

The dynamics of promotions, quits and layoffs

Odilon Câmara*
Marshall School of Business
University of Southern California

Dan Bernhardt
University of Illinois
at Urbana-Champaign

July, 2009

Abstract

Which manager should a firm promote to CEO? How do the attributes of a managerial workforce affect firms' placement decisions and wage offers, and managers' quit decisions? The existing literature offers trivial answers to this nuanced question. In one class of models, placement decisions are determined on an individual-by-individual basis, so that promotion of one worker does not crowd out another. In the other class — rank order tournaments — the answer is simply the manager who is believed to be most able. The real world is far more complicated, as the firm-wide distribution of all managerial attributes — including ability, age and complementarities — enters the placement/wage decisions by firms and the quit decisions by managers.

This paper develops a rich dynamic model to get at these issues. Our OLG model features two division managers and a CEO, where each executive may be at a different point in his employment horizon. Taking into account the attributes of *all* executives, the firm decides which manager to promote to CEO, which manager(s) to lay off, and what wage offers to make to different executives; while managers who were not laid off decide whether to stay at the firm or to quit and take an outside employment offer. We analyze how the decisions of the firm and managers vary according to the age-skill profiles of the firm's executives.

We provide sufficient conditions for the existence of a unique equilibrium, in which: (i) the firm has a bias toward promoting older managers to CEO, even in the absence of learn-by-doing; (ii) the firm's strategy is monotone on manager's own skill, so that if a manager of a given age/ability is promoted, then so are more able managers of the same age, but (iii) the firm's strategy is a complex, possible non-monotonic function of the attributes of other executives; (iv) a middleaged manager is more likely to stay at the firm when the other manager is older or sufficiently less talented, and when (v) the CEO is older. We then characterize how the firm trades off between wages and placement offers to align managerial incentives. In particular, the firm uses the flexibility in wage offers to "smooth" the discrete incentives generated by placement offers.

Keywords: Tournament, promotion, executive compensation, CEO.

JEL classification: C73, D21, J31, J63, L2, M5.

*Corresponding Author: Marshall School of Business, University of Southern California, 3670 Trousdale Parkway, BRI Room 308, Los Angeles, CA 90089-0804. ocamara@marshall.usc.edu

1 Introduction

Which manager should a firm promote to CEO? How do the attributes of a managerial workforce affect firms' placement decisions and wage offers, and managers' quit decisions? The existing literature offers trivial answers to these nuanced questions. In one class, placement decisions are made by the firm on an individual-by-individual basis, so that promotion of one worker does not crowd out another — the firm promotes workers whose abilities exceed possibly individual-specific ability thresholds. Classic papers in this class include Waldman (1984), Bernhardt and Scoones (1993), Bernhardt (1995), Gibbons and Waldman (1999) and Bai and Wang (2003). In the other class — rank order tournaments — the answer is simply the manager who is believed to be most able. Classic references are Lazear and Rosen (1981), Carmichael (1983), Rosen (1986), Prendergast (1993), Demougin and Siow (1994), and Zájbojník and Bernhardt (2001). Moreover, introducing quits to these models would have no qualitative effect as long as promoted workers become less likely to leave the firm: it would remain optimal to promote the most able manager.

The real world is far more complicated, as the firm-wide distribution of all managerial attributes — including ability, age and complementarities — enters the placement/wage decisions by firms and the quit decisions by managers. For example, promoting a young “hot shot” manager is likely to lock up the CEO position for years. Many firms lose valuable top executives who quit after the appointment of a young CEO — the top executives know they are unlikely to be promoted to CEO within the firm before their work horizon ends¹. A middleaged manager who could be promoted to CEO in the future also considers the attributes of fellow managers. If the other managers are older and likely to retire soon, they will not compete for the CEO position. However, if the other managers are also middleaged, then they will be competitors for the promotion if they do not depart before the CEO position opens up. In this case, the probability that a middleaged manager leaves the firm will depend on the likelihood that the other managers stay — managerial decisions become intertwined in complex ways.

This paper develops a rich dynamic model to get at these and other issues. Our OLG model features two division managers and a CEO, where each executive may be at a different point in his employment horizon. Taking into account the attributes — skill and age — of *all* executives, the firm decides which manager to promote to CEO, which manager(s) to lay off, and what wage offers to make to different executives; while managers who were not laid off decide whether to stay at the firm or to quit and take an outside employment offer. We analyze how the decisions of the

¹Hayes, Oyer and Schaefer (2006) find that both the smallest probability of CEO turnover and the largest probability of turnover of top managers occur when the CEO is young.

firm and managers vary according to the age-skill profiles of the firm's executives.

In our model, the firm solves a complex dynamic profit-maximization problem. A firm's placement/wage offer to a manager and the manager's attributes affect not only the manager's decision of whether to stay or quit, but also the decisions of all *other* current and future managers. When determining whom to promote and whom to lay off, the firm must consider the ramifications for what its managers do in the future, in particular their quit decisions, which, in turn, will affect the future set of CEO candidates. As a result, the firm's employment decisions must take into account the entire cross-sectional distribution of the attributes of its executives. Managers also have a hard problem to solve. A manager must take into account how his attributes, the attributes of other managers, and the firm's placement/wage offers affect the quit decisions of other managers, which influence the attributes of those managers who will compete for the next CEO promotion.

The firm has two tools at hand with which to provide incentives to executives — its placement decisions and its wage offers to experienced managers. To highlight the trade off between providing incentives via placement or wage offers, and the implications of constraints on wage offers, we contrast (a) a perfectly rigid wage setting, where the firm has no wage flexibility and must offer all managers the same fixed wage offer, so that placement offers are the only instrument to provide incentives for workers, with (b) a perfectly flexible wage setting, where the firm can combine optimally-designed wage offers with placement to provide incentives to executives. In particular, with flexible wages, the firm's wage offer to an experienced manager can and will depend not only on his own attributes and placement, but also on the attributes and placement of others in the firm.

We solve the firm's and managers' problems and provide sufficient conditions on primitives such that a Perfect Bayesian Equilibrium exists. We then provide sufficient conditions on primitives such that the equilibrium is unique and features several key placement and payoff monotonicity properties. Placement monotonicity implies that, given the attributes of *all* executives, the firm's optimal placement offer is characterized by two cutoff rule functions. A manager is promoted to CEO if his skill exceeds the higher promotion cutoff; is offered the managerial position if his skill is intermediate; and is laid off if his skill falls below the lower cutoff. Payoff monotonicity implies that the firm's optimal wage offer to a manager gives him an expected payoff from staying in the firm that is weakly increasing in his own skill; and, in equilibrium, firm profits are weakly increasing in the skill of each executive.

We then present five analytical results that characterize central features of equilibrium outcomes when the CEO position is important, that is, when the CEO's wage and relative contribution to the

firm's output are high. The results describe how the probability that each manager quits vary with the attributes of *all* executives and the probability of CEO promotion. They also characterize how the firm combines placement and wage offers to provide incentives to executives — by exploiting the effect of offers on quit-decisions of all managers — and how the firm's use of these incentives varies with the attribute profiles of executives. The firm uses these tools to search for future CEOs, to nurture talented executives, and to optimally adjust the level of competition for the CEO position.

R1. *The firm has a bias toward promoting old executives to CEO.*

The model dynamics imply a complicated trade-off between promoting a middleaged or an older manager to CEO. A middleaged manager has a longer remaining work horizon. Hence, if promoted, he could stay longer in the more important CEO position. Furthermore, if not promoted, a middleaged manager may receive a better outside offer and quit. These considerations could bias the firm in favor of promoting the younger manager. However, retention concerns for talented older managers who are not promoted are even greater: older managers will retire next period and will never receive the CEO prize from the firm, so they are especially likely to leave if not promoted. In contrast, able middleaged managers have good chances of promotion after the retirement of an old CEO — they have a high expected payoff from staying in the firm, and hence are less likely to quit. Thus, our theory reconciles Belzil and Bognanno's (2004) empirical finding that for executives in the top six hierarchical levels in the firm, controlling for tenure and hierarchical level, an executive's promotion probability increases with his age².

When wages are perfectly flexible, the firm uses wages to partially compensate an old manager who is not promoted to CEO. However, the bias remains because of the same principle: an old manager who is not promoted lost the competition. He lost the expected payoff in the form of probability of promotion, while a middleaged manager still considers that expected payoff in his quit decision. Hence, it is more profitable to promote the old manager and marginally change the current wage of the younger manager to adjust his probability of quitting.

R2. *A talented middleaged manager is less likely to leave the firm if the CEO is old.*

An old CEO will retire soon, and the prospect of imminent promotion to CEO increases the incentives of a talented middleaged manager to stay. Moreover, a firm with an older CEO is more

²Belzil and Bognanno also find that without controlling for other variables, the average age of CEOs is 55.8, while that of executives reporting directly to the CEO is 50.4. Notice that without controlling for hierarchical level, younger workers are more likely to be promoted than older workers simply because younger workers are more likely to be at lower hierarchical levels in the firm, and probability of promotion decreases with hierarchical level.

concerned with retaining talented middleaged managers who can become the next CEO, and hence will offer wages to middleaged managers that yield higher expected discounted payoffs from staying in the firm. This equilibrium result can reconcile Hayes, Oyer and Schaefer’s empirical finding that, controlling for many other observable characteristics including tenure, turnover of top managers is higher when the CEO is younger. More generally, their results support the idea that not only do the specific characteristics of a top executive play a significant role in explaining his career dynamics, but so too do the characteristics of the other top executives.

R3. *An old manager is more likely to leave the firm if the CEO is old.*

The firm must consider the necessity of seeking and training future CEOs. These incentives cause the firm to aggressively replace old managers with potentially highly talented young managers — there is an option value embedded in firm’s on-the-job skill assessment of new managers. Because the firm has a limited number of “slots” to place and train managers, it may be optimal to lay off an existing old manager who will retire before having another chance of promotion to CEO and replace him with a new hire, so the young manager can acquire learn-by-doing experience in the firm and the firm can assess his skill level. This incentive to replace old managers with new hires is stronger if the current CEO is closer to retirement.

Managers must consider this possibility of being laid off in the future when evaluating their internal and external employment alternatives. When wages are rigid, the firm simply lays off the undesirable old managers. When wages are perfectly flexible, the firm can profitably keep some less talented old managers by offering them lower wages.

R4. *Given any wage offer, a middleaged manager is more likely to stay if the firm retains another manager who is either (i) older and closer to retirement, or (ii) middleaged and sufficiently less talented.*

For a middleaged manager, an older or less able manager represents lesser competition for future CEO promotions. If the old manager stays, he will then retire and hence not compete for the promotion. If a less able middleaged manager stays, both executives will be the same age, and the firm will promote the most able. In both cases, if the less competitive manager quits, he is replaced by a young manager, who could turn out to be very talented and win the CEO promotion. Hence, a middleaged manager must account for the probability that the other manager will quit. In particular, a less able middleaged manager knows that he will not be promoted if the able manager stays in the firm; and this increases his probability of leaving the firm, so that quit decisions become intertwined.

When wages are fixed, the firm's discrete choice of whether or not to keep the less competitive manager creates a discontinuous incentive for a talented middleaged manager. When wages are flexible, the firm uses its wage offers to *both* managers to optimally adjust incentives and quit probabilities.

R5. *The firm's optimal placement/wage offers to a manager are complex, potentially non-monotonic functions of the attributes of all other executives.*

Although we identify conditions that ensure the firm's optimal strategy is monotone with respect to a manager's own skill, this strategy may be a non-monotonic function of the attributes of other executives. To understand why, consider a firm with an old CEO, an old manager, and an able middleaged manager whom the firm wants to retain. To increase the probability that the middleaged manager stays, the firm combines the direct incentive of a higher wage offer with the indirect incentive of its placement/wage offer to the old manager. The firm uses its knowledge of the equilibrium behavior of the middleaged manager: he is more likely to stay if the old manager is likely to stay. The firm increases the wage of the old manager as a function of the skill of the middleaged manager up to some point, in order to increase the probability that the old manager will stay and provide incentives to the able middleaged manager. After this point, however, the middleaged manager is so likely to be promoted to CEO that the firm no longer needs to keep the old manager as an indirect incentive. The firm's wage offer to the old manager then becomes decreasing in the skill of the middleaged manager. Hence, the firm's placement/wage offer to the old manager is a non-monotonic function of the middleaged manager's skill.

The paper is organized as follows. Section 2 describes the model, presents the existence result and provides sufficient conditions on primitives such that the equilibrium is unique and features several key monotonicity properties. Section 3 provides the analytical results that characterize important features of equilibrium outcomes when the CEO position is important. We numerically solve the model and characterize the equilibrium for a series of economically relevant primitives. Section 4 concludes. All proofs are in the Appendix A; Appendix B describes a full-fledged version of our model, and shows how the model becomes intractable, but that our simplified, tractable model maintains the key incentive considerations of the full model.

2 The Model

2.1 Agents and Preferences

We consider a risk neutral firm with two divisions that maximizes discounted expected lifetime profits. Each period it uses three executives to produce output: one CEO and two managers — one manager for each of its two divisions, $d = 1, 2$. The firm operates for N periods. For most of our paper, we consider N arbitrarily large, but finite. Production starts at period $t = 1$, and at period $t = N$ the firm closes and lays off all remaining employees.

