to be paid above their outside options, 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 concealing a deteriorating fit. 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, the model 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 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 of 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
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.\(^{21}\)

A commitment to avoid renegotiations can be achieved in the present context if \( \delta \) is sufficiently high. 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 then prevent the board from deviating from its commitment to avoid renegotiations.

### 3.3.4 Replacement Strategies When Boards are Not Career Concerned

As Section 3.1 has made clear, the main novel economic forces and implications described in the paper have little to do with Assumption 1. To summarize, a first-best replacement policy would require retaining managers who are a good fit and replacing managers who are a bad fit. However, following this policy requires eliciting the manager’s fit in every period, which may allow managers to extract too much information rent. To reduce this rent, the board may introduce some inefficiency in the replacement policy. In the baseline case, analyzed under Assumption 1, this meant introducing periods in which the manager’s contract does not offer incentives for reporting a bad fit, forcing the board to rely on noisy cash flow performance to make replacement and retention decisions.

Dispensing with Assumption 1 naturally increases the contractual possibilities to reduce the manager’s rent. First, the contract may now stipulate that also managers who are a good fit are replaced without severance pay following the cash flow realizations at the end of some periods. The incentive constraint that the manager truthfully reports a bad fit becomes then

\[
w_{s,t} + \sum_{j=t}^{T} \delta^{j-t} U \geq w_t + \theta_N \Delta w_t + \delta U_{t+1}^e (\theta_N, w) .
\]  

(20)

where \( U_{t+1}^e (\theta_N, w) = \theta_N U_{t+1} (\theta_N, w) + (1 - \theta_N) \sum_{j=t+1}^{T} \delta^{j-t-1} U \) if the manager would be fired following a low cash flows realization. Indeed, under some conditions, replacing a manager who is a good fit following low cash flows (but not following high cash flows) could help relax both (20) and the constraint that the manager invests in firm-specific human capital.\(^{22}\) Thus,

---

\(^{21}\)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.

\(^{22}\)Note that replacing a manager who is a good fit was implicitly allowed also in the baseline case when the contract’s expiration in \( T \) is chosen to coincide with the manager’s retirement age.
similar to a period without incentives for truthful reporting, the board reduces the manager’s rent by making it impossible for the manager to stay at the firm regardless of its cash flows by just lying about her fit. As a result, just as in the baseline model, turnover becomes more (under-)performance-sensitive. Also here, the cost is a potentially inefficient replacement. Contract length would again correspond to the distance between periods in which the board can replace the manager without paying severance at the period’s end. However, the main economic forces are the same. In particular, high outside options reduce the board’s cost of offering severance pay and, thus, reduce the need to rely on alternative strategies involving potentially inefficient replacement. This speaks in favor of longer contracts.

A second possibility to reduce the manager’s rent (again introducing inefficiency in replacement) is to retain managers that report a bad fit, but offer them different continuation contracts. In this case, a contract’s length could become conditional on the history of reports. However, there is no reason to believe that the insight that a higher $U$ and $\Delta x$ reduce the need to rely on such inefficient strategies would be affected. As a practical matter, however, it should be noted that contracts featuring such a level of completeness are typically not observed in practice. The most likely reason is that it is hard to define what represents a good fit. Thus, the simplification introduced by Assumption 1 that contracts do not condition on the history of reports about the manager’s fit does not seem to be very restrictive.

Third, it may be ex ante suboptimal to replace the manager following low cash flow realizations in a period without truthful reporting. Indeed, in the more general case in which relation (1) does not hold, the board’s learning from cash flows about changes in the manager’s fit takes more time. However, introducing this possibility has no bearing on the insight that a contract’s length would be defined as the distance between two periods in which the board can terminate the contract without severance pay. Also in this case, the insight that the board’s willingness to offer severance pay would decrease in the manager’s outside option $U$ and the firm’s cash flow upside $\Delta x$ would be unaffected. Finally, it should be noted that in a more general formulation of the model, the manager and the board could learn about the manager’s fit not only from the information that the manager observes at the interim date $\tau_l = 1$, but also from the firm’s cash flows. Finding a tractable dynamic

\footnote{Indeed, the literature on relational contracts has argued that it may be too complex to condition even on the firm’s cash flows (Levin, 2003; Malcomson, 2016).}

\footnote{An example in which the board would learn also from the firm’s cash flows is when the manager’s fit is continuously distributed on $[\theta_L, \theta_H]$. In the one-shot analogy of this problem, Inderst and Mueller (2010) show that the optimal replacement policy will feature a cutoff type $\theta_c$, such that all types below this cutoff leave with the ex ante stipulated severance pay. In this case, the firm’s cash flow realizations at the end of the period would be further informative about how the manager’s type is distributed on $[\theta_L, \theta_H]$. The first renewal period would correspond 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 $\theta_c$. High cutoffs would imply a weaker relation between turnover and

27
framework to incorporate all these possibilities goes beyond the scope of this paper, but would be a natural step towards extending the baseline model.

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 manager’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 predic-

underperformance. Intuitively, there would be less-negative updating about the manager’s fit following low cash flows since her fit would be coming from a “better” distribution.
tion, 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 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


[38] 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.


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


Appendix A  Omitted Proofs

Proof of Proposition 1.  Part (i) and the first part of part (ii) of the Proposition follows largely from the main text. The proof that conditions (7) is binding and that (8) is binding for $\theta_{t-1} = \theta_C$ are straightforward and subsumed by case 1 of Lemma B.1 in Appendix B. In what follows, it is shown that seeking truthful reporting in all periods while having $T \to \infty$ cannot be optimal for the board. This is shown in two steps. Step 1 introduces some notation, which would be useful throughout the paper, and Step 2 shows the claim by arguing 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 (tr.rep.) and depending on the manager’s
the period, and the two terms cancel out in board’s expected payoff. For market in the period in which she is …red, but the …rm hires a new manager for 

Similarly, we can define the expected amount that the outgoing manager would be paid by 

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

\[ p(w) := \delta p_1 + \sum_{i=2}^{T-1} \delta^i \left( \Pi_{k=1}^{i-1} e_k \right) p_i + \delta^T \left( \Pi_{k=1}^{T-1} e_k \right) 1, \quad (A.1) \]

which corresponds to $E \left[ \sum_{i=1}^{T} \delta^{i-1} q_i \right] \delta = E \left[ \delta q_1 + \sum_{i=2}^{T-1} \delta^i q_i + \delta^T q_T \right]$ in expression (4).

Note that the replacement probability in the manager’s retirement period $T$ is $1 = (1; 1)$. 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

\[ h(U, w) := \delta p_1 \sum_{j=2}^{T} \delta^{j-2} U + \sum_{i=2}^{T-1} \left( \delta^i \left( \Pi_{k=1}^{i-1} e_k \right) p_i \sum_{j=i+1}^{T} \delta^{j-i-1} U \right), \quad (A.2) \]

which corresponds to $E \left[ \sum_{i=1}^{T} \delta^{i-1} (U - q_i U) \right]$ in expressions (3) and (4). It is convenient to

---

25Note that the board will always offer sufficient incentives for $\theta_G$ to invest in firm-specific human capital, as, otherwise, it would be better off hiring a new manager.

26For $T = 1$, we have $p(w) = \delta$, and for $T = 2$, we have $p(w) = \delta p_1 + \delta^2 e_1$.