Each executive maximizes discounted expected lifetime wages. Workers share a common time discount factor $\beta \in (0, 1)$ with the firm. An executive’s productive life has three periods, $a \in \mathcal{A} \equiv \{1, 2, 3\}$, and we refer to executives as young ($a = 1$), middleaged ($a = 2$), and old ($a = 3$). Executives must retire after $a = 3$. We use $a = 0$ to indicate that an executive position is vacant, and refer to middleaged and old managers as “experienced” executives. This framework is the simplest one in which age and ability enter non-trivially into placement decisions.

2.2 Output Technology and Endowment

The productive characteristics of each executive are captured by two variables:

Executive Learn-by-Doing: Via on-the-job learning-by-doing, an executive develops productivity $l_a \in \mathcal{L} \equiv \{l_1, l_2, l_3\}$ during his a^{th} period working as an executive. This productivity is non-decreasing with experience level, $l_3 \geq l_2 \geq l_1$.

Executive-Skill Level: When hired by the firm, a young employee draws skill $s \in \mathcal{S}_1 = [\underline{S}, \bar{S}]$, $0 < \underline{S} < \bar{S}$, during his first term as an executive. Skill s is a random variable, independently and identically distributed according to the continuous c.d.f. F_1 , with compact support \mathcal{S}_1 , and mean \bar{s} . This skill productivity level is specific to a given manager and evolves over time as follows:

$$s_{t+1} = s_t + \tau_{t+1}, \tag{1}$$

where τ_{t+1} is a mean zero random variable, independently and identically distributed according to the continuous c.d.f. T , with compact support $[\underline{T}, \bar{T}]$, $\underline{T} < 0 < \bar{T}$. The strict inequality ensures that, at any period t , expected period $t + 1$ profits and payoffs are continuous in skill s_t . One can interpret such exogenous improvement/decline of a worker’s ability as a random component of learn-by-doing. From an ex ante perspective, it is straightforward to define the probability distributions F_2 and F_3 over the skills of middleaged and old executives, with supports $\mathcal{S}_2 = [\underline{S} + \underline{T}, \bar{S} + \bar{T}]$ and $\mathcal{S}_3 = [\underline{S} + 2\underline{T}, \bar{S} + 2\bar{T}]$. We assume skills are non-negative, $\underline{S} + 2\underline{T} \geq 0$.

The interpretation of whether skill s is primarily firm-specific (executive-firm match) or general in nature will depend on the correlation of skill s with the executive's wage offer from outside firms, which we describe momentarily. In particular, a higher correlation between s and the outside offer indicates that a higher fraction of s is of a transferable, general-skill nature.

The value of the output of a manager d with characteristics (a^d, s^d) is $l_{a^d} + s^d$, and a CEO with characteristics (a^c, s^c) produces $\rho[l_{a^c} + s^c]$ for $a^c > 1$, and B for $a^c = 1$, where $B < 0$ is assumed to be low enough that it is never optimal for the firm to promote an untested, raw young executive to CEO. The technology parameter $\rho > 1$ captures the importance of the CEO's skill for the firm's output relative to those of managers. We represent the distribution of characteristics of the firm's executives by the executive-profile vector $z = (a^c, s^c, a^1, s^1, a^2, s^2)$. For $a^c > 1$, the total value of period output, measured in dollars, net of costs other than wages is

$$Y(z) = \rho[l_{a^c} + s^c] + [l_{a^1} + s^1] + [l_{a^2} + s^2].$$

We have intentionally designed the period output value function so that it exhibits constant returns to scale and the contribution of one executive is independent of the characteristics of all other executives — from a productive standpoint, there are *no* cross-manager interlinkages. In particular, from a static output maximization standpoint (ignoring the wage-setting process and executives' quit decisions), it would immediately follow that the firm should place the most productive worker as CEO, and that the decision to layoff a manager should be independent of the characteristics of other executives. Nonetheless, we will show that the dynamic aspects of the model cause managerial attributes to become intertwined in complicated ways both in the firm's placement decisions and wage offers, and in the quit decisions of managers.

Our central characterizations extend to more general production technologies as long as (a) the value of period output strictly increases in the skill and learn-by-doing of each executive; (b) output is more sensitive to the CEO's skill and learn-by-doing than the skill and learn-by-doing of a manager; and (c) the functional form of output does not vary across periods. If we drop the additive structure across different executives, other mild assumptions are needed on cross derivatives. We assume that parameters are such that, when we integrate the wage setting process and quit decision of managers, expected profits are positive and finite, so that the firm always wants to operate³.

³Ceteris paribus, expected profits are always positive if the expected skill \bar{s} of young executives is sufficiently high.

2.3 Insiders, Outsiders, and Simplifying Assumptions

Insiders: At the beginning of each period, each of the firm’s three executive positions is either occupied by an executive who worked in the position last period, or it is vacant. The firm must choose which insiders to retain, which insiders to lay off, and which executive positions to fill with outside hires. To each inside executive whom the firm chooses not to layoff, the firm makes a wage offer and a placement offer — either his current position or a promotion to CEO.

After receiving the firm’s placement/wage offer, each insider receives one employment (wage) offer from an outside firm. The realized value of the offer is private information to the executive, but its distribution is common knowledge — we describe this outside wage offer distribution in Section 2.4. If an executive is laid off, he receives and takes the outside offer. If an executive is not laid off, he can quit and take the outside wage offer, or stay at the firm.

Outsiders: Outside executives are divided into two groups: young executives ($a = 1$) and experienced ones ($a \in \{2, 3\}$). Each period, there is a large set of ex ante identical young executives, willing to work at the firm for initial wage W_t^m . One can interpret such young executives as recently-graduated MBA students. A young executive’s skill is not revealed until after he works for one period as a manager. At the end of his first period, a young manager’s skill s is learned by the firm and all of its executives.

In Appendix B we describe a full-fledged version of our model, in which there is also a large set of experienced executives working in other firms. The firm can conduct a costly search to find outside experienced executives who are potential candidates for the firm’s CEO or managerial positions, and there is relevant uncertainty about the outcome of the process. We show how this search model becomes intractable. This leads us to assume that this search cost is so high that it is optimal for the firm to make external hires only at the entry level: outside executives hired for the managerial positions are always young executives, and CEOs are always promoted from within the firm. Consequently, we cannot fully endogenize CEO compensation and CEO turnover. This is because, when promoting an inside manager to CEO, if the firm does not offer a sufficiently high wage, then with positive probability the manager will reject the promotion and leave for a better outside offer. Then the firm could face a situation where it had no inside experienced managers to promote. To ensure that this does not occur, we take the CEO wage and turnover as exogenous. However, we *fully endogenize* both the firm’s placement and wage offers to experienced managers, and managers’ quit decisions.

We then show how our simplified, tractable model maintains the key incentive considerations of the full model. Inside managers must take into account, when evaluating their own probability of promotion to CEO, the probability that the firm might find or not a “good” outside candidate — potential competitor for the CEO promotion. The firm must take into account both (a) the option value embedded in its on-the-job skill assessment of managers, and (b) the value of maintaining a pool of talented inside executives, suitable candidates for the CEO position. These incentives are present in our tractable model via the uncertainty regarding the true skill of young executives.

In practice, most new CEOs are hired internally. Bognanno (2001) studies a database with more than 600 firms for up to 8 years and finds that only 17% of the CEO’s were hired directly from outside firms. Hayes, Oyer and Schaefer (2006) find a similar result — only 17% of their CEO-years observations had less than 5 years of tenure in their firms. Ang and Nagel (2007) find that “CEOs appointed from inside the firm deliver greater cumulative and more persistent performance than those hired from outside the firms.” Although empirically most CEOs are promoted from within, and inside CEOs are on average better than outside CEOs, we believe that fully understanding the search process and trade-offs involved in the choice between insiders and experienced outsiders is important, but simply beyond the scope of this paper.

CEO Wage and Turnover: The CEO wage is exogenously given by the function $W_t^c(a^c, s^c)$, which we allow to depend exogenously on the CEO’s age and skill. We only require the CEO wage to be: (a) high enough that a manager always accepts a CEO promotion, (b) non-decreasing in the CEO’s skill, (c) increasing less quickly in the CEO’s skill than his marginal contribution to output, ρ , and (d) differentiable with respect to skill s^c . In particular, consider a constant $\overline{W}_t^c \geq 0$ and function $w_t^c(a^c, s^c) \geq 0$. Then, for any $s^c \in \mathcal{S}_{a^c}$,

$$\begin{aligned} W_t^c(a^c, s^c) &= \overline{W}_t^c + w_t^c(a^c, s^c), \\ \rho &> \frac{\partial w_t^c(a^c, s^c)}{\partial s^c} \geq 0. \end{aligned} \tag{2}$$

Given the wage $W_t^c(a^c, s^c)$, an old manager always accepts the promotion, works as CEO and retires the next period. A middleaged manager also accepts promotion to CEO and works for at least one period. At the end of the period, the middleaged CEO leaves the firm⁴ with exogenous probability $\pi_t \in [0, 1)$.

⁴If a middleaged CEO leaves the firm at the end of period t , the value of his period $t + 1$ expected outside wage offer is irrelevant for our qualitative results, as long as it is sufficiently high — so that it is always optimal to accept the CEO promotion when middleaged. Hence, to simplify notation, we assume that this expected outside wage offer is given by the CEO wage function $W_{t+1}^c(3, s^c)$.

After we present Theorem 2, we discuss the extent to which our results extend when the CEO wage and turnover can vary exogenously with the CEO's attributes *and* the attributes of other inside managers.

2.4 Personnel Management Technology

At the beginning of each period t , the firm observes which executive positions are vacant and the characteristics of those executives from the previous period who remain in the firm. The executive-profile vector $z_t = (a_t^c, s_t^c, a_t^1, s_t^1, a_t^2, s_t^2)$ summarizes this information. We sometimes refer to z_t as the state of the firm. We define the set of possible executive-profiles \mathcal{Z} , or state space, in the appendix. The initial state z_1 is irrelevant for our qualitative results, as long as $z_1 \in \mathcal{Z}$.

After observing the state z_t , the firm announces a feasible period **executive employment offer** e_t . The vector e_t has seven dimensions and can be divided into two components: a feasible period **executive placement offer** p_t and a feasible period **executive wage offer** w_t .

An executive placement offer is a four-dimensional vector $p_t = (\Psi_t^1, \Psi_t^2, \Gamma_t^1, \Gamma_t^2) \in \{0, 1\}^4$. Here, $\Psi_t^d = 1$ if the firm offers the manager of division $d \in \{1, 2\}$ the CEO promotion, and is zero otherwise; and $\Gamma_t^d = 1$ if the firm offers manager d the opportunity to stay as manager of that division, and is zero otherwise.

To be feasible, a placement offer must satisfy four technological conditions: (i) if the CEO position is vacant, then one existing manager must be promoted to CEO; (ii) if the CEO position is not vacant, then no manager can be promoted; (iii) promotion to CEO cannot occur from a vacant division; and (iv) if the manager from division d is promoted or if that division is vacant, then it must be that $\Gamma_t^d = 0$. For each realized state $z_t \in \mathcal{Z}$, $t < N$, the set of feasible placement offers is

$$\mathcal{P}(z_t) = \left\{ p_t = (\Psi_t^1, \Psi_t^2, \Gamma_t^1, \Gamma_t^2) \in \{0, 1\}^4 \text{ s.t. } \begin{array}{ll} (i) & \Psi_t^1 + \Psi_t^2 = 1 \quad \text{if } a_t^c = 0 \\ (ii) & \Psi_t^1 = \Psi_t^2 = 0 \quad \text{if } a_t^c = 3 \\ (iii) & \Psi_t^d = 0 \quad \text{if } a_t^d = 0 \\ (iv) & \Gamma_t^d = 0 \quad \text{if } \Psi_t^d = 1 \text{ or } a_t^d = 0 \end{array} \right\}.$$

The firm closes at period $t = N$ and lays off all remaining executives. Therefore, the unique placement offer at $t = N$ is $z_N = (0, 0, 0, 0)$.

An executive wage offer is a three-dimensional vector of wage offers to the CEO and managers of divisions 1 and 2, $w_t = (W_t^{CEO}, W_t^1, W_t^2)$. We consider very general settings for wage offers to experienced managers. For example, we will consider the possibility that the firm faces no constraints on wage offers. With this perfectly flexible wage structure, the endogenous wage the firm offers an experienced manager can vary in arbitrary ways not only with the manager's attributes,

but also the attributes of *all* other executives. But it may also be that the firm faces constraints on the wages it can offer experienced managers. For example, workers with the same job may have to receive the same wage, or, due to equity concerns, wages cannot differ by too much across workers, or seniority rules may demand that older workers be paid at least as much as younger workers. To capture such possibilities, we introduce a general non-empty/closed wage constraint set $\mathcal{H}_t^d(z_t, p_t, w_t^{-d})$ that limits the possible wages that the firm can offer an experienced manager of division d . The wage constraint set can be conditional on the executive profile and the firm's placement offer, and a continuous correspondence of the firm's wage offer w_t^{-d} to other executives. To highlight the potential impact of such constraints, we focus primarily on contrasting equilibrium outcomes when the firm is completely unconstrained, $\mathcal{H}_t = \mathbb{R}$, with those that obtain when the firm has no discretion on wage offers to experienced managers at date t , and must offer them the common wage W_t^m . Contrasting equilibrium outcomes in these two extreme scenarios reveals the extent to which the firm's use of placement offers to provide incentives to executives varies with its discretion in wage offers.

At period t , given the executive-profile vector $z_t = (a^c, s^c, a^1, s^1, a^2, s^2)$, the set of feasible period executive employment offers is

$$\mathcal{E}_t(z_t) = \left\{ \begin{array}{l} e_t = (\Psi_t^1, \Psi_t^2, \Gamma_t^1, \Gamma_t^2, W_t^{CEO}, W_t^1, W_t^2) | \\ s.t. \quad (i) \quad (\Psi_t^1, \Psi_t^2, \Gamma_t^1, \Gamma_t^2) \in \mathcal{P}(z_t) \\ \quad (ii) \quad W_t^d = W_t^m \\ \quad (iii) \quad W_t^d \in \mathcal{H}_t \\ \quad (iv) \quad W_t^{CEO} = W_t^c(a^c, s^c) \\ \quad (v) \quad W_t^{CEO} = W_t^c(a^d, s^d) \end{array} \quad \left. \begin{array}{l} \\ \\ \\ if \quad \Gamma_t^d = 0 \\ if \quad \Gamma_t^d = 1 \\ if \quad \Psi_t^1 = \Psi_t^2 = 0 \\ if \quad \Psi_t^d = 1 \end{array} \right\}.$$

First, the placement offer must be feasible. Second, when managerial position d is vacant, the offer to the *young* manager is W_t^m . Third, wage offers to experienced managers must be in \mathcal{H}_t . Fourth and fifth, the CEO wage offer must be consistent with the CEO wage function, both when there is a CEO from last period and when a new CEO is promoted at period t .

After the firm announces an executive employment offer e_t , each manager $d \in \{1, 2\}$ who is not promoted to CEO receives an employment/wage offer from an outside firm, q_t^d , drawn from the c.d.f. $G_t^d(q_t^d | z_t, e_t)$, with compact support $\mathcal{Q} = [0, \hat{q}]$. This outside wage offer equals the executive's expected discounted compensation, were he to quit, over the remainder of his productive life. The realized value of the offer is private information to the manager, but its distribution is common knowledge. We allow the possibility that the distribution over outside wage offer *might* be a function

of (a) manager’s own characteristics⁵, (b) characteristics of other executives⁶, and the (c) firm’s actions⁷. We assume $G_t^d(\cdot)$ is differentiable with respect to q_t^d and the firm’s wage offer.

If a manager is laid off, he receives and takes the outside offer. If a manager is not laid off, he can quit and take the outside wage offer, or stay at the firm. We assume that a manager always stays in the zero probability event that he is indifferent between staying and quitting. After managers have accepted or rejected the firm’s offers, the firm and managers cannot change their decisions.

Each vacant managerial position is then filled by a young outside executive. Nature draws an independent random realization of managerial skill s for each new manager according to F_1 . Then production takes place, profits are realized and wages are paid. If an executive has experience l_3 , he then retires and the position becomes vacant. If the CEO is middleaged, he leaves the firm with exogenous probability π_t . Period t then ends, agents get older and draw an independent random skill component τ according to T ; period $t + 1$ begins. The sequence of events and decisions is:

1. Period starts. Firm chooses managers to promote and lay off, announces wage offers.
2. Managers not promoted receive wage offer from outside firm.
3. Managers simultaneously choose to quit/stay.
4. New managers are hired for the vacant positions, and draw their skills.
5. Output is produced and wages paid.
6. Old executives retire; if the CEO is middleaged, he leaves the firm with exogenous probability π_t .
7. Remaining executives get older and random skill component τ of each executive is realized.

2.5 Equilibrium

Our three period-lived OLG model generates a rich variety of management structures (age-profiles of executives) that allow us to determine how different management structure realizations affect the firm’s placement/wage offers and managers’ quit decisions. Figure 1 depicts five of the eight possible representative partitions detailed in Appendix A. Example (a) shows an age-profile of executives that is comparable to the standard models in the literature in two dimensions: (i) all workers in the same hierarchical level have the same age, and (ii) all workers in the next hierarchical level are one period older. Examples (b) and (c) illustrate one dimension in which our model features

⁵If part of the skill is general, transferable to the outside firm, then the distribution of the external wage offer should be positively correlated with the manager’s skill.

⁶Consider an outside firm that cannot observe the exact skill level of a manager. If (a) the outside firm can observe the firm’s total output, but not individual contributions; or (b) the outside firm can only compare managers and deduce who is better, then the skill of manager d is relevant for the outside wage offer distribution of manager $-d$.

⁷For example, the outside wage distribution might be lower if the manager is laid off. The firm’s wage offer might also convey information to the outside firm, influencing the outside wage offer.

economically interesting possibilities missing in existing models, where a middleaged manager works with an older or younger manager. Examples (d) and (e) illustrate the economically-relevant scenario where the CEO can be younger than a manager or be the same age.

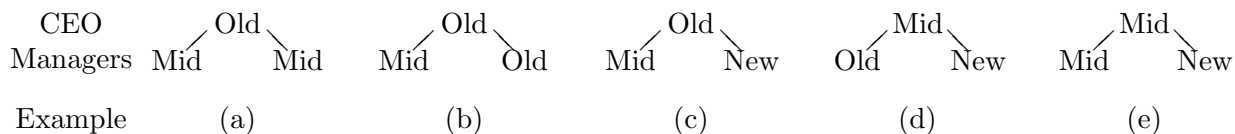


Figure 1: Examples of management structures (age-profile of executives).

As this is a dynamic game, the firm and each manager must form expectations about the actions all other agents take not only in the current period, but also their likely actions in future periods. The firm-wide distribution of ages and skills enters incentives and strategies non-trivially, as they influence the probability distribution over future actions and states. Our model is the most intricate and comprehensive model of management dynamics in the literature. Our framework allows us to address: (a) how the firm's placement/wage offer to experienced managers depend on the attributes of the other executives; (b) how the firm trades off between age and ability in its CEO promotion decisions; and (c) how the firm trades off between wages and placement offers to align managerial incentives, for different managerial age profiles and skill distributions.

Strategies and Beliefs: At the beginning of period t , the firm observes the executive-profile vector z_t and then announces a feasible period employment offer $e_t \in \mathcal{E}_t(z_t)$. The firm knows the distribution of outside offers $G_t^d(\cdot)$, has beliefs about the quit strategies of managers, and can calculate the probability $Pr_t^d(z_t, e_t)$ that manager d will stay given the state z_t and offer e_t . The expected profit function along the equilibrium path is given in Appendix A.

Let $\Delta\mathcal{E}_t(z_t)$ be the set of probability distributions over feasible period executive employment offers $\mathcal{E}_t(z_t)$. A period executive employment strategy for the firm is a function $\sigma_t : \mathcal{Z} \rightarrow \Delta\mathcal{E}_t(z_t)$ that defines for each state $z_t \in \mathcal{Z}$ a probability distribution $\sigma_t(z_t)$ over the set of feasible executive employment offers $\mathcal{E}_t(z_t)$. An executive employment strategy for the firm is a collection σ of period strategies, $\{\sigma_1, \dots, \sigma_N\}$. The firm chooses σ to maximize expected discounted lifetime profits.

Each executive maximizes discounted expected lifetime wages. A manager working at division $d \in \{1, 2\}$ at the beginning of period t observes the state z_t , the firm's placement offer e_t and his own realized outside wage offer q_t^d . Each existing manager who was neither promoted to CEO nor laid off must decide whether to accept the job offer and stay at the firm, or to quit and take the outside option q_t^d . Managers make their stay/quit decisions simultaneously. Manager d has beliefs

about the quit strategy of the other manager and the firm's employment strategy in the next period, and can calculate next period's expected payoff $\mu_{t+1}^d(z_t, e_t)$ if he stays at the firm in the current period. For an old manager, this value is zero since he will retire. Manager d 's period strategy is a function $\varphi_t^d : \mathcal{Z}^m \rightarrow \{0, 1\}$, $\mathcal{Z}^m = \mathcal{Z} \times \mathcal{E}_t(z_t) \times \mathcal{Q}$, where $\varphi_t^d(z_t, e_t, q_t^d) = 1$ if the manager decides to stay at the firm and $\varphi_t^d(z_t, e_t, q_t^d) = 0$ if he quits. By convention, we set $\varphi_t^d(z_t, e_t, q_t^d) = 0$ when the manager d is fired, promoted to CEO, or the position is vacant. In particular, since the firm never keeps managers at the terminal period $t = N$, $\varphi_N^d(z_N, e_N, q_N^d) = 0$. Let φ^d be the collection of period strategies $\{\varphi_1^d, \dots, \varphi_N^d\}$ of managers of division d . Each manager of division d chooses a period strategy $\varphi_t^d(\cdot)$ that maximizes his discounted expected lifetime wages.

Definition: A Perfect Bayesian Equilibrium (PBE) of the model is a collection of beliefs and strategies σ^* , φ^{1*} and φ^{2*} , such that:

- i) Given strategies φ^{1*} and φ^{2*} and the firm's beliefs, the strategy σ^* solves the firm's problem;
- ii) Given strategies σ^* and φ^{-d*} and division d managers' beliefs, the strategy φ^{d*} solves the problem of division d managers;
- iii) Beliefs of the firm and managers are consistent, wherever possible, with equilibrium strategies.

Theorem 1 Consider any feasible collection of firm technology parameters and sequence of outside wage offer distributions. If either

- (i) The set of feasible wage offers $\mathcal{H}_t^d(\cdot)$ is finite, or
- (ii) $\frac{\partial G_t^d(q_t^d | z_t, p_t, w_t)}{\partial q_t^d}$ and $\frac{\partial G_t^d(q_t^d | z_t, p_t, w_t)}{\partial w_t}$ are sufficiently small,

then there exists a Perfect Bayesian Equilibrium.

We solve the model recursively to prove existence of a PBE⁸. In the last period $t = N$, the firm closes and lays off all executives; hence, computing expected period payoffs is trivial. At period $t = N - 1$, after observing the firm's employment offer, each manager chooses whether to stay or quit, generating equilibrium quit probabilities. At the beginning of period $N - 1$, the firm forms consistent beliefs regarding managers' quit probabilities for each feasible employment offer, and chooses the expected profit-maximizing offer. If (i) the set of feasible wage offers is finite, then the

⁸We assumed that the CEO wage function and the firm's output technology are such that it is always optimal for an experienced manager to accept the CEO promotion, and that it is never optimal for the firm to promote a young, unexperienced executive to the CEO position. To keep the paper's length manageable, we do not characterize the off equilibrium path in which the firm promotes a young executive to CEO, as this would greatly increase the size of the state space. All other off-equilibrium paths are characterized.

set of feasible employment offers is also finite and an optimal offer always exists. Alternatively, we prove that an optimal employment offer also exists if condition (ii) is satisfied. Condition (ii) is a gross sufficient condition for managers' quit probabilities and the firm's expected profits to be continuous functions of the firm's wage offer. Therefore, we can compute expected payoffs for period $t = N - 1$ and solve for optimal strategies at period $t = N - 2$, and so on.

The strategic interaction between managers and firm could conceivably give rise to multiple equilibria. Also it is not obvious that the firm's optimal strategy must treat the two divisions symmetrically. Indeed, we will show numerically that if for some exogenous reason the firm (or managers) must behave asymmetrically at some period $t < N$ (e.g., promote the manager of division 1 at date t), then optimal strategies for all periods $t' < t$ are also asymmetric. We now investigate when equilibria *will* be unique and appropriately symmetric and monotone.

Definition: A PBE is **monotone** and features **pure**-strategies if:

1. **[Firm's Placement Strategy]** The firm's placement offer takes the form of cutoff functions on the manager's own skill.

That is, given the age of experienced manager d and the characteristics of *all* other current executives, the firm defines a promotion cutoff, $\gamma_t^c(\cdot)$, and a layoff cutoff, $\gamma_t^m(\cdot)$. When the CEO position is vacant, the manager of division d is promoted to CEO if his skill exceeds the promotion cutoff, i.e., $s^d > \gamma_t^c(\cdot)$; otherwise, he is offered the managerial position if his skill exceeds the layoff cutoff, i.e., $s^d > \gamma_t^m(\cdot)$, and he is laid off if his skill falls below the layoff cutoff, i.e., $s^d < \gamma_t^m(\cdot)$.

2. **[Firm's Wage Strategy]** The firm's equilibrium wage offer provides an experienced managers d an expected payoff from staying in the firm that increases weakly in his skill s^d .
3. **[Firm's Profits]** In equilibrium, the firm's expected profits increase weakly in each of manager's skill.
4. **[Managers' Quit Strategies]** A manager's equilibrium quit strategy defines a cutoff on the outside wage offer; that is, a manager quits and accepts the outside offer if and only if the offer exceeds the cutoff.
5. **[Managers' Payoff]** In equilibrium, the expected lifetime payoff of an experienced manager increases weakly in his own skill.