27To be precise, in case the manager reveals that she is a bad fit, she receives $U$ from the outside labor market in the period in which she is fired, but the firm hires a new manager for $U$ for the remainder of the period, and the two terms cancel out in board’s expected payoff. For $T = 1$, $h(w) = 0$; for $T = 2$, $h(w) = \delta p_1 U$. 
rewrite this expression as

$$
h(U, w) = \frac{\overline{U}}{1-\delta} \left( \delta p_1 (1 - \delta^{T-1}) + \sum_{i=2}^{T-1} \delta^i (\Pi_{k=1}^{i-1} e_k) p_i (1 - \delta^{T-i}) \right)
$$

$$
= \frac{\overline{U}}{1-\delta} \left( -p_1 \delta^T + \sum_{i=2}^{T-1} \delta^i (\Pi_{k=1}^{i-1} e_k) p_i - \delta^T \sum_{i=2}^{T-1} (\Pi_{k=1}^{i-1} e_k) p_i \right)
$$

$$
= \frac{\overline{U}}{1-\delta} \left( -p_1 \delta^T + p(w) - \delta^T \Pi_{k=1}^{T-1} e_k 1 - \delta^T (e_1 - \Pi_{k=1}^{T-1} e_k 1) \right)
$$

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

(A.3)

Denoting further $s(w) := E \left[ \sum_{i=1}^{T} \delta^{i-1} q_i (x_i - c) \right]$ in the board’s equilibrium expected payoff given by expression (4), expression (4) can be stated as

$$
V^* = -U_1(w) + s(w) + h(U, w) + p(w) V^*,
$$

(A.4)

Using (A.3) and expressing $U_1(w)$ as $v_1(w) + \sum_{j=1}^{T} \delta^{j-1} U_j$ using (5), we can simplify (A.4) to

$$
V^* = \frac{s(w) + h(U, w) - U_1(w)}{1 - p(w)} = \frac{s(w) - v_1(w)}{1 - p(w)} - \frac{\overline{U}}{1-\delta}.
$$

(A.5)

The functional dependence on $w$ 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 contract.

**Step 2.** Suboptimality of Seeking Truthful Reporting in all Periods When $T \to \infty$.

We argue to a contradiction. Suppose that the board incentivizes truthful reporting in all periods. We have

$$
s(w) = x + (\bar{\theta} + e_{1} (\theta_G - \bar{\theta})) \Delta x - c
$$

$$
+ \left( \delta e_{1} + \sum_{j=2}^{T-1} \delta^j e_1 \Pi_{i=2}^{j} e_{1} (\theta_G) \right) (x + (\bar{\theta} + e (\theta_G) (\theta_G - \bar{\theta})) \Delta x - c)
$$

$$
= x + (\bar{\theta} + e_{1} (\theta_G - \bar{\theta})) \Delta x - c
$$

$$
+ \left( \delta e_{1} \left( 1 - e_{G}^{T-1} \delta^{T-1} \right) \right) \left( x + (\bar{\theta} + e_G (\theta_G - \bar{\theta})) \Delta x - c \right).
$$

(A.6)

where, given the Markov structure, it is without loss of generality to write for brevity $e_G \equiv \ldots$
Furthermore
\[
1 - p(w) = 1 - \left( \delta (1 - e_1) + \sum_{j=2}^{T-1} \delta^j e_1 e(\theta_G)^{j-2} (1 - e(\theta_G)) + \delta^T e_1 e(\theta_G)^{T-2} \right)
\]
\[
= (1 - \delta) \left( 1 + \frac{\delta e_1 (1 - e(\theta_G)^{T-1})}{1 - e(\theta_G)^{T-1}} \right)
\]
(A.7)

Plugging in for \( s(w) \), \( p(w) \), as well as for \( h(w) \) from (A.3) and \( v_1(w) \) from (11), (A.5) becomes
\[
V^* = \left( \frac{x + (\bar{v} + e_1 (\theta_G - \bar{v})) \Delta x - c}{1 - \Delta \theta e_1 (\bar{v} + e_1 (\theta_G - \bar{v})) \Delta x - c} \right)
\]
\[
+ \left( \frac{(\bar{v} + e_1 (\theta_G - \bar{v})) \Delta x - c}{1 - \Delta \theta e_1 (\bar{v} + e_1 (\theta_G - \bar{v})) \Delta x - c} \right)
\]
\[
- \left( \frac{\theta_{NC}}{e_1 \Delta \theta} - \mathcal{U} + \frac{\delta - \Delta \theta}{1 - \Delta \theta} \left( \frac{\theta_{NC}}{e_1 \Delta \theta} - \mathcal{U} \right) \right)
\]
\[
= (1 - \delta) \left( 1 + \frac{\delta e_1 (1 - e(\theta_G)^{T-1})}{1 - e(\theta_G)^{T-1}} \right)
\]
(A.8)

It is now sufficient to show that letting \( T \to \infty \) can make (A.8) negative. This is the case if
\[
\frac{\theta_{NC}}{e_1 \Delta \theta} - \mathcal{U} > 1 - \delta \left( \frac{(x + (\bar{v} + e_1 (\theta_G - \bar{v})) \Delta x - c)}{1 - e_1 \Delta \theta (\bar{v} + e_1 (\theta_G - \bar{v})) \Delta x - c - \mathcal{U}} \right)
\]
(A.9)

The key observation now is that the RHS of (A.9) decreases towards zero as \( \delta \) increases towards one (in particular, note that \( \frac{1 - \delta}{\Delta \theta} \) and \( \frac{\delta e_1}{1 - e_1 \Delta \theta} \) do not cancel out). By contrast, the LHS of (A.9) is independent of \( \delta \) and positive if (12) is not satisfied. Thus, for any parameter constellation, there is a threshold \( \delta^* \) such that for \( \delta > \delta^* \), this condition is satisfied. Hence, in these cases the board must deviate from a policy of pursuing truthful reporting in all periods and choosing \( T \) as high as possible. Q.E.D.

The following straightforward result is helpful for the proofs of Propositions 2 and 4.

**Lemma A.1** Consider a contract offer \( w_n \) that gives the board an expected payoff of \( V_n^* \) and features \( n \) periods with truthful reporting. Compare this offer to an alternative \( w_{n+k} \) that gives the board \( V_{n+k}^* \) and features \( n + k \) periods with truthful reporting. The attractiveness of
offer \( w_{n+k} \) for the board increases in \( U \) and \( \Delta x \) if

\[
\frac{\partial}{\partial U} (V^*_{n+k} - V^*_n) = \frac{\partial}{\partial U} \left( \frac{-v_1(w_{n+k})}{1 - p(w_{n+k})} - \frac{-v_1(w_n)}{1 - p(w_n)} \right) > 0 \tag{A.10}
\]

\[
\frac{\partial}{\partial \Delta x} (V^*_{n+k} - V^*_n) = \frac{\partial}{\partial \Delta x} \left( \frac{s(w_{n+k})}{1 - p(w_{n+k})} - \frac{s(w_n)}{1 - p(w_n)} \right) > 0. \tag{A.11}
\]

**Proof of Lemma A.1.** The claim follows by standard comparative statics arguments. The increasing differences in (A.10) and (A.11) imply that the \( V^*_{n+k} \) offer becomes increasingly more attractive for the board as \( U \) and \( \Delta x \) increase. To obtain the equalities in (A.10) and (A.11), plug in for \( V^* \) from (A.5). Q.E.D.