We say that a monotone equilibrium is unique if the equilibrium strategies define unique cutoffs and wage offers for each period t . An equilibrium is symmetric/anonymous if (i) for any state $z_t \in \mathcal{Z}$, where it is a best response to promote (lay off) the manager of division d , then it would also be a best response to promote (lay off) the manager of division $-d$ in state $z'_t \in \mathcal{Z}$ that is equal to state z_t except that the division labels of the managers have been switched, and (ii) the optimal wage offer, the equilibrium strategies of managers, expected present value of profits, and expected lifetime payoff of a manager in equilibrium are symmetric/anonymous in the same sense.

Establishing uniqueness simplifies comparative statics analysis. Establishing monotonicity eases the analytical characterization of optimal strategies — in a non-monotonic setting, the firm’s promotion strategy defines multiple cutoffs on a manager’s own skill, rather than one. Because we believe that monotonicity is empirically relevant, we focus our subsequent analytical characterization of equilibrium outcomes on this class of distributions. However, we also numerically characterize economies where monotonicity does not hold.

Some structure on the distribution of outside wage offers is required for monotonicity to hold. For example, if the outside wage offers that an old manager receives are too sensitive to his skill (e.g., rise more than one-for-one with skill), then profits will be non-monotone. Less intuitively, monotonicity may fail if the outside wage offer increases *too slowly* with skill. To see why, suppose that the outside offer of a middleaged manager *fell* with his skill. The firm might then find it profitable to delay promotion to CEO, resulting in a non-monotone placement strategy. More generally, consider a change in the state of the firm and its employment offer from (z_t, e_t) to (z'_t, e'_t) . If, for manager d , the change in his expected outside wage offer diverges by too much from the change in his expected payoff from staying in the firm, then monotonicity might fail.

By exploiting the recursive nature of the model, one can solve the model recursively and construct a non-empty class of outside wage offer distribution, possibly correlated to managerial attributes and firm actions, such that the PBE exhibits these properties: the PBE is unique, symmetric, monotone and features pure-strategies. For each partition of the state space at each period t , given the next period’s expected payoff of managers and the firm’s expected discounted profits, one can define upper and lower bounds on the sensitivity of the outside wage offer distribution $G_t^d(\cdot)$ to the attributes of executives and the actions of the firm, for the reasons explained earlier.

In this broad class of distributions, the degree of correlation between a manager’s outside wage offer distribution, the attributes of other executives and the firm’s employment offer can vary significantly. We are particularly interested in determining how a firm’s placement and wage offer

strategies and a manager's quit decisions vary when the attributes of other executives are changed. When the outside wage offer is correlated with these attributes and firm's endogenous optimal employment offer, it is difficult to unravel which results are driven by the endogenous dynamics of the model, and which results are driven by the specific correlation structure of the outside wage offer distribution. Accordingly, this leads us to consider a setting where the distributions of outside wage offers are possibly correlated with manager's own skill, but are independent of the characteristics of other managers and the firm's employment offer.

Old Managers: The outside wage offer to an old manager with skill s is given by the random variable

$$q_t^{old} = \lambda s + \xi_t^{old}, \quad (3)$$

where $\lambda \in [0, 1)$ and the random variable ξ_t^{old} is uniformly distributed on $[0, \hat{q}^{old}]$. The parameter λ captures the proportion of skill that is general (transferable to the outside firm) in nature⁹, and the realized value of ξ_t^{old} captures the realized match of the worker with an outside firm.

Middleaged Managers: The outside wage offer to a middleaged manager with skill s is given by the random variable

$$q_t^{mid} = (1 + \beta)(\lambda s + \xi_t^{mid}), \quad (4)$$

where the random variable ξ_t^{mid} is uniformly distributed on $[0, \hat{q}^{mid}]$. The term $(\lambda s + \xi_t^{mid})$ represents the per period payoff, and $(1 + \beta)$ captures the fact that the middleaged manager still has two productive periods left in his work horizon.

Theorem 2 *Consider a perfectly flexible wage structure, any feasible collection of firm technology parameters, and let the distribution of outside wage offers be given by equations (3) and (4). If*

C1. *There are sufficiently attractive outside wage offers, i.e., the supports of ξ_t^{mid} and ξ_t^{old} are sufficiently large; and*

C2. *The CEO position is sufficiently important, i.e., ρ , \bar{W}_t^c and $\rho - \frac{\partial w_t^c(a^c, s^c)}{\partial s^c}$ are sufficiently large,*

*then there is a **unique** PBE. Further, this equilibrium is **symmetric**, **monotone** and features **pure-strategies**.*

⁹The qualitative results from Theorem 2 and Proposition 1, which we present momentarily, extend to the more general case where the expected outside wage offer is a non-linear function $\Lambda(s)$ of a manager's skill s , as long as $\Lambda'(s) \in [0, 1)$. We focus on the linear case to simplify presentation.

Condition C1 serves three roles. First, large upper bounds $\{\hat{q}^{old}, \hat{q}^{mid}\}$ on the outside wage offers ensure that the firm’s optimal wage offers are interior — that is, a manager who is not laid off stays in the firm with probability less than one: circumventing consideration of boundary possibilities eases presentation. Second, when the firm wants to keep two managers, condition C1 bounds the externalities caused by wage offers. This bound guarantees that the optimal wage offers are unique and yield an expected payoff that is increasing in a manager’s own skill.

Third, C1 is necessary to restrain the firm’s incentives to conceal talented executives in the managerial position, and ensure promotion monotonicity. To understand the firm’s incentives to conceal talented executives, recall that we have not bounded the probability π_t of CEO turnover. If the exogenous probability of CEO turnover is too high relative to the endogenous probability that a manager chooses to quit, then the firm will maximize profits by concealing talented executives in the managerial position, violating CEO promotion monotonicity¹⁰.

Moreover, we need to ensure that the firm’s expected profit from promoting a manager to CEO increases in the executive’s skill at least as fast as the firm’s expected profit from keeping the executive as a manager. Equation (2) only states that $\rho - \frac{\partial w_t^c(a^c, s^c)}{\partial s^c} > 0$; that is, the firm’s per period profit from promoting a manager to CEO increases in the executive’s skill, but this increase might be small. *Conditional* on the executive staying as a manager, the firm’s discounted profit may increase in his skill faster than $\rho - \frac{\partial w_t^c(a^c, s^c)}{\partial s^c}$. Therefore, condition C1 bounds the increase in the firm’s *expected* discounted profit, by ensuring that the manager is sufficiently likely to leave the firm if not promoted, while condition C2 ensures that $\rho - \frac{\partial w_t^c(a^c, s^c)}{\partial s^c}$ is sufficiently high, i.e., the firm’s expected profit increases in the CEO’s skill sufficiently fast. Notice that $\rho - \frac{\partial w_t^c(a^c, s^c)}{\partial s^c}$ can be small if \hat{q}^{mid} and \hat{q}^{old} are high, and vice versa.

Finally, a sufficiently high \bar{W}_t^c ensures that it is always optimal for managers to accept the CEO promotion even when \hat{q}^{old} and \hat{q}^{mid} are high, and a sufficiently high ρ ensures that the firm’s expected profits are positive. Numerically we find that monotonicity is violated only when $\{\hat{q}^{old}, \hat{q}^{mid}\}$ and $\rho - \frac{\partial w_t^c(a^c, s^c)}{\partial s^c}$ are small, and CEO turnover is implausibly high.

Robustness: One can show that the results in Theorem 2 extend to the rigid wage setting, where all managers receive the same exogenously given wage W_t^m , *provided* that the level of transferable skill λ is sufficiently small. When the firm’s wage offers are perfectly rigid, λ must be bounded; otherwise, a manager’s likelihood of quitting might increase too quickly with his skill — as his

¹⁰The key is to ensure that an executive is more likely to leave the firm if he is not promoted to CEO. Empirically, Hayes, Oyer and Schaefer (2006) find that the CEO turnover probability (10.3%) is less than the turnover probability of other top firm’s executives (13.8%).

expected outside wage offer increases with his skill at rate λ but his inside wage offer is constant — reducing the firm’s expected profits, so that payoff monotonicity does not hold. In the next section we numerically solve economies with flexible and rigid wage structures.

Theorem 2 also holds even if the CEO’s wage varies in a limited way with the characteristics of inside managers. The bounds are as follows. If a manager’s skill s^d is below the layoff cutoff, then the CEO wage must be non-increasing in s^d ; otherwise, the firm’s profits fall with the skill of a manager who will be laid off. For similar reasons, if s^d exceeds the layoff cutoff, then the rate at which the CEO wage increases in s^d must be bounded from above, where the strictly positive upper bound increases with s^d . For partitions where the firm chooses a manager to promote to CEO, the rate at which the CEO wage decreases in the skill of the manager not promoted must be bounded from below; otherwise, when the firm has two talented managers of the same age, it might be optimal to promote the slightly less talented manager. Lastly, the difference between the rates at which the CEO’s wage increases with the CEO’s skill when division d is vacant and the rate at which the CEO’s wage rises with the CEO’s skill when division d has a manager must be bounded appropriately, otherwise profits might be decreasing in the CEO’s skill.

Theorem 2 also holds if we allow the likelihood of CEO turnover to depend exogenously on the CEO’s skill in a limited way. We require $\pi_t(s^c)$ to be differentiable and identify a sufficiently small upper bound $\bar{\pi} > 0$ such that, for any $s^c \in \mathcal{S}_{a^c}$,

$$\left| \frac{\partial \pi_t(s^c)}{\partial s^c} \right| \leq \bar{\pi}. \quad (5)$$

In practice, the impact of CEO skill on CEO turnover is not a priori obvious. For example, the probability of turnover may rise with CEO skill since other firms would be more willing to hire the CEO away; but a valid claim could be made in the opposite direction — the probability may be decreasing in CEO skill since the firm would compete more aggressively to keep the executive. More generally, CEO turnover may be higher for low and high values of CEO skill, and lower for intermediate skill values. Similar to the CEO wage, we can also allow the CEO turnover probability to vary in a limited way with the attributes of other inside executives.

3 Analysis

3.1 Wage offers

In the proof of Theorem 2, we solve for the firm’s optimal wage offer to managers who are not laid off. We define the *total surplus* created when the manager stays in the firm as the sum of the firm’s

change in profit plus the manager's change in discounted expected lifetime payoff, and find that

Corollary 1 *When the firm makes a wage offer to a single manager, the firm's optimal wage offer gives the manager an expected lifetime payoff from staying in the firm equal to half of the total surplus plus the deterministic component of the outside wage offer¹¹.*

One can view this framework as a single buyer (the firm) that makes a take-it-or-leave-it offer (the wage offer) to a single seller (the executive), who has a privately-known opportunity cost (the outside wage offer) of "selling" his work. The optimal wage offer is then a surplus sharing agreement, and the uniform distribution allows us to compute the sharing rule explicitly. When the firm makes wage offers to two managers, the offer described in Corollary 1 is adjusted by the equilibrium probability that the other manager stays, the change in total surplus conditional on both managers staying, and the externalities in managers' expected payoffs derived from the wage offers.

The fact that the firm's optimal wage offer takes into account the manager's own contribution to output and the distribution of his outside wage offer is not surprising. However, the results that the current period's wage offer also takes into account (i) the manager's expected next period payoff from staying in the firm, and (ii) the characteristics of the other executives in the firm deserves further analysis. Recall that we intentionally designed the period output value function so that it exhibits CRTS and the contribution of one executive is independent of the characteristics of all other executives. Therefore, the connection between the firm's wage offer to one executive with the characteristics of other executives emerges endogenously from the dynamic aspects of the model. The connection between current wage offer and a manager's future expected payoff also emerges endogenously. The firm knows that a manager's decision whether to stay at the firm or quit is a function of the present value of his payoff from staying in the firm, which includes not only the current wage offer, but also expected future payoff, which is a function of future probabilities of promotion and layoff, as well as expected inside and outside wage offers. In the following section we further characterize the firm's placement and wage offers.

3.2 Promotions, quits and layoffs

In this section we show how the attributes of the managerial workforce affect the firm's placement/wage offers and managers' quit decisions, and how the technology parameters affect equilibrium outcomes. To focus on the central economics, we consider an outside wage distribution given by equations (3) and (4). A manager's outside wage distribution is orthogonal to the attributes of

¹¹ λs for old managers and $(1 + \beta)\lambda s$ for middleaged managers.

the other executives; therefore changes in optimal strategies across different attribute profiles are endogenously driven by the dynamics of the model.

Proposition 1 below summarizes five *analytical* results that characterize central features of equilibrium outcomes when the CEO position is important, that is, when the CEO’s wage and relative contribution to firm’s output are high. The results describe how the probability that each manager quits varies with the attributes of *all* executives and the probability of CEO promotion. They also characterize how the firm combines placement and wage offers to provide incentives to executives — by exploiting the effect of offers on quit-decisions of all managers — and how the firm’s use of these incentives varies with the attribute profiles of executives. The firm uses these tools to search for future CEOs, to nurture talented executives, and to optimally adjust the level of competition for the CEO position.