**Proof of Proposition 2.** Let the maximum contract length of \( w \) be \( T \), while that of \( w' \) be \( T + 1 \). Using (A.5) to express \( V^* \), we have to show that

\[
\frac{\partial}{\partial U} (V^*(w') - V^*(w)) = \frac{\partial}{\partial w} h(w') (1 - p(w)) - \frac{\partial}{\partial w} h(w) (1 - p(w')) \tag{A.12}
\]

To obtain the equality in (A.12), we use that when the first-best condition (12) does not hold, the manager’s expected payoff \( U_1(w) \) is independent of \( U \) (to see this, add \( \sum_{j=1}^{T} \delta^{j-1} U \) to the manager’s rent in (11)). The increasing difference in (A.12) will imply that the \( T' \)-offer becomes increasingly more attractive as \( U \) increases.

Recalling from (A.3) that \( h(w) = \frac{\partial}{\partial U} \left( -\delta^T + p(w) \right) = \frac{\partial}{\partial U} \left( 1 - \delta^T + p(w) - 1 \right) \) and plugging in from (A.7), we can express the numerator in (A.12) as

\[
\frac{1}{1 - \delta} \left( (1 - \delta^{T+1}) + p(w') - 1 \right) (1 - p(w)) - (1 - \delta^T + p(w) - 1) (1 - p(w'))
\]

\[
= \frac{1}{1 - \delta} \left( (1 - \delta^{T+1}) (1 - p(w)) - (1 - \delta^T) (1 - p(w')) \right)
\]

\[
= (1 - \delta^{T+1}) \left( 1 + \frac{\delta e_1 \left( 1 - (\delta e_G)^{T-1} \right)}{1 - \delta e_G} \right) - (1 - \delta^T) \left( 1 + \frac{\delta e_1 \left( 1 - (\delta e_G)^T \right)}{1 - \delta e_G} \right)
\]

which after some transformations becomes

\[
\frac{\delta^T}{1 - \delta e_G} \left( (1 - \delta) (1 - \delta e_G) + e_1 \left( \delta (1 - \delta) + (e_G)^{T-1} \left( \delta^{T+1} (1 - e_G) - (1 - e_G) \right) \right) \right) \tag{A.13}
\]

\[
\geq 0.
\]

To see the last inequality, observe that expression (A.13) is positive if the term in brackets following \( e_1 \) is positive. If it is negative, expression (A.13) would decrease in \( e_1 \). Thus, at its
minimum for \( e_1 = e_G \); that value would be

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

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 \( \delta = 1 \).\textsuperscript{28} Hence, for any \( T, \delta \in [0, 1] \), \( e_G \in [0, 1] \), \( e_1 \in [0, e_G] \), (A.13) is (weakly) positive, and it is strictly positive for \( e_G < 1 \) and \( \delta < 1 \), implying that we have strictly increasing differences in (A.12). By Lemma A.1, this proves the claim.

Next, we argue that

\[
\frac{\partial}{\partial \Delta x} (V^* (w') - V^* (w)) = \frac{\partial}{\partial \Delta x} s (w') (1 - p (w)) - \frac{\partial}{\partial \Delta x} s (w) (1 - p (w')) (1 - p (w)) (1 - p (w')) > 0. \tag{A.14}
\]

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

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

Plugging in for \( p (w) \), the numerator of (A.14) becomes

\[
\left( (\bar{\theta} + e_1 (\theta_G - \bar{\theta})) + \delta e_1 \left( \frac{1 - e (\theta_G)^{T-1} \delta^{T-1}}{1 - e \theta_G \delta} \right) (\bar{\theta} + e (\theta_G) (\theta_G - \bar{\theta})) \right) (1 - \delta) \left( 1 + \frac{\delta e_1 \left( 1 - e (\theta_G)^{T-1} \delta^{T-1} \right)}{1 - e \theta_G \delta} \right) \\
- \left( (\bar{\theta} + e_1 (\theta_G - \bar{\theta})) + \delta e_1 \left( \frac{1 - e (\theta_G)^{T-1} \delta^{T-1}}{1 - e \theta_G \delta} \right) (\bar{\theta} + e (\theta_G) (\theta_G - \bar{\theta})) \right) (1 - \delta) \left( 1 + \frac{\delta e_1 \left( 1 - e (\theta_G)^{T-1} \delta^{T-1} \right)}{1 - e \theta_G \delta} \right)
\]

\[
= \delta^T e_1 (e_G - e_1) e^{T-1} (1 - \delta) (\theta_G - \bar{\theta}) > 0
\]

proving the claim. Q.E.D.

**Proof of Proposition 3.** For any given reporting policy, the dynamics of the manager’s compensation contract is pinned down by determining which of the constraints bind. The technical derivations are relegated to Lemma B.1 in Appendix B, as they bring little further

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

This term is nonpositive, as the maximum value (of zero) of the numerator is obtained for \( \delta e_G = 1 \).
insight relative to the sketch of the proof presented here and the intuition contained in the main text.

(i) 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 (15) 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 (15) more difficult to satisfy. The key dynamic element is that it is optimal to defer the manager’s bonus — i.e., the board satisfies (13) for $\Delta w_t = 0$ by promising a higher continuation payoff— as this helps relax the constraint that the manager invests in firm-specific human capital in the subsequent period(s).

(ii) 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. Expressing $w_{s,t}$ from (7) and plugging into the constraint that the manager invests in firm-specific human capital, given by (8), we have two cases. First, if there is truthful reporting in $t - 1$, the board’s payoff is increased by reducing $\Delta w_t$ until (8) or $\Delta w_t \geq 0$ becomes binding. Note that it is without loss of generality to stipulate that $w_t$ 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 $\theta_G$ (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 $t - 1$.\(^{29}\) If, instead, period $t$ follows $n \geq 1$ periods in which the board does not offer incentives for truthful reporting (from $t - n$ to $t - 1$), we have an additional constraint, as the bonus and respective continuation payoffs in $t$ must be high enough to satisfy not only (8), but also (13) in the preceding $n$ periods (for which $\Delta w_{t-n} = \ldots = \Delta w_{t-1} = 0$). The latter condition is stated as $\Delta w_t \geq \Delta w_{t-n}^{-1}(0)$. An explicit expression for $\Delta w_{t-n}^{-1}(0)$ is given in Appendix B. Thus, $\Delta w_t$ is determined by the more stringent of condition (8), $\Delta w_t \geq \Delta w_{t-n}^{-1}(0)$, and $\Delta w_t \geq 0$. Also in this case, the manager’s continuation payoff reduces the need to offer her a bonus in $t$, but as Proposition 1 illustrates, the need to satisfy the manager’s incentive constraint (7) could prevent the board from fully deferring the bonus in $t$. Since the manager is dismissed when generating low cash flows in $t - n$ until $t - 1$, the history dependence of $\Delta w_t$ on the cash flows from $t - n$ to $t - 1$ is again trivial.

It remains to argue that the board always seeks truthful reporting in the manager’s

\(^{29}\)Such history dependence may become important in some of the model’s extensions in Section 3.3.
retirement period $T$. Observe, first, that if the board seeks truthful reporting in period $T$, we have

$$\Delta w_T = \max \left\{ \Delta w_{T-n}^{-1} (0), \frac{c}{e_T (\theta_{T-1}^*) \Delta \theta} \right\}$$

$$w_{s,T} = \theta_N \Delta w_T - \overline{U}.$$  

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 (13)–(15), while also making it optimal to set $\Delta w_t = 0$ in the immediately preceding periods without truthful reporting (in case there are such periods). Thus,

$$\Delta w_T = \max \left\{ \Delta w_{T-n}^{-1} (0), \frac{c}{e_T (\theta_{T-1}^*) \Delta \theta} \right\}.$$  

Plugging into the manager’s payoff $U_T$, we obtain that this payoff is identical in both cases. From (16)–(19), 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 4.** Suppose that the manager lives for at most two periods and that the first-best condition (12) 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 3). Thus, consider the case in which $T = 2$. We start by deriving the separate components of (A.5) and then plug into (A.10) and (A.11).