Proposition 1 *Consider a perfectly flexible wage structure, any feasible collection of firm technology parameters, and let the distribution of outside wage offers be given by equations (3) and (4). If*

C1’. There are sufficiently attractive outside wage offers, i.e., the supports of ξ_t^{mid} and ξ_t^{old} are sufficiently large; and

C2’. The CEO position is sufficiently important, i.e., ρ , \bar{W}_t^c and $\rho - \frac{\partial w_t^c(a^c, s^c)}{\partial s^c}$ are sufficiently large,

then the following five results hold¹² for any period $t < (N - 1)$:

R1. *Consider an old manager with skill s^{old} and a talented middleaged manager with skill s^{mid} , who has a sufficiently high chance to win internal promotion to CEO when facing a young manager with unknown talent. The firm promotes the old manager to CEO unless s^{mid} is sufficiently greater than s^{old} , even in the absence of learn-by-doing and when middleaged managers tend to receive better outside wage offers, $\hat{q}^{old} < \hat{q}^{mid}$.*

R2. *A talented middleaged manager, who has a sufficiently high chance to win internal promotion to CEO when facing a young manager with unknown talent, is more likely to leave the firm if the CEO is middleaged than if the CEO is old.*

R3. *An old manager close to retirement, who no longer has a chance of internal promotion to*

¹²The tightness of the bounds in Theorem 2 and Proposition 1 may differ. For example, Proposition 1 holds even if the monotonicity of the firm’s strategy with respect to middleaged managers is violated, as illustrated in the example at the end of this section. Obviously, both results hold when we tighten the bounds.

CEO when facing a young manager with unknown talent, is more likely to leave the firm if the CEO is old than if the CEO is middleaged.

R4. Consider an old CEO and fix any wage offer to a middleaged manager. The middleaged manager is more likely to stay if the firm retains another manager who is either (i) older and closer to retirement, or (ii) middleaged and sufficiently less talented.

When wages are perfectly flexible, a middleaged manager is more likely to stay if the firm keeps the old manager than if the firm lays the old manager off.

R5. The firm's optimal placement/wage offers to a manager are complex, potentially non-monotonic functions of the attributes of all other executives.

To illustrate and understand the incentives underlying these fundamental results, we numerically characterize equilibrium outcomes in the stationary, long-run economy, for the following technology parameters:

- A young executive keeps his skill for his entire career; that is, $\underline{T} = \bar{T} = 0$ ¹³.
- Uniform distribution F_1 of a young executive skill, with support $[0, \bar{S}]$.
- A CEO with skill s receives wage $\bar{W}^c + \lambda s$. \bar{W}^c is large enough that it is always optimal for a manager to accept promotion.
- Young managers receive exogenous wage W^m .
- With perfectly rigid wages, all experienced managers also receive W^m .
- With perfectly flexible wages, the firm optimally chooses the wages of experienced managers.
- The probability of middleaged CEO turnover is a constant, $\pi \in [0, 1)$.

The related figures are equilibrium strategies at period $t = 1$ when the firm operates for a large number N of periods. Our numerical analysis shows that as N becomes large, the equilibrium strategy at the initial period $t = 1$ always converges. Moreover, if any exogenous asymmetry is introduced at some period t (e.g., the firm has to promote the manager of division 2 at that date), then optimal period strategies in periods $t' < t$ are also asymmetric. However, the bias in the optimal strategies falls as $t - t'$ increases, and strategies converge to the case without exogenous asymmetry¹⁴.

¹³The assumption $\underline{T} < 0 < \bar{T}$ ensures that expected profits and payoffs are continuous in skill s_t , and eases the proof of Theorem 2. The assumption is not necessary for the numerical characterization.

¹⁴For example, suppose that at period t , for some exogenous reason, the firm always promotes the manager from division 2 when the CEO position is vacant. Then at period $t - 1$, a middleaged manager from division 2 has a strong incentive to stay at the firm if not promoted to CEO: he will be promoted in the next period. Conversely, a

R1. *The firm has a bias toward promoting old executives to CEO.*

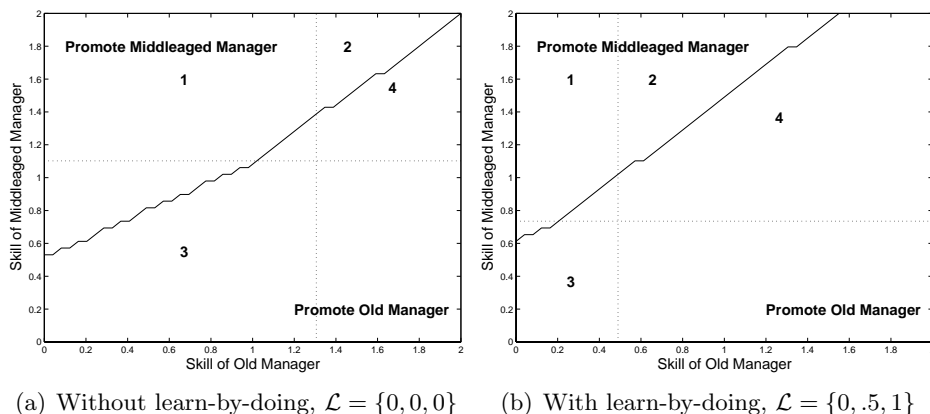


Figure 2: Optimal CEO promotion when managers have different age and wages are rigid. Parameters: $\rho = 3$, $\bar{S} = 2$, $\hat{q}^{mid} = \hat{q}^{old} = 3$, $\bar{W}^c = 3$, $W^m = 1.5$, $\pi = .1$, $\beta = .8$, $\lambda = 0$.

Figures 2(a) and 2(b) illustrate the firm’s promotion strategy bias — the difference between the promotion cutoff rule and the forty-five degree line. When the middleaged manager is sufficiently more able, he is promoted to CEO; the firm fires low skilled old managers (region 1) and keeps high skilled old managers (region 2). When the middleaged manager is not sufficiently more able, the old manager is promoted to CEO; the firm fires low skilled middleaged managers (region 3) and keeps high skilled middleaged managers (region 4).

When both managers are high-skilled, the bias is mainly driven by the differences in managers’ expected payoffs from staying in the firm. A talented old manager who is not promoted to CEO lost the tournament and will never receive the CEO wage prize. As a result, he is especially likely to leave the firm. In contrast, a talented middleaged manager who is not promoted to CEO is likely to be promoted once the old CEO retires, and hence is likely to stay. When both managers are low-skilled, the firm prefers to promote the untalented old executive who will retire next period than the untalented middleaged executive, who would then likely drag down the firm as CEO for longer. This result can reconcile the empirical finding that, controlling for tenure and hierarchical level, the promotion probability of top executives increases with age.

R2. *A talented middleaged manager is less likely to leave the firm if the CEO is old.*

Figures 3(a) and 3(b) show the equilibrium probability that the middleaged manager stays

middlesaged manager from division 1 has a strong incentive to leave the firm: he will not be promoted in the next period. The firm takes this into account, and at period $t - 1$ optimally introduces a bias to promote able managers from division 1. However, this optimal asymmetry at period $t - 1$ is smaller than the exogenous asymmetry from period t : the firm will not promote every manager of division 1 at period $t - 1$. A similar argument applies to period $t - 2$ and so on, in a manner that the asymmetry switches from period to period, but becomes smaller each time.

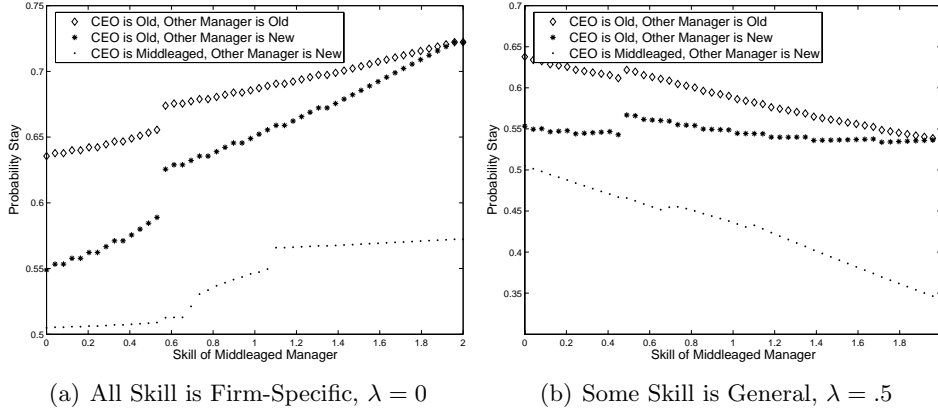


Figure 3: Quitting Strategies of a middleaged manager when wages are perfectly rigid. Parameters: $\rho = 3$, $\bar{S} = 2$, $\hat{q}^{mid} = \hat{q}^{old} = 3$, $\bar{W}^c = 3$, $W^m = 1.5$, $\mathcal{L} = \{0, 1, 1\}$, $\pi = .1$, $\beta = .8$.

under different age-profile of executives, when wages are perfectly rigid. In each figure, the bottom line shows the equilibrium probability that the middleaged manager stays when the manager of the other division is young and the CEO is middleaged. The middle line shows the higher probability that the middleaged manager stays when the CEO is old. An old CEO will retire soon, and the prospect of imminent promotion to CEO increases the incentives of a talented middleaged manager to stay. Moreover, a firm with an older CEO is more concerned with retaining talented middleaged managers who can become the next CEO. This equilibrium result can reconcile the empirical finding that turnover is higher when the CEO is younger.

The figures also show that the difference between the two lines increases with the middleaged manager's skill: both the manager's probability of promotion and firm's incentive to keep him increase faster with the manager's skill when the CEO is old, closer to retirement. The jumps in the probability that a middleaged manager stays reflect the discreteness of placement offers. There is a discontinuous increase in the middleaged manager's expected payoff from staying in the firm when his skill approaches the cutoff for non-layoff when old. When wages are fully flexible, the firm optimally adjusts its wage offer to smooth incentives, as Figure 6(a) illustrates.

Comparing Figures 3(a) and 3(b) illustrates how an increase in the correlation between the outside wage offer and manager's skill can affect equilibrium quit probabilities. In Figure 3(a), all skill development is firm specific, i.e., the outside wage offer distribution is orthogonal to manager's skill. In this case, monotonicity in the manager's expected payoff from staying in the firm implies that the probability a middleaged manager stays increases in his own skill. Figure 3(b) illustrates that, when wages are rigid and the outside wage offer is correlated with skill, if the outside offer increases faster in skill than the expected payoff from staying in the firm, then the probability

that a manager stays is non-monotonic. In this example, although the probability of staying is non-monotonic, all defined conditions for equilibrium monotonicity (firm placement/wage offers, profits and managerial cutoffs) still hold.

R3. *An old manager is more likely to leave the firm if the CEO is old.*

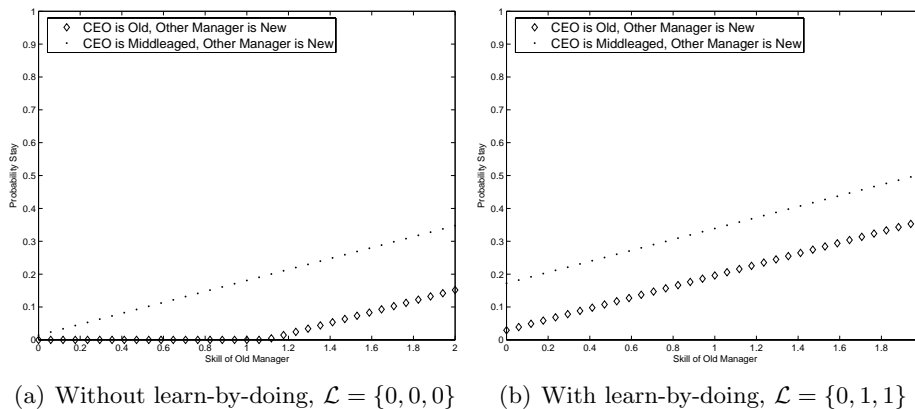


Figure 4: Quitting Strategies of an old manager when wages are perfectly flexible. Parameters: $\rho = 3$, $\bar{S} = 2$, $\hat{q}^{mid} = \hat{q}^{old} = 3$, $\bar{W}^c = 3$, $W^m = 1.5$, $\pi = .1$, $\beta = .8$, $\lambda = 0$.

Figures 4(a) and 4(b) illustrate the equilibrium probability that an old manager stays when wage offers are perfectly flexible: the old manager is more likely to stay when the CEO is middleaged (top line) than when the CEO is old (bottom line). When the contribution of the CEO's attributes to firm's output is very important, the firm has a strong incentive to replace old managers (who were not promoted to CEO and will retire before having another chance of promotion) with new young managers, who could turn out to be highly talented candidates for future CEO promotion. Result R3 then holds because the incentive to replace old managers with new hires is stronger if the current CEO is closer to retirement.