In the final period, the board always seeks truthful reporting by Proposition 3. By (9), this implies that

$$U_2 \left( \theta_1, w \right) = \max \left\{ e_2 \left( \theta_1 \right) (w_2 + \theta_G \Delta w_2) + (1 - e_2 \left( \theta_1 \right)) \left( w_{s,2} + \overline{U} \right) - c, w_{s,2} + \overline{U} \right\}$$

$$= \max \left\{ \left( \theta_N + e_2 \left( \theta_1 \right) \Delta \theta \right) \Delta w_2 - c, \theta_N \Delta w_2 \right\}.$$  

where we use that $w_{s,2} = \theta_N \Delta w_2 - \overline{U}$ and $w_2 = 0$. Furthermore, we have that $\Delta w_2 = \max \left\{ \Delta w_{1}^{-1} (0), \frac{c}{e_2 (\theta_{T-1}^*) \Delta \theta} \right\}$, 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

\[
0 = \frac{c}{\epsilon_1 \Delta \theta} + \delta \frac{U_2^e(\theta_N, w) - U_2^e(\theta_G, w)}{\Delta \theta} \tag{A.15}
\]

\[
= \frac{c}{\epsilon_1 \Delta \theta} + \delta \frac{\theta_N \max \{(\theta_N + e_2(\theta_N) \Delta \theta) \Delta w_2 - c, \theta_N \Delta w_2\} + (1 - \theta_N) \overline{U}}{\Delta \theta} - \theta_G \max \{(\theta_N + e_2(\theta_G) \Delta \theta) \Delta w_2 - c, \theta_N \Delta w_2\} - (1 - \theta_G) \overline{U} \tag{A.16}
\]

We can avoid the max-operator in (A.16) by using the simplified notation that \(\bar{e}_N = e_2(\theta_N)\) (\(\bar{e}_N = 0\) and \(\bar{c} = c\) (\(\bar{c} = 0\)) if the manager invests (does not invest) in firm-specific human capital in \(t = 2\) if her fit in \(t = 1\) is \(\theta_N\). Then, from (A.16), we obtain that \(\Delta w_1^{-1}(0) = \frac{\frac{c}{\epsilon_1 \Delta \theta} + \frac{\theta_G e_2 - \theta_N \bar{c}}{\epsilon_1 \Delta \theta} + \overline{U}}{(\frac{e_G - \bar{e}_N}{\epsilon_G e_G} + \frac{1 - \bar{e}_N}{\theta_N}) - 1} \).

Consider, first, the case in which \(\Delta w_2 = \Delta w_1^{-1}(0)\). We have

\[
\nu_1(w^{nr}) = U_1(w^{nr}) - \sum_{j=1}^{2} \delta^{j-1} \overline{U} = -c + \delta E_1 [U_2^e(\theta_t, w)] - \overline{U} (1 + \delta) \tag{A.17}
\]

\[
= -c + \delta [U_2^e(\theta_N, w) + e_1 (U_{t+1}^e(\theta_G, w) - U_2^e(\theta_N, w))] - \overline{U} (1 + \delta)
\]

\[
= \delta \left( \theta_N \left( \frac{\bar{e}_N \theta_G + (1 - \bar{e}_N) \theta_N}{\frac{e_G - \bar{e}_N}{\epsilon_G e_G} + \frac{1 - \bar{e}_N}{\theta_N} - 1} \right) - c \right) + (1 - \theta_N) \overline{U}
\]

\[
= A_1 + \left( -\delta \theta_N \theta_G \left( \frac{e_G - \bar{e}_N}{\theta_G e_G} \right) - 1 \right) \overline{U}
\]

where \(A_1\) stands for terms that do not depend on \(\overline{U}\); the third equality follows after using from (A.15) that \(U_2^e(\theta_N, w) - U_2^e(\theta_G, w) = \frac{c}{\epsilon_1 \Delta \theta}\) and after plugging for \(\Delta w_2 = \Delta w_1^{-1}(0)\) (cf. (A.16)).

If the board seeks truthful reporting in the first period, then, by (11), the manager’s rent is simply

\[
\nu_1(w^r) = \frac{\theta_N}{\epsilon_1 \Delta \theta} c - \overline{U} + \delta \left( \frac{\theta_N}{\epsilon_G \Delta \theta} c - \overline{U} \right),
\]

where the superscript \(r\) stands for truthful reporting.

\(\text{Note that the board will always offer sufficient incentives for } \theta_G \text{ to invest in firm-specific human capital, as, otherwise, it would be better off hiring a new manager.}\)
By plugging $\nu_1 (w^r)$ and $\nu_1 (w^{nr})$ into (A.10), we obtain that

$$\frac{\partial}{\partial U} \left( -\frac{\nu_1 (w^r)}{1 - p (w^r)} - \frac{-\nu_1 (w^{nr})}{1 - p (w^{nr})} \right)$$

(A.18)

$$= \frac{1 + \delta}{1 - \delta (1 - e_1 - \delta^2 e_1)} \delta \theta_N \theta_G \left( \frac{e_G - \tilde{e}_N}{(\theta_G - \theta_N + (1 - e_1) \theta_N)} + 1 \right)$$

$$= \frac{(1 + \delta) (1 + \delta E_1 \theta)}{(1 - \delta) (1 + \delta e_1) (1 + \delta E_1 \theta)}$$

$$> (1 + \delta) \delta \theta_N \left( \frac{1 - \delta \theta_G}{(1 - \delta) (1 + \delta e_1) (1 + \delta E_1 \theta)} \right) > 0$$

where we use that $E_1 \theta := e_1 \theta_G + (1 - e_1) \theta_N > \theta_N$ and $e_1 < 1$. Hence, (A.18) is positive.

Consider, second, the case in which $\Delta w_2 = \frac{c}{e_2 (\theta_1 - \lambda)}$. As in (A.17), we have then

$$\nu_1 (w^{nr}) = -c + \delta \left[ U_2^G (\theta_N, w) + e_1 (U_2^{nr+1} (\theta_G, w) - U_2^G (\theta_N, w)) \right] - \overline{U} (1 + \delta)$$

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

Approaching (A.11) similarly, we have

$$s (w^r) \left( 1 - p (w^r) \right) - s (w^{nr}) \left( 1 - p (w^{nr}) \right)$$

$$= \frac{x + (e_1 \theta_G + (1 - e_1) \overline{\theta}) \Delta x - c + \delta e_1 (x + (e_G \theta_G + (1 - e_G) \overline{\theta}) \Delta x - c)}{(1 - \delta) (1 + \delta e_1) (1 + \delta E_1 \theta)}$$

$$= \frac{x + (e_1 \theta_G + (1 - e_1) \theta_N) \Delta x - c + \delta \left( e_1 \theta_G (x + (e_G \theta_G + (1 - e_G) \overline{\theta}) \Delta x - c) + (1 - e_1) \theta_N (x + (\tilde{e}_N \theta_G + (1 - \tilde{e}_N) \overline{\theta}) \Delta x - \tilde{c}) \right)}{(1 - \delta) (1 + \delta E_1 \theta)}$$.

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

$$\frac{\Delta x}{(1 - \delta) (1 + e_1 \delta) (1 + \delta E_1 \theta)} \left( \delta (\theta_G - \overline{\theta}) ((e_G - e_1) (e_1 - E \theta) + \theta_N \Delta e (1 - e_1) (1 + \delta e_1)) + (\overline{\theta} - \theta_N) (1 - e_1) \theta_N \right)$$

which is strictly positive as

$$(e_G - e_1) (e_1 - E \theta) + \theta_N \Delta e (1 - e_1) (1 + \delta e_1)$$

$$> (e_G - e_1) (e_1 - e_1 \theta_G - (1 - e_1) \theta_N + \theta_N (1 - e_1))$$

$$= (e_G - e_1) (e_1 - e_1 \theta_G) > 0$$

Hence, expression (A.11) is positive, proving also the second statement. Q.E.D.
Proof of Proposition 5. Recall that from expressions (A.3) and (A.5), we can express

\[ V^* (w) = \frac{s(w) + \frac{T}{1-\delta} (p(w) - \delta^T) - U_1(w)}{1 - p(w)}. \]

Hence, \( \frac{\partial}{\partial \delta} V^*(w) > 0 \) as long as \( (p(w) - \delta^T) > (1 - \delta) \frac{\partial}{\partial \delta} U_1(w) \). This is trivially satisfied if the board stimulates truthful reporting in all periods and the first-best condition (12) is not satisfied. In this case, \( U_1(w) \) is independent of \( \overline{U} \). Q.E.D.
Appendix B  For Online Publication: Supplementary Material

Binding constraints in Proposition 3. In what follows, we take the board’s truthful reporting policy as given and analyze which of the conditions (7)–(15), \(w_t, \Delta w_t, w_{s,t} \geq 0\) are binding when minimizing the manager’s period one payoff \(U_1(w)\).

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 the more stringent of conditions (15) and \(w_t \geq 0\) (if the first-best condition (12) does not hold, \(w_t \geq 0\) is lax). (ii) If the board seeks truthful reporting in period \(t\), \(w_{s,t}\) is determined by the more stringent of condition (7) and \(w_{s,t} \geq 0\) (if the first-best condition (12) does not hold, \(w_{s,t} \geq 0\) is lax); \(\Delta w_t\) is determined by condition (8) if the board seeks truthful reporting in \(t - 1\). Without truthful reporting from \(t - n\) to \(t - 1\) \((n \geq 1)\), \(\Delta w_t\) is determined by the most stringent of conditions (8), the analogue of condition (13) for the preceding \(n\) periods, and \(\Delta w_t \geq 0\) (if the first-best condition (12) does not hold, \(\Delta w_t \geq 0\) is lax).

Proof of Lemma B.1. The Lemma is shown by induction by arguing first that it must be always satisfied in period \(t\) when minimizing \(U_t(\theta_{t-1}, w)\) with respect to the pay components in \(t\). Since minimizing this payoff is also the objective in \(t = 1\), it is without loss of generality to make the argument for \(t = 1\). It is then argued that if the conditions stated in the Lemma are satisfied in \(t\), then minimizing \(U_t(\theta_{t-1}, w)\) with respect to the manager’s pay components in \(t + 1\) requires also that the \(t + 1\) analogue of these conditions be satisfied. Note that these are the same conditions that would have to be satisfied when minimizing \(U_{t+1}(\theta_t, w)\) with respect to the pay components in \(t + 1\). By induction, this means that minimizing the manager’s payoff in period one corresponds to minimizing the manager’s payoff, while satisfying the conditions in the Lemma, in all periods.

If the board offers incentives for truthful reporting in \(t - 1\) (if there is such a period), there are four main cases depending on whether the board offers incentives for truthful reporting in period \(t\), and \(t + 1\). These cases are considered first. The end of the proof considers the case in which the board does not offer incentives for truthful reporting in \(t - 1\).

Truthful reporting in period \(t\). The induction hypothesis for this case is that \(\Delta w_t, w_t, w_{s,t}\) are given by (17)–(19). For completeness, we can state \(w_{s,t}\) as \(\max\{0, \theta_N \Delta w_t + \delta U_{t+1}(\theta_N, w) - \sum_{j=t}^{T} \delta^j U\}\), but we show that the lower bound is never binding if (12) is not satisfied. 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, it is optimal for the board to set \( \Delta w_1 \) minimal subject to the more stringent of conditions (8) and \( \Delta w_1 \geq 0 \).

**Case 1: Truthful reporting in periods \( t \) and \( t + 1 \).** Consistent with the notation in Proposition 1, define \( \tilde{c}_t (\theta_{t-1}) = c_t (\theta_{t-1}) \) (\( \tilde{c} = c \)) if the manager, whose fit in \( t - 1 \) is \( \theta_{t-1} \), invests in firm-specific human capital in \( t \), and \( \tilde{c}_t (\theta_{t-1}) = 0 \) (\( \tilde{c} = 0 \)) 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 \( \theta_N \). The manager’s expected payoff in period \( t \) is then

\[
U_t (\theta_{t-1}, w) = \tilde{c}_t (\theta_{t-1}) (w_t + \theta_G \Delta w_t + \delta U_{t+1} (\theta_G, w)) + (1 - \tilde{c}_t (\theta_{t-1})) \left( w_{s,t} + \sum_{j=t}^{T} \delta^{j-t-1} U_j \right) - \tilde{c}.
\]

Using the induction hypothesis (17)–(19) to plug into \( w_t, \Delta w_t, \) and \( w_{s,t} \), this payoff becomes

\[
\tilde{c}_t (\theta_{t-1}) \left( \frac{\theta_G c}{e_t (\theta_G) \Delta \theta} + \frac{\theta_G \delta U_{t+1} (\theta_N, w) - \theta_N \delta U_{t+1} (\theta_G, w)}{\Delta \theta} \right) + (1 - \tilde{c}_t (\theta_{t-1})) \left( \frac{\theta_N c}{e_t (\theta_G) \Delta \theta} + \frac{\theta_G U_{t+1} (\theta_N, w) - \theta_N U_{t+1} (\theta_G, w)}{\Delta \theta} \right) - \tilde{c}.
\]

if the zero lower bounds of \( \Delta w_t \) 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 3 minimizes

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

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

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

\[
w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2} (\theta_G, w) - \frac{c}{e_{t+1} (\theta^*_t)} \geq w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} U_j. \tag{B.2}
\]

Truthful reporting in period \( t + 1 \) would further require that

\[
w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} U_j \geq w_{t+1} + \theta_N \Delta w_{t+1} + \delta U_{t+2} (\theta_N, w). \tag{B.3}
\]

To find the pay components \( \{w_{s,t+1}, w_{t+1}, \Delta w_{t+1}\} \) 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

\[
L_1(w, \Lambda) = -(\theta_G \tilde{e}_{t+1}(\theta_N) - \theta_N e_{t+1}(\theta_G))(w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2}(\theta_G, w)) - (\theta_G (1 - \tilde{e}_{t+1}(\theta_N) - \theta_N (1 - e_{t+1}(\theta_G))) \left( w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} U_j \right) + \theta_G \tilde{c} - \theta_N c + \lambda \left( w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2}(\theta_G, w) - \frac{c}{e_{t+1}(\theta_G)} - w_{s,t+1} - \sum_{j=t+1}^{T} \delta^{j-t-1} U_j \right) + \mu \left( w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} U_j - w_{t+1} - \theta_N \Delta w_{t+1} - \Delta U_{t+2}(\theta_N, w) \right) + \kappa \Delta w_{t+1} + \rho w_{t+1} + \chi \Delta w_{t+1}
\]