The bottom line of Figure 4(a) shows how strong this incentive can be. In this example, the outside wage offer is orthogonal to skill, $\lambda = 0$. Therefore, were the firm to offer any positive wage close to zero, there is some chance that the firm could induce any old manager to stay. Nonetheless, the firm prefers to lay off with probability one even old managers with slightly above-average skills, in order to hire a young manager at wage $W^m = 1.5$. Comparing Figures 4(a) and 4(b) reveals how learn-by-doing might decrease the turnover of old managers, since they become more productive. However, one can contemplate an economy where increasing learn-by-doing increases the turnover of old managers *who are not promoted* to CEO. To see this, suppose ρ and β are high, and π is low; i.e., the CEO position is sufficiently important, the firm is sufficiently patient, and CEO turnover

is low. In this case, greater learn-by-doing increases the value of replacing an old manager who will not be promoted to CEO with a new hire, so the young executive, who is a candidate for the CEO position, can acquire learn-by-doing. The result follows if this gain exceeds the increase in the firm's expected profits from keeping the old manager.

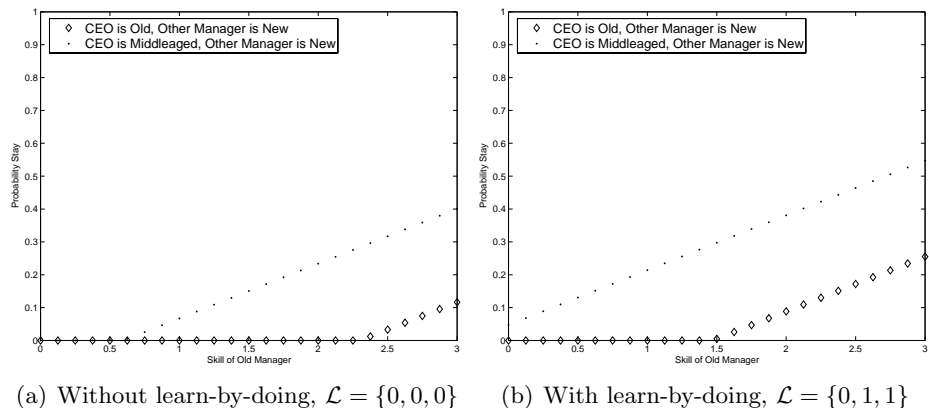


Figure 5: Quitting Strategies of a old manager when wages are perfectly flexible. Parameters: $\rho = 3$, $\bar{S} = 3$, $\hat{q}^{mid} = \hat{q}^{old} = 3$, $\bar{W}^c = 3$, $W^m = 1.5$, $\pi = .1$, $\beta = .8$, $\lambda = 0$.

The incentives to replace an experienced manager with a new hire also depend on the skill distribution of untried executives, as Figures 4 and 5 illustrate. In these examples, the young executives' maximum skill increases from $\bar{S} = 2$ to $\bar{S} = 3$. With more dispersion in skills, the firm has a greater incentive to fire less talented old executives in order to search. The bottom line of Figure 5(a) reveals that the firm optimally lays off most old managers when the CEO is old. This equilibrium feature may underlie the high turnover of executives in firms such as Merrill Lynch. Although many executives who leave the firm have high skills, the firm values the possibility of drawing extraordinarily skilled executives by so much that it is willing to replace relatively talented executives with new draws.

R4. *Given any wage offer, a middleaged manager is more likely to stay if the firm retains another manager who is either (i) older and closer to retirement, or (ii) middleaged and sufficiently less talented.*

Figure 3 illustrates result (i): the top line represents the equilibrium probability that a middleaged manager stays in the firm when the other manager is old, while the middle line represents the lower probability when the other manager is young. The result holds because the middleaged manager is more likely to win future CEO promotion if the other manager is old — an old manager will retire and hence not compete for future promotions, while a young manager might turn out to

be a talented competitor. Notice that as the middleaged manager’s skill increases, the incentives derived from keeping the old manager fall. As the middleaged manager becomes more talented, the probability that a young manager will be so skilled that he is promoted to CEO next period decreases. In particular, the bias toward promoting older managers implies that a sufficiently talented middleaged manager will be promoted with probability one when old independently of the young manager’s skill realization.

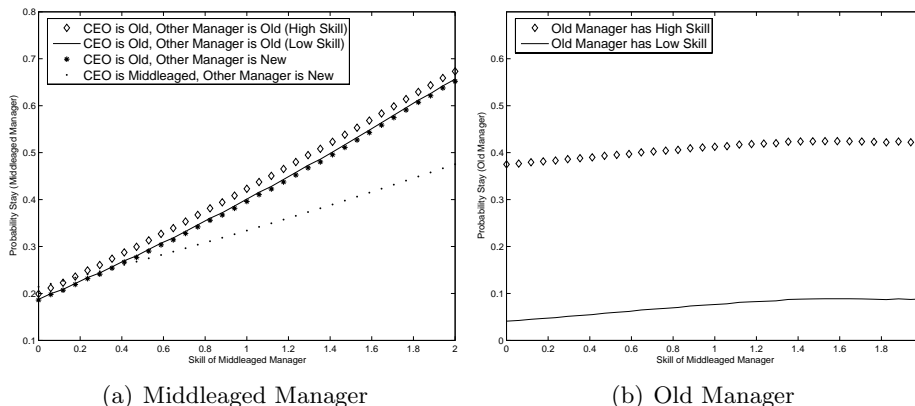


Figure 6: Quitting Strategies when wages are perfectly flexible and skill is firm-specific. Parameters: $\rho = 3$, $\bar{S} = 2$, $\hat{q}^{mid} = \hat{q}^{old} = 3$, $\bar{W}^c = 3$, $W^m = 1.5$, $\mathcal{L} = \{0, 1, 1\}$, $\pi = .1$, $\beta = .8$, $\lambda = 0$.

Figures 6(a) and 6(b) illustrate the probabilities that middleaged and old managers stay when wages are flexible. The firm’s wage offer to an old manager and his probability of staying increase with his skill. Therefore, the middleaged manager’s probability of staying rises with the old manager’s skill, as the top two lines of Figure 6(a) illustrate. Comparing Figures 3(a) and 6(a) reveals how the firm uses the flexibility in the wage offer to experienced executives to smooth incentives. The firm optimally adjusts managers’ probability of staying, closing the discontinuities created by the placement cutoffs and reducing the gaps between different attribute profiles.

Result R4(ii) holds trivially when wages are fixed: for a more able middleaged manager, a less able middleaged manager represents lesser competition for future CEO promotion. That is, if the firm keeps both middleaged managers, it promotes to CEO the most able executive in the following period. However, result R4(ii) might not hold when wages are flexible. When the firm has only one talented middleaged candidate, it is willing to offer him a high wage to raise the probability of keeping him. When the firm has two talented middleaged managers with similar skills, one can conceive of the possibility that the firm might optimally decrease its wage offer to both candidates: the firm gambles, expecting that at least one talented manager will stay to fill the single CEO position. This wage reduction could increase the quit probability of the more talented manager.

R5. *The firm's optimal placement/wage offers to a manager are complex, potentially non-monotonic functions of the attributes of all other executives.*

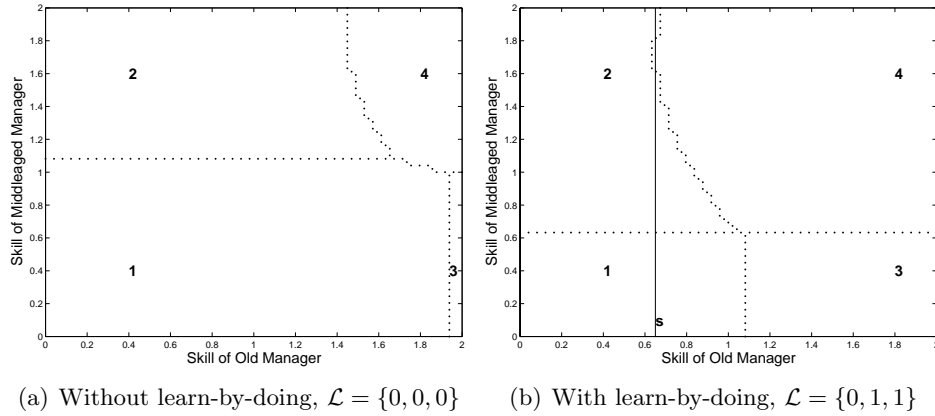


Figure 7: Optimal placement offer when the CEO is old, one manager is middleage and one manager is old, and wages are rigid. Parameters: $\rho = 3$, $\bar{S} = 2$, $\hat{q}^{mid} = \hat{q}^{old} = 3$, $\bar{W}^c = 3$, $W^m = 1.5$, $\pi = .1$, $\beta = .8$, $\lambda = 0$.

Although we identify conditions that ensure the firm's optimal strategy is monotone with respect to a manager's own skill, this strategy can easily be a non-monotonic function of the attributes of other executives. To understand why this is so, we first focus on the pure placement offer. Figure 7 shows the optimal placement offer when wage offers are perfectly rigid, the CEO is old, one manager is middleaged and one manager is old. The firm's optimal placement offer is to lay off both managers if they are unable (region 1); to keep only the able middleaged manager (region 2); to keep only the able old manager (region 3); or to keep both able managers (region 4).

When the firm cannot give incentives via wage adjustment, it uses its knowledge of the equilibrium behavior of middleaged managers to provide incentives via its placement offers. Consider an old manager with moderate skill s in Figure 7(b). If the middleaged manager has low skill, the firm optimally fires both managers in order to draw young executives, who might be suitable CEO candidates. Once the skill of the middleaged manager is sufficiently high, he becomes a suitable candidate for the CEO position and the firm wants to increase the probability that he will stay. The firm then would like to keep a moderately talented old manager s as an incentive to the able middleaged manager, in order to reduce the probability that the able middleaged manager quits. Moreover, since the firm already has a good candidate for the CEO position, the value of replacing the old manager with a young executive is reduced. If, instead, the middleaged manager is very skilled, he is so likely to be promoted in the next period that the firm no longer wants to keep old manager s as an incentive. As a result, the firm's equilibrium placement offer to the old manager

is a non-monotonic function of the middleaged manager’s skill.

When the firm has no flexibility to adjust wage offers, placement incentives to the middleaged manager have a discrete nature — the firm can only keep the old manager or not. When wages are flexible, then to adjust optimally the probability that the middleaged manager stays, the firm uses *both* the direct incentive of the middleaged manager’s wage offer, and the indirect incentive of the old manager’s placement/wage offer. Figure 6(b) depicts the equilibrium probability that the old manager quits as a function of his skill and that of the middleaged manager, and reflects the firm’s optimal wage offers. The firm increases the wage of the old manager as a function of the skill of the middleaged manager up to some point, in order to increase the probability that the old manager will stay and provide incentives to the able middleaged manager to stay. After this point, however, the middleaged manager is so likely to be promoted to CEO that the firm no longer needs to keep the old manager as an indirect incentive. The firm’s wage offer to the old manager then becomes decreasing in the skill of the middleaged manager. Hence, the firm’s wage offer to the old manager is a non-monotonic function of the middleaged manager’s skill.

In our framework, the firm defines placement and wages at each period, and cannot credible commit to long term contracts. In an alternative framework where the firm can offer long term contracts, the firm could conceivably prefer other incentives to keep the middleaged manager, instead of the indirect incentive of keeping the old manager. For example, the firm might commit to (a) always promote a middleaged manager of sufficient ability to CEO next period when he becomes old, or alternatively (b) give the middleaged manager a high wage in the cases where he is not promoted to CEO when old. However, if in the following period the younger manager turns out to be a better candidate for the CEO position, then the firm has a significant incentive to break the contract.

Non-Monotonicity: We finish this section by describing an economy for which the CEO promotion monotonicity does not hold. To generate non-monotonicities, we require implausible parameters: the upper bounds on the outside wage offer distribution ($\hat{q}^{mid} = \hat{q}^{old} = 1.8$) are quite small relative to the firm’s fixed wage offer to managers ($W^m = 1.5$), and the probability of CEO turnover is very high, $\pi = .9$. As a result, the firm worries sufficiently more about losing an executive promoted to CEO than about losing an executive whom it can “hide” as a manager.

Figure 8(a) represents the firm’s optimal placement offer when the CEO position is vacant and both managers are middleaged. In regions $M1$ and $M2$, the firm promotes the most able executive to CEO and lays off the other manager. In regions $M1^*$ and $M2^*$, the firm promotes the less able

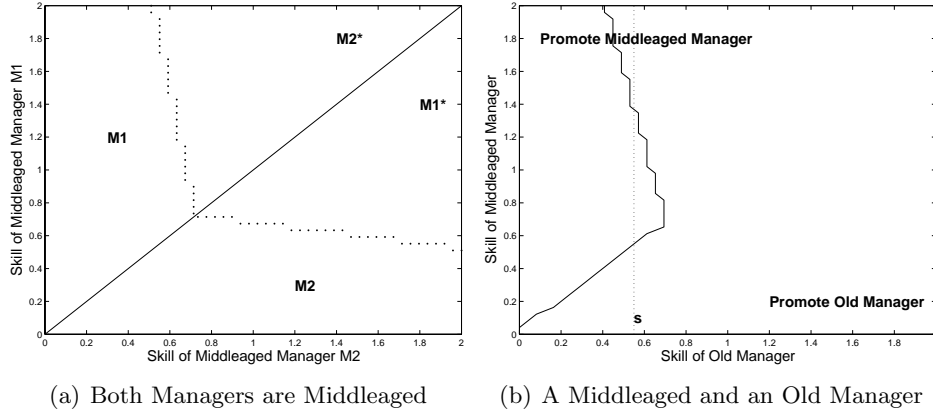


Figure 8: Optimal CEO promotion when wages are rigid. Parameters: $\rho = 2$, $\bar{S} = 2$, $\hat{q}^{mid} = \hat{q}^{old} = 1.8$, $\bar{W}^c = 3$, $W^m = 1.5$, $\pi = .9$, $\beta = .8$, $\lambda = 0$, $\mathcal{L} = \{0, 1, 1\}$.

executive and keeps the most able executive as a manager. Figure 8(b) presents the firm’s optimal placement offer when the CEO position is vacant, one manager is middleaged and one manager is old. Consider an old manager with skill s . If the middleaged manager has low skill, he is laid off and the old manager is promoted. If the middleaged manager has intermediate skill, he is promoted to CEO; but if his skill is high, then the old manager is promoted while the middleaged executive is kept as a manager, in order to raise the probability that he stays.