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

\[
\begin{align*}
\frac{\partial L_1(w, \Lambda)}{\partial w_{s,t+1}} &= 0 = -(\theta_G (1 - \tilde{e}_{t+1}(\theta_N)) - \theta_N (1 - e_{t+1}(\theta_G))) - \lambda + \mu + \kappa \\
\frac{\partial L_1(w, \Lambda)}{\partial \Delta w_{t+1}} &= 0 = -(\theta_G \tilde{e}_{t+1}(\theta_N) - \theta_N e_{t+1}(\theta_G)) \theta_G + \lambda \theta_G - \mu \theta_N + \chi \\
\frac{\partial L_1(w, \Lambda)}{\partial w_{t+1}} &= 0 = -(\theta_G \tilde{e}_{t+1}(\theta_N) - \theta_N e_{t+1}(\theta_G)) + \lambda - \mu + \rho,
\end{align*}
\]

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 \( \lambda = \theta_G \tilde{e}_{t+1}(\theta_N) + \theta_N (1 - e_{t+1}(\theta_G)) > 0 \). Thus, \( \rho, \mu, \lambda > 0 \), imply that the board minimizes \( w_{s,t+1}, \Delta w_{t+1}, \) 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 \( \theta_G \), this implies that the board will choose to satisfy (B.2) for \( \theta_G^* = \theta_G \). For completeness, using that (B.3) is binding, note that we obtain

\[
\begin{align*}
w_{s,t+1} &= \theta_N \Delta w_{t+1} + \delta U_{t+2}(\theta_N, w) - \sum_{j=t+1}^{T} \delta^{j-t-1} U_j \\
\Delta w_{t+1} &= \frac{c}{e_{t+1}(\theta_G)} \Delta \theta + \frac{\delta U_{t+2}(\theta_N, w) - \delta U_{t+2}(\theta_G, w)}{\Delta \theta},
\end{align*}
\]

\footnote{Note that this implies that the manager has no incentives to invest in firm-specific human capital if \( \theta_t = \theta_N \), i.e., \( \tilde{e}_{t+1}(\theta_N) = 0 \) and \( \tilde{c} = 0 \).}
which are the $t+1$ analogues of (17) and (19).

Consider, now, the case in which the zero lower bound of $\Delta w_t$ is binding. Then the manager’s expected payoff is simply

$$\tilde{\pi}_t (\pi_{t-1}) \delta U_{t+1} (\pi_G, w) + (1 - \tilde{\pi} (\pi_{t-1})) U_{t+1} (\pi_N, w) - \tilde{c},$$

which is clearly minimized when $w_{t+1}$, $\Delta w_{t+1}$, 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, expression (9) shows 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.$^{32}$

**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

$$U_{t+1} (\pi_t, w) = \max \left\{ w_{t+1} + (\pi_N + e_{t+1} (\pi_t) \Delta \pi) \Delta w_{t+1} - c + \delta E_{\theta_t} [U_{t+2}^e (\pi_{t+1}, w)], \right\}$$

(B.4)

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

$$w_{t+1} + (\pi_N + e_{t+1} (\pi_t^*) \Delta \pi) \Delta w_{t+1} + \delta E_{\theta_t} [U_{t+2}^e (\pi_{t+1}, w)] - c \geq w_{t+1} + \pi_N \Delta w_{t+1} + \delta U_{t+2}^e (\pi_N, w).$$

which can be restated as

$$\Delta w_{t+1} \geq \frac{c}{e_{t+1} (\pi_t^*) \Delta \pi} + \delta \frac{U_{t+2}^e (\pi_N, w) - U_{t+2}^e (\pi_G, w)}{\Delta \pi}.$$  

(B.5)

In analogy to (15), we further need to satisfy the interim participation constraint in $t+1$

$$w_{t+1} + \pi_N \Delta w_{t+1} + \delta U_{t+2}^e (\pi_N, w) \geq w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} U_j.$$  

(B.6)

$^{32}$It is straightforward to verify that $\Delta w_t > 0$ and $\Delta w_{t+1} > 0$ when the constrain that the manager invests in firm-specific human capital in $t+1$ is binding.
Hence, when the zero lower bound of $\Delta w_t$ is not binding, we need to minimize (B.1), subject to (B.5), (B.6), and feasibility. Define\textsuperscript{33} 

$$L_2 = -\theta_G \left( w_{t+1} + (\theta_N + \tilde{c}_{t+1} (\theta_N) \Delta \theta) \Delta w_{t+1} + \delta \tilde{E}_{\theta_N} \left[ U_{t+2}^e (\theta_{t+1}, w) \right] - \tilde{c} \right)$$  
\[+ \theta_N \left( w_{t+1} + (\theta_N + c_{t+1} (\theta_G) \Delta \theta) \Delta w_{t+1} + \delta E_{\theta_G} \left[ U_{t+2}^e (\theta_{t+1}, w) \right] - c \right)\]  
\[+ \lambda \left( \Delta w_{t+1} - \frac{c}{e_{t+1} (\theta_G^*)} \Delta \theta - \delta \frac{U_{t+2}^e (\theta_N, w) - U_{t+2}^e (\theta_G, w)}{\Delta \theta} \right)\]  
\[+ \mu \left( w_{t+1} + \theta_N \Delta w_{t+1} + \delta U_{t+2}^e (\theta_N, w) - w_{s,t+1} - \sum_{j=t+1}^{T} \delta^{j-t-1} U_j \right)\]  
\[+ \kappa w_{s,t+1} + \rho w_{t+1} + \chi \Delta w_{t+1}\]

giving the first-order conditions

$$\frac{\partial L_2}{\partial w_{t+1}} = 0 = -\Delta \theta + \mu + \rho$$  
$$\frac{\partial L_2}{\partial w_{s,t+1}} = 0 = -\mu + \kappa$$  
$$\frac{\partial L_2}{\partial \Delta w_{t+1}} = 0 = -\left( \theta_G (\theta_N + \tilde{c}_{t+1} (\theta_N) \Delta \theta) - \theta_N (\theta_N + e_{t+1} (\theta_G) \Delta \theta) \right) + \lambda + \theta_N \mu + \chi$$  
\[= -\left( \theta_N - e_{t+1} (\theta_G) \theta_N + \tilde{c}_{t+1} (\theta_N) \theta_G \right) \Delta \theta + \lambda + \theta_N \mu + \chi\]

The first FOC implies that $\mu$ or/and $\rho$ 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 $\lambda$, $\mu$, and/or $\chi$ are positive. If we don’t have first-best ($\mu = 0$), this means that $\Delta w_{t+1}$ 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.\textsuperscript{34} Similar to the final paragraph in 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 (14), subject to (13), (15), $w_t, \Delta w_t, w_{s,t} \geq 0$. The constraint that the manager invests in firm-specific human capital (13), together with the feasibility restriction $\Delta w_t \geq 0$ can be

\textsuperscript{33}\(\tilde{E}_{\theta_N}\) is defined as $E_{\theta_N}$ but for $\tilde{c}_{t+1} (\theta_N)$.

\textsuperscript{34}As in case 1, the board will satisfy (B.5) only for $\theta_G^* = \theta_G$. 

47
stated as
\[
\Delta w_t \geq \max \left\{ 0, \frac{c}{e_t(\theta_{t-1}^e)} + \delta \left( U_{t+1}^e (\theta_N, w) - U_{t+1}^e (\theta_G, w) \right) \right\}.
\]  
(B.7)

Using that \( U_{t+1}^e (\theta_t, w) = \theta_t U_{t+1} (\theta_t, w) + (1 - \theta_t) \sum_{j=t+1}^T \delta_j \delta^{j-1} \), the zero lower bound of \( \Delta w_t \) in (B.7) is binding if
\[
\frac{c}{e_t(\theta_{t-1}^e)} + \delta \left( \theta_N U_{t+1} (\theta_N, w) - \theta_G U_{t+1} (\theta_G, w) + \Delta \theta \sum_{j=t+1}^T \delta_j \delta^{j-1} \right) \leq 0.
\]  
(B.8)

Given the assumption of truthful reporting in \( t-1 \), the induction hypothesis is that \( w_{s,t} = 0 \), \( \Delta w_t \) is given by (B.7), and \( w_t = 0 \). Clearly, in \( t = 1 \), the choice of \( w_1 \), \( \Delta w_1 \), 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 \( \Delta w_1, w_{s,1} \) and \( w_1 \), subject to (13) and (15) proving the first induction step.

Using the induction hypothesis to plug into the manager’s payoff (14) in period \( t \), we have
\[
\left( \theta_N + \tilde{c}_t \right) \left( \theta_{t-1} \Delta \theta \right) \max \left\{ 0, \frac{c}{e_t(\theta_{t-1}^e)} + \delta \left( U_{t+1}^e (\theta_N, w) - U_{t+1}^e (\theta_G, w) \right) \right\} - \tilde{c}
\]
\[
+ \tilde{c}_t \left( \theta_{t-1} \right) \delta U_{t+1}^e (\theta_G, w) + (1 - \tilde{c}_t \left( \theta_{t-1} \right)) \delta U_{t+1}^e (\theta_N, w)
\]
\[
= \begin{cases} 
\tilde{c}_t \left( \theta_{t-1} \right) \delta U_{t+1}^e (\theta_G, w) + (1 - \tilde{c}_t \left( \theta_{t-1} \right)) \delta U_{t+1}^e (\theta_N, w) - \tilde{c} & \text{if } \Delta w_t \geq 0 \\
\tilde{c}_t \left( \theta_{t-1} \right) \delta U_{t+1}^e (\theta_G, w) + (1 - \tilde{c}_t \left( \theta_{t-1} \right)) \delta U_{t+1}^e (\theta_N, w) - \tilde{c} & \text{if } \Delta w_t = 0
\end{cases}.
\]  
(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 \( U_{t+1}^e (\theta_t, w) \) 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 \( \Delta w_{t+1} \) 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
\[
\theta_G U_{t+1}^e (\theta_N, w) - \theta_N U_{t+1}^e (\theta_G, w)
\]  
(B.10)

\[
= \theta_G \theta_N \left( U_{t+1} (\theta_N, w) - U_{t+1} (\theta_G, w) \right) + \Delta \theta \sum_{j=t+1}^T \delta_j \delta^{j-1} \]
or, thus, equivalently minimize $U_{t+1}(\theta_N, w) - U_{t+1}(\theta_G, w)$, 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

$$L_3(w, \Lambda)$$

$$= - (\bar{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G)) (w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2}(\theta_G, w)) + \bar{c} - c$$

$$- ((1 - \bar{e}_{t+1}(\theta_N)) - (1 - e_{t+1}(\theta_G))) \left( w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} U_j \right)$$

$$+ \lambda \left( w_{t+1} + \theta_G \Delta w_{t+1} + \delta U_{t+2}(\theta_G, w) - \frac{c}{e_{t+1}(\theta^*_t)} - w_{s,t+1} - \sum_{j=t+1}^{T} \delta^{j-t-1} U_j \right)$$

$$+ \mu \left( w_{s,t+1} + \sum_{j=t+1}^{T} \delta^{j-t-1} U_j - w_{t+1} - \theta_N \Delta w_{t+1} - \delta U_{t+2}(\theta_N, w) \right)$$

$$+ \kappa w_{s,t+1} + \rho w_{t+1} + \chi \Delta w_{t+1}$$

$$+ \sigma \left( \theta_N U_{t+1}(\theta_N, w) - \theta_G U_{t+1}(\theta_G, w) + \Delta \theta \sum_{j=t+1}^{T} \delta^{j-t-1} U_j + \frac{c}{e_{t+1}(\theta^*_t)} \right),$$

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

$$\frac{\partial L_3(w, \Lambda)}{\partial w_{s,t+1}} = 0 = - ((1 - \bar{e}_{t+1}(\theta_N)) - (1 - e_{t+1}(\theta_G)))$$

$$+ \sigma (\theta_N (1 - \bar{e}_{t+1}(\theta_N)) - \theta_G (1 - e_{t+1}(\theta_G))) - \lambda + \mu + \kappa$$

$$\frac{\partial L_3(w, \Lambda)}{\partial \Delta w_{t+1}} = 0 = - (\bar{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G)) \theta_G$$

$$+ \sigma (\theta_N \bar{e}_{t+1}(\theta_N) - \theta_G e_{t+1}(\theta_G)) \theta_G + \theta_G \lambda - \theta_N \mu + \chi$$

$$\frac{\partial L_3(w, \Lambda)}{\partial w_{t+1}} = 0 = - (\bar{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G))$$

$$+ \sigma (\theta_N \bar{e}_{t+1}(\theta_N) - \theta_G e_{t+1}(\theta_G)) + \lambda - \mu + \rho$$

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 $\Delta w_{t+1}$ (and the continuation payoffs $U_{t+2}(\theta_{t+1}, w)$) must be chosen such that the more stringent of (B.8) and (B.2) is binding. Furthermore, if $\mu > 0$ and $\kappa = 0$ (i.e., first-best is
not achievable and \(w_{s,t+1}\) is given by (B.6), then \(\rho > 0\) and so \(w_t = 0\). Finally, in this case,

\[
\lambda = \frac{(\tilde{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G)) \theta_G + \sigma (\theta_G e_{t+1}(\theta_G) - \theta_N \tilde{e}_{t+1}(\theta_N)) \theta_G + \theta_N \mu - \chi}{\theta_G}
\]

and so \(\lambda\) is positive unless only \(\chi\) is positive. Thus, \(\Delta w_{t+1}\) is given by the more stringent of (B.2) and \(\Delta w_{t+1} \geq 0\) as was to be shown. If, instead, \(w_{s,t+1} = 0\) (i.e., \(\mu = 0\) and \(\kappa > 0\)), then as in the previous cases, first-best is achievable (this occurs if (12) is satisfied).