The five results from Proposition 1 still hold in this non-monotonic example.

4 Conclusion

This paper develops a dynamic model in which we explore how the attributes of a managerial workforce affect firms’ placement decisions and wage offers, and managers’ quit decisions. We first provide sufficient conditions on primitives such that the equilibrium is unique and features several key placement and payoff monotonicity properties. We then present five analytical results that characterize central features of equilibrium outcomes when the CEO position is important:

R1. The firm has a bias toward promoting old executives to CEO, since an old executive is more likely to leave the firm if not promoted to CEO, while a middleaged executive is likely to stay due to the prospects of future CEO promotion — this result can reconcile the empirical finding that, controlling for tenure and hierarchical level, the promotion probability of top executives increases with age;

R2. The prospect of imminent promotion to CEO makes a talented middleaged manager less likely

to leave the firm if the CEO is old — this equilibrium result can reconcile the empirical finding that turnover is higher when the CEO is younger;

R3. An old manager is more likely to leave the firm if the CEO is old, since the firm has a stronger incentive to replace an old manager who will retire soon with a younger executive, who might be a talented candidate for future CEO promotion;

R4. Given any wage offer, a middleaged manager is more likely to stay if the firm retains another manager who is either older and closer to retirement, or middleaged and sufficiently less talented, as these managers represent less competition for future CEO promotion;

R5. The firm’s optimal placement/wage offers to a manager might be a non-monotonic function of the attributes of all other executives.

In our model, the firm uses result 4 to optimally adjust the level of competition for the CEO promotion: the firm sometimes keeps less talented managers in order to reduce the turnover of talented executives. Result 5 is a direct consequence of this equilibrium interaction between the attributes of different managers. In real life, the firm can also use other instruments to reduce the expected competition for the CEO position and decrease turnover of talented executives, such as creating a reputation of typically promoting inside managers to CEO.

We also characterize how the firm trades off between wages and placement offers to align managerial incentives. In particular, the firm uses the flexibility in wage offers to “smooth” the discrete incentives generated by placement offers. The firm’s wage offer to a manager is not only a function of the manager’s own characteristics, but also the attributes of other executives in the firm.

An interesting feature that we leave for future research is the effects of inter-manager skill complementarities on equilibrium outcomes. In particular, consider a CEO who has a higher level of complementarities with one manager than another. This CEO and the managers are likely to take this into account when making placement and investment decisions. A CEO will tend to invest — give more important assignments — to managers whom he knows well and interacts with. In turn, a manager with whom the CEO has not worked understands that he is likely to receive less investment and as a consequence will find outside alternatives more attractive, and is more likely to quit. The firm understands these consequences and will take them into account when making its choice of which manager to promote to CEO and which managers to fire. Hayes, Oyer and Schaefer (2006) find evidence that such complementarity effects are empirically relevant.

A Appendix

A.1 State Space

In this section we define the state space \mathcal{Z} and firm's state variable $z_t \in \mathcal{Z}$. Young executives ($a = 1$) are hired by the firm during the current period. Therefore, along the equilibrium path, there is never a young CEO nor a young manager at the beginning of any period t . It is never optimal to hire a young executive directly to the CEO position; therefore, along the equilibrium path, there is never a middleaged CEO at the beginning of any period t , since he would have been a young executive promoted to CEO at period $t - 1$. Executive positions may be vacant because the employee retired — we use $\emptyset \equiv (0, 0)$ to represent a vacant executive position. The set of possible manager characteristics is $\mathcal{M} = \{2\} \times \mathcal{S}_2 \cup \{3\} \times \mathcal{S}_3 \cup \emptyset$, and the set of possible CEO characteristics is $\mathcal{C} = \{3\} \times \mathcal{S}_3 \cup \emptyset$. The state describing the firm at the beginning of period t (the distribution of the characteristics of the firm's executives) is represented by the executive-profile vector $z_t = (a_t^c, s_t^c, a_t^1, s_t^1, a_t^2, s_t^2)$.

We assume that at the beginning of the initial period $t = 1$: (i) the firm has at least one experienced executive; (ii) if the CEO is old, then at least one manager is middleaged; and (iii) if both managers are old, then the CEO position is vacant. The model's structure of promotions, quits and layoffs guarantees that these properties are preserved along the equilibrium path, i.e., (i) $\max \{a_t^c, a_t^1, a_t^2\} \in \{2, 3\}$; (ii) $a_t^c = 3 \Rightarrow a_t^1 = 2$ or $a_t^2 = 2$; and (iii) $a_t^1 = a_t^2 = 3 \Rightarrow a_t^c = 0$. Condition (ii) holds because an old CEO at the beginning of period t was promoted to CEO in period $t - 1$. Therefore, at the end of period $t - 1$, at least one manager was young (the one who replaced the manager promoted to CEO), which implies that at least one manager is middleaged at the beginning of period t . Consequently, condition (i) also holds: we never have three old executives, so at least one executive will stay for the following period. Finally, if both managers are old at period t , then both were middleaged managers at period $t - 1$, which implies that the CEO was old — otherwise, if the CEO position was vacant, one of the middleaged managers would have been promoted to CEO. Therefore, the old CEO will be retired at period t .

The relevant state space is

$$\mathcal{Z} = \left\{ z_t \in \mathcal{C} \times \mathcal{M} \times \mathcal{M} \text{ s.t. } \begin{array}{l} (i) \quad \max \{a_t^c, a_t^1, a_t^2\} \in \{2, 3\} \\ (ii) \quad a_t^c = 3 \Rightarrow a_t^1 = 2 \text{ or } a_t^2 = 2 \\ (iii) \quad a_t^1 = a_t^2 = 3 \Rightarrow a_t^c = 0 \end{array} \right\}. \quad (6)$$

It is useful to decompose states $z_t \in \mathcal{Z}$ into 8 representative partitions of the possible management structures at the beginning of the period, along the equilibrium path. To simplify presentation,

for a middleaged executive with skill x , we represent his productive characteristics $\{2, x\}$ as $2x$, and for an old executive we use $3x$; we use $3x + \tau$ when integrating over the possible changes τ in the old manager's skill.

Partition 1) Firm starts the period with no CEO and one middleaged manager. All $z_t \in \mathcal{Z}$ such that either $z_t = (\emptyset, \emptyset, 2x)$ or $z_t = (\emptyset, 2x, \emptyset)$, where $x \in \mathcal{S}_2$.

Partition 2) Firm starts the period with no CEO and one old manager. All $z_t \in \mathcal{Z}$ such that either $z_t = (\emptyset, \emptyset, 3x)$ or $z_t = (\emptyset, 3x, \emptyset)$, where $x \in \mathcal{S}_3$.

Partition 3) Firm starts the period with an old CEO and one middleaged manager. All $z_t \in \mathcal{Z}$ such that either $z_t = (3x, \emptyset, 2y)$ or $z_t = (3x, 2y, \emptyset)$, where $x \in \mathcal{S}_3$ and $y \in \mathcal{S}_2$.

Partition 4) Firm starts the period with no CEO, an old manager and a middleaged manager. All $z_t \in \mathcal{Z}$ such that either $z_t = (\emptyset, 3x, 2y)$ or $z_t = (\emptyset, 2y, 3x)$, where $x \in \mathcal{S}_3$ and $y \in \mathcal{S}_2$.

Partition 5) Firm starts the period with no CEO and two middleaged managers. All $z_t \in \mathcal{Z}$ such that $z_t = (\emptyset, 2x, 2y)$, where $x, y \in \mathcal{S}_2$.

Partition 6) Firm starts the period with no CEO and two old managers. All $z_t \in \mathcal{Z}$ such that $z_t = (\emptyset, 3x, 3y)$, where $x, y \in \mathcal{S}_3$.

Partition 7) Firm starts the period with an old CEO, a middleaged manager and an old manager. All $z_t \in \mathcal{Z}$ such that either $z_t = (3z, 3x, 2y)$ or $z_t = (3z, 2y, 3x)$, where $z, x \in \mathcal{S}_3$ and $y \in \mathcal{S}_2$.

Partition 8) Firm starts the period with an old CEO and two middleaged managers. All $z_t \in \mathcal{Z}$ such that $z_t = (3z, 2x, 2y)$, where $z \in \mathcal{S}_3$ and $x, y \in \mathcal{S}_2$.

Table 1 summarizes the eight representative partitions of executive age-profile at the beginning of the period, and the feasible firm's placement offers for each partition. We omit the symmetric state for each partition — where the divisions of managers are switched. After firm's placement/wage offer, there are other eight intermediate representative partitions, which are observed by the managers before they make quit decisions: $(2, 1, 1)$, $(3, 1, 1)$, $(3, 1, 2)$, $(2, 3, 1)$, $(2, 2, 1)$, $(3, 3, 1)$, $(3, 3, 2)$ and $(3, 2, 2)$. The same eight intermediate partitions are possible when production takes place, after managers make their quit decisions.

A.2 Equilibrium Characterization

Strategies: The equilibrium of the model is fully characterized by the collection of period strategies of managers and the firm, and by consistent beliefs about these strategies. A manager's period quit strategy is a function $\varphi_t^d : \mathcal{Z}^m \rightarrow \{0, 1\}$, and the firm's period employment

Initial partition	Firm's feasible placement offers	Intermediate partition	Managers' quit decision	Final partition
P1 (0,0,2)	Promote M2	(2,1,1)		(2,1,1)
P2 (0,0,3)	Promote M2	(3,1,1)		(3,1,1)
P3 (3,0,2)	(i) Lay off M2	(3,1,1)	M2 quits M2 stays	(3,1,1)
	(ii) Keep M2	(3,1,2)		(3,1,2)
P4 (0,3,2)	(i) Promote M1, lay off M2	(3,1,1)	M2 quits M2 stays	(3,1,1)
	(ii) Promote M1, keep M2	(3,1,2)		(3,1,2)
	(iii) Promote M2, lay off M1	(2,1,1)	M1 quits M1 stays	(2,1,1)
	(iv) Promote M2, keep M1	(2,3,1)		(2,3,1)
P5 (0,2,2)	(i) Promote M1, lay off M2	(2,1,1)	M2 quits M2 stays	(2,1,1)
	(ii) Promote M1, keep M2	(2,1,2)		(2,1,2)
	(iii) Promote M2, lay off M1	(2,1,1)	M1 quits M1 stays	(2,1,1)
	(iv) Promote M2, keep M1	(2,2,1)		(2,2,1)
P6 (0,3,3)	(i) Promote M1, lay off M2	(3,1,1)	M2 quits M2 stays	(3,1,1)
	(ii) Promote M1, keep M2	(3,1,3)		(3,1,3)
	(iii) Promote M2, lay off M1	(3,1,1)	M1 quits M1 stays	(3,1,1)
	(iv) Promote M2, keep M1	(3,3,1)		(3,3,1)
P7 (3,3,2)	(i) Lay off M1, lay off M2	(3,1,1)	M1 quits M1 stays	(3,1,1)
	(ii) Keep M1, lay off M2	(3,3,1)		(3,3,1)
	(iii) Lay off M1, keep M2	(3,1,2)	M2 quits M2 stays	(3,1,1) (3,1,2)
	(iv) Keep M1, keep M2	(3,3,2)	M1 quits, M2 quits M1 stays, M2 quits M1 quits, M2 stays M1 stays, M2 stays	(3,1,1) (3,3,1) (3,1,2) (3,3,2)
P8 (3,2,2)	(i) Lay off M1, lay off M2	(3,1,1)	M1 quits M1 stays	(3,1,1)
	(ii) Keep M1, lay off M2	(3,2,1)		(3,2,1)
	(iii) Lay off M1, keep M2	(3,1,2)	M2 quits M2 stays	(3,1,1) (3,1,2)
	(iv) Keep M1, keep M2	(3,2,2)	M1 quits, M2 quits M1 stays, M2 quits M1 quits, M2 stays M1 stays, M2 stays	(3,1,1) (3,2,1) (3,1,2) (3,2,2)

Table 1: Representative partitions of executive age-profile.

strategy is a function $\sigma_t : \mathcal{Z} \rightarrow \Delta \mathcal{E}_t(z_t)$. When the equilibrium is monotone and symmetric, the firm's placement strategy can be fully describe by two cutoff functions. Consider an executive profile $\{a^c, s^c, a^1, s^1, a^2, s^2\}$. When the CEO position is not vacant, $\{a^c, s^c\} \neq \emptyset$, the firm defines an optimal layoff cutoff $\gamma_t^m(a^d, a^c, s^c, a^{-d}, s^{-d})$, such that, with probability one, experienced manager d is laid off if $s^d < \gamma_t^m(a^d, a^c, s^c, a^{-d}, s^{-d})$, and is offered the managerial position if $s^d > \gamma_t^m(a^d, a^c, s^c, a^{-d}, s^{-d})$. When the CEO position is vacant, $\{a^c, s^c\} = \emptyset$, the firm's optimal promotion cutoff is a function $\gamma_t^c(a^d, a^{-d}, s^{-d})$. With probability one, experienced manager d is promoted to CEO if $s^d > \gamma_t^c(a^d, a^{-d}, s^{-d})$, while the experienced manager $-d$ is promoted if $s^d < \gamma_t^c(a^d, a^{-d}, s^{-d})$. When experienced manager d is not promoted to CEO, the firm uses the layoff cutoff function $\gamma_t^m(a^d, a^{-d}, s^{-d}, \emptyset)$ to keep or not the manager.

In the following sections we solve for the firm's optimal wage offer to the experienced manager of division d , $W_t^d(\cdot)$, as a function of his characteristics, the realized feasible placement offer p_t and the characteristics of other executives in the firm.

Beliefs: Three functions define consistent beliefs regarding strategies of the firm and managers: manager's quit probability and expected payoff, and the firm's expected profit.

A consistent belief about the probability that the manager of division d will stay in the firm is a function of the initial state z_t and firm's realized employment offer e_t , $Pr_t^d : \mathcal{Z} \times \mathcal{E}_t(z_t) \rightarrow [0, 1]$. By the definition of a manager's quit strategy, $\varphi_t^d(\cdot) = 1$ if the manager stays, and $\varphi_t^d(\cdot) = 0$ if he quits. Integrating over the possible outside wage offers, consistency implies

$$Pr_t^d(z_t, e_t) = \int \varphi_t^d(z_t, e_t, q_t^d) dG_t^d(q_t^d | z_t, e_t).$$

For each period t , a consistent belief about the expected discounted present value of profits along the equilibrium path — that is, assuming that firm and managers behave optimally thereafter — is a function $V_t : \mathcal{Z} \rightarrow \mathbb{R}$ of the firm's current state z_t . From now on we refer to V_t as the firm's profit function. When the manager of division d is old, a consistent belief about his expected period payoff — before the realization of the firm's employment offer and the outside wage offer — is a function $\mu_t^d(z_t) : \mathcal{Z} \rightarrow \mathbb{R}$ of the state z_t .

A.3 Proof of Theorem 1

We solve the model recursively to prove existence of a PBE. In the last period $t = N$, the firm closes and lays off all executives; hence, computing expected period payoffs V_N and μ_N^d is trivial.

At period $t = N - 1$, we solve the model for each relevant partition. Partitions 1 and 2

are trivial. For all other partitions, we first solve each manager's optimal quit strategy after observing the firm's employment offer. When the firm only keeps one experienced manager, it is straightforward to solve for his quit strategy. The manager observes the firm's employment offer, the CEO's attributes and he knows the skill distribution F_1 of the young manager hired for the other division. The experienced manager computes his expected discounted payoff from staying in the firm, and he will quit if and only if the realized outside wage offer yields higher payoff. When the firm keeps two experienced managers, they simultaneously decide whether to stay or quit. We need to show that this static game has a Nash Equilibrium on quit probabilities. Notice that manager $M1$ knows his expected payoff from staying in the firm conditional on manager $M2$ staying, and conditional on manager $M2$ quitting. Manager $M1$'s expected payoff is a continuous function of the probability that manager $M2$ quits, hence $M1$'s best response yields a quit probability for $M1$ that is continuous on $M2$'s quit probability. Quit probabilities must be in the compact set $[0, 1]$, hence an equilibrium exists to this stay/quit game. If there are multiple equilibria, fix one equilibrium.

We now solve for the firm's optimal employment offer. At the beginning of period $N-1$, the firm forms consistent beliefs regarding managers' quit probabilities for each feasible employment offer, and chooses the expected profit-maximizing offer. Recall that there is a finite number of feasible placement offers; if the set of feasible wage offers $\mathcal{H}_t^d(\cdot)$ is also finite, then a profit-maximizing employment offer exists. To show that such optimal employment offer exists when the *closed* set $\mathcal{H}_t^d(\cdot)$ is not finite, it is sufficient to show that, for any given placement offer, there exists an optimal wage offer. Fix a placement offer. The managers' outside wage offer distributions are bounded, $q_t^d \in [0, \hat{q}]$, hence there exists bounds on the firm's relevant set of feasible wage offers: (a) there exists a lower bound such that the manager quits with probability one if the firm offers a wage below this lower bound, and (b) there exists an upper bound such that the manager stays with probability one if the firm offers a wage above this upper bound. Therefore, we only need to consider the intersection of the *closed* feasible wage offer set $\mathcal{H}_t^d(\cdot)$ with these bounds, that is, a *compact* set of relevant feasible wage offers. Recall that the outside wage offer distribution $G_t^d(q_t^d | z_t, p_t, w_t)$ is continuous in q_t^d and the wage offer w_t . Differentiating the best response functions of managers with respect to the firm's wage offers, it is routine to show that equilibrium quit probabilities change continuously with the firm's wage offer if the sensitivity of $G_t^d(\cdot)$ to q_t^d and w_t is sufficiently small. Therefore, the firm's expected profit is a continuous function of its wage offers, which are in a compact set — an optimal wage offer exists. We then compute expected period payoffs V_{N-1} and μ_{N-1}^d .

This argument extends inductively to earlier periods.

A.4 Proof of Theorem 2

Consider any feasible collection of firm technology parameters. Let the distribution of outside wage offers be given by equations (3) and (4). To prove Theorem 2, we show that the following lemmas hold at the terminal period, $t = N$.

Lemma 1 (Firm's Placement Strategy) *The firm's optimal placement strategy takes the form of unique cutoff functions on the manager's own skill.*

That is, given the age of experienced manager d and the characteristics of all other current executives, the firm defines a promotion cutoff, $\gamma_t^c(\cdot)$, and a layoff cutoff, $\gamma_t^m(\cdot)$. When the CEO position is vacant, the manager of division d is promoted to CEO if his skill exceeds the promotion cutoff, i.e., $s^d > \gamma_t^c(\cdot)$; otherwise, he is offered the managerial position if his skill exceeds the layoff cutoff, i.e., $s^d > \gamma_t^m(\cdot)$, and he is laid off if his skill falls below the layoff cutoff, i.e., $s^d < \gamma_t^m(\cdot)$.

Lemma 2 (Firm's Wage Strategy) *The firm's optimal wage offer to experienced managers is a unique function that yields to manager d an expected payoff from staying in the firm that increases weakly in his skill s^d .*

Lemma 3 (Firm's Profits) *In equilibrium, the firm's expected profits increase weakly in each of manager's skill.*

Lemma 4 (Managers' Quit Strategies) *A manager's equilibrium quit strategy defines a unique cutoff on the outside wage offer; that is, a manager quits and accepts the outside offer if and only if the offer exceeds the cutoff.*

Lemma 5 (Managers' Payoff) *In equilibrium, the expected lifetime payoff of an experienced manager increases weakly in his own skill.*

Lemma 6 (Symmetry) *The firm's employment strategy, expected profits, managers' quit strategies and expected payoffs are symmetric.*

Lemma 7 (Technical) *From an ex ante perspective, i.e., before the realization of the change τ in each executive skill, the firm's expected profit function and an old managers' expected payoff*

- $\iiint V_{t+1}(3x + \tau, 2s_1, 2s_2) dT(\tau) dF_2(s_1) dF_2(s_2)$

- $\iint V_{t+1}(\emptyset, 3x + \tau, 2s) dT(\tau) dF_2(s)$
- $\iint V_{t+1}(\emptyset, 3x + \tau_1, 3y + \tau_2) dT(\tau_1) dT(\tau_2)$
- $\iint \mu_{t+1}^1(\emptyset, 3x + \tau, 2s) dT(\tau) dF_2(s)$
- $\iiint \mu_{t+1}^1(3y + \tau_1, 3x + \tau_2, 2s) dT(\tau) dT(\tau_1) dT(\tau_2) dF_2(s)$
- $\iint \mu_{t+1}^1(\emptyset, 3x + \tau_1, 3y + \tau_2) dT(\tau_1) dT(\tau_2)$

are differentiable with respect to the skill x of the old manager.

For periods $t < N$, we use an induction argument to prove that if Lemmas 1-7 hold in period $t + 1$, then they also hold for each partition in period t when (C1) upper bounds $\{\hat{q}^{mid}, \hat{q}^{old}\}$ are sufficiently large, and (C2) CEO position is important, i.e., ρ , \bar{W}_t^c and $\rho - \frac{\partial w_t^c(a^c, s^c)}{\partial s^c}$ are sufficiently large. We define and fully characterize the optimal strategies and the consistent beliefs.

A.4.1 Terminal Period

At the terminal period $t = N$ the firm closes, lays off all remaining executives and receives zero profits. The laid-off managers receive and take their outside wage offers. If the firm starts period N with an old CEO, he receives and takes the outside wage offer $W_N^c(z_N)$. The value of the old CEO's outside wage offer at period N is unimportant for our qualitative results as long as it is always optimal for a middleaged manager to accept the promotion at period $t = N - 1$.

Equilibrium strategies are $\sigma_N(z_N) = (0, 0, 0, 0, 0, 0, 0)$ and $\varphi_N^d(z_N, e_N) = 0$. Expected profits are zero, $V_N(z_N) = 0$. The firm's wage offer function $W_t^d(\cdot)$, placement cutoff functions $\gamma_N^c(\cdot)$ and $\gamma_N^m(\cdot)$, and consistent belief $Pr_t^d(\cdot)$ are trivial. Lemmas 1 to 4 hold trivially.

At period $t = N$, the expected lifetime payoff of an experienced manager is his expected outside wage offer, which is symmetric and weakly increasing in his own skill — strictly increasing if $\lambda > 0$. The expected payoff of an old manager of division d is $\mu_N^d(z_N) = \lambda s^d + \bar{q}^{old}$.

Hence, Lemmas 5 to 7 hold, concluding the proof for period $t = N$.

A.4.2 Period $t < N$

Lemma 8 (Induction Lemma) *If Lemmas 1-7 hold in period $t + 1$, then Lemmas 1-7 also hold in period t .*

Using the induction argument, assume Lemmas 1-7 hold in period $t + 1$. To complete the proof, we show that Lemmas 1-7 hold in each of the 8 partitions at period t .

Partition 1) No CEO, unique middleaged manager: First consider $z_t = (\emptyset, \emptyset, 2x)$ and $x \in \mathcal{S}_2$. The firm has a unique feasible placement offer — promote the sole inside manager and hire two young managers. The executive promoted to CEO receives wage $W_t^c(2x)$ and each young manager receives W_t^m . The firm's unique optimal strategy is $\sigma_t(z_t) = (0, 1, 0, 0, W_t^c(2x), W_t^m, W_t^m)$. Promotion/layoff cutoffs and wage offers are trivial. All conditions of Lemmas 1, 2 and 5 are satisfied. There is no quit decision, so Lemma 4 holds trivially.

The promoted CEO produces $\rho(l_2 + x)$. Each young manager has an expected period output of $(l_1 + \bar{s})$, contributing an expected net period profit $\eta_t \equiv l_1 + \bar{s} - W_t^m$. Total expected period profit is then $\rho(l_2 + x) - W_t^c(2x) + 2\eta_t$. With probability $1 - \pi_t$, the CEO stays in the firm next period and becomes old; with probability π_t , the CEO leaves and the position becomes vacant. The skill of each executive evolves according to equation (1): the CEO's skill x becomes $x + \tau$, where τ follows the distribution T ; and period $t + 1$ skill of a young manager hired at period t follows the distribution F_2 . Integrating over the possible skills of each executive, the firm's expected discounted future profit is

$$\bar{v}_{t+1}(x) \equiv \iiint \left[[1 - \pi_t] V_{t+1}(3x + \tau, 2s_1, 2s_2) + \pi_t V_{t+1}(\emptyset, 2s_1, 2s_2) \right] dT(\tau) dF_2(s_1) dF_2(s_2).$$

Discounting future profits by β , the firm's profit function for partition 1 is

$$V_t(\emptyset, \emptyset, 2x) = \rho(l_2 + x) - W_t^c(2x) + 2\eta_t + \beta \bar{v}_{t+1}(x).$$

Next we show that $V_t(\emptyset, \emptyset, 2x)$ strictly increases in x . For any x ,

$$\frac{\partial V_t(\emptyset, \emptyset, 2x)}{\partial x} = \rho - \frac{\partial W_t^c(2x)}{\partial x} + \beta \left[1 - \pi_t \right] \frac{\partial}{\partial x} \iiint V_{t+1}(3x + \tau, 2s_1, 2s_2) dT