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

\[
L_4(w, \Lambda) = -(\tilde{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G)) \Delta \theta \Delta w_{t+1} - \delta \tilde{E}_{\theta_N} [U_{t+1}^c(\theta_{t+1}, w)] + \delta E_{\theta_G} [U_{t+1}^c(\theta_{t+1}, w)] - \tilde{c} + c \]

\[
+ \lambda \left( \frac{\Delta w_{t+1} - c}{\epsilon_{t+1}(\theta^*_t) \Delta \theta} - \delta \frac{U_{t+1}^c(\theta_N, w) - U_{t+1}^c(\theta_G, w)}{\Delta \theta} \right) \]

\[
+ \mu \left( w_{t+1} + \theta_N \Delta w_{t+1} + \delta U_{t+1}^c(\theta_N, w) - w_{s,t+1} - \sum_{j=t+1}^T \delta^{j-t-1} U_j \right) \]

\[
+ \kappa w_{s,t+1} + \rho w_{t+1} + \chi \Delta w_{t+1} \]

\[
+ \sigma \left( \theta_N U_{t+1}(\theta_G, w) - \theta_G U_{t+1}(\theta_N, w) + \Delta \theta \sum_{j=t+1}^T \delta^{j-t-1} U_j + \frac{c}{\delta e(\theta^*_t)} \right).
\]

Taking the first order condition with respect to \(\Delta w_{t+1}\), we have

\[
\frac{\partial L_4(w, \Lambda)}{\partial \Delta w_{t+1}} = 0 = -(\tilde{e}_{t+1}(\theta_N) - e_{t+1}(\theta_G)) \Delta \theta + \lambda + \mu \theta_N + \chi
\]

\[
+ \sigma (\theta_N (\theta_N + e_{t+1}(\theta_G) \Delta \theta) - \theta_G (\theta_N + \tilde{e}_{t+1}(\theta_N) \Delta \theta)).
\]  

Since the first line in (B.11) is positive, it must be that \(\sigma > 0\) as the term following \(\sigma\) can be rewritten as \(- (\theta_N (1 - e_{t+1}(\theta_G)) + \theta_G \tilde{e}_{t+1}(\theta_N)) \Delta \theta\), 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 \(\Delta w_t = 0\). Furthermore, \(\frac{\partial L_4(w, \Lambda)}{\partial w_{t+1}} = \mu + \rho - \Delta \theta \sigma\) and \(\frac{\partial L_4(w, \Lambda)}{\partial w_{s,t+1}} = -\mu + \kappa\). 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

$$0 \geq \frac{c}{e_{t-1}(\theta^*_t)} + \delta (U^e_t(\theta_N, w) - U^e_t(\theta_G, w))$$

$$= \frac{c}{e_{t-1}(\theta^*_t)} + \delta \begin{pmatrix} \theta_N \\ \theta_G \end{pmatrix} \begin{pmatrix} -\bar{c} + (\bar{c}_t(\theta_N) \theta_G + (1 - \bar{c}_t(\theta_N)) \theta_N) \Delta w_t \\ -c + e_t(\theta_G) \theta_G + (1 - e_t(\theta_G)) \theta_N) \Delta w_t \\ + \delta (\bar{c}_t(\theta_N) U_{t+1}(\theta_G, w) + (1 - \bar{c}_t(\theta_N)) U_{t+1}(\theta_N, w)) \end{pmatrix} + \delta \Delta \theta \sum_{j=t}^T \delta^{j-t} U_j$$

Defining

$$\Delta w_{t-1}^{-1}(0) := \frac{\frac{c}{e_{t-1}(\theta^*_t)} + \delta (\theta_G - \theta_N \bar{c}) + \delta \Delta \theta \sum_{j=t}^T \delta^{j-t} U}{(\theta_N e_t(\theta_N) - \theta_G e_t(\theta_G)) U_{t+1}(\theta_G, w)}$$

we, therefore, need in addition to the analysis in Cases 1 and 2 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

$$\Delta w_t = \max \left\{ \frac{c}{e_t(\theta^*_t) \Delta \theta} + \frac{U_{t+1}(\theta_N, w) - U_{t+1}(\theta_G, w)}{\Delta \theta}, \Delta w_{t-1}^{-1}(0), 0 \right\} . \tag{B.13}$$

After some transformations, it is now straightforward to show that, regardless of $\theta^*_t$ and whether the first or second term is larger, the objective to minimize $U_t$ again boils down to minimizing (B.1). If $\Delta w_t = 0$, the result is immediate as in Case 1.

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

$$\Delta w_{t-1}^{-1}(0) := \frac{\frac{c}{e_{t-1}(\theta^*_t)} + \delta (\theta_G - \theta_N \bar{c}) + \delta \Delta \theta \sum_{j=t}^T \delta^{j-t} U}{(\theta_N e_t(\theta_N) - \theta_G e_t(\theta_G)) U_{t+1}(\theta_G, w)}$$

We can now augment the induction hypothesis with

$$\Delta w_t = \max \left\{ 0, \frac{c}{e_t(\theta^*_t) \Delta \theta} + \frac{U^e_{t+1}(\theta_N, w) - U^e_{t+1}(\theta_G, w)}{\Delta \theta}, \Delta w_{t-1}^{-1}(0) \right\} . \tag{B.14}$$

After some transformations, it can be shown that, 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

\[
\frac{e_N \theta_G}{e_N \theta_G + (1 - e_N) \theta_N} e_G + \frac{(1 - e_N) \theta_N}{e_N \theta_G + (1 - e_N) \theta_N} e_N > e_1 > \frac{e_G (1 - \theta_G)}{e_G (1 - \theta_G) + (1 - e_G) (1 - \theta_N) e_N} + \frac{(1 - e_G) (1 - \theta_N)}{e_G (1 - \theta_G) + (1 - e_G) (1 - \theta_N) e_N}.
\]

(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 \( \theta_G \). 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 \( \theta_N \). These conditions are captured by (B.15). To see that these conditions can be satisfied, let \( e_G = e_1 + \varepsilon \), and \( e_N = e_1 - \varepsilon \). For \( \varepsilon = 0 \), the LHS and the RHS of (B.15) are equal to \( e_1 \). Furthermore, there is a threshold \( \tilde{\varepsilon} \), such that the LHS increases in \( \varepsilon \), while the RHS decreases in \( \varepsilon < \tilde{\varepsilon} \). Q.E.D.

Calculating the Board’s Expected Payoff in Section 3.2.2 In equilibrium, the board’s expected payoff is \( V^* \). Hence, we can rewrite (4) as

\[
V^* = \sum_{i=1}^{2} \delta^{i-1} \frac{\bar{U} - U_1(w) + E \left[ \sum_{i=1}^{2} \delta^{i-1} \tilde{q}_i (x_i - c - \bar{U}) \right]}{1 - E \left[ \sum_{i=1}^{2} \delta^{i-1} \tilde{q}_i \right] \delta}.
\]

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

\[
V^* = \frac{-U_1(w) + \left( x + (\bar{\theta} + e_1 (\theta_G - \bar{\theta})) \Delta x - c \right) + \delta e_1 (\bar{\theta} + e_2 (\theta_G - \bar{\theta})) \Delta x - c}{1 - \delta (1 - e_1) - \delta^2 e_1}.
\]

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

\[
V^* = \frac{-U_1(w) + \left( (x + (\theta_N + e_1 (\theta_G - \theta_N)) \Delta x - c \right) + \delta e_1 (\bar{\theta} + e_2 (\theta_G - \bar{\theta})) \Delta x - c) \right) + \delta (1 - e_1) \theta_N (x + (\bar{\theta} + e_2 (\theta_N) (\theta_G - \bar{\theta})) \Delta x - c)}{1 - \delta (e_1 (1 - \theta_G) + (1 - e_1) (1 - \theta_N)) \bar{U} + \delta^2 (e_1 \theta_G + (1 - e_1) \theta_N))}.
\]
By plugging in for \( \{x, \Delta x\} \), \( \bar{U} \), and \( U_1(w) \), we obtain the values in Table 1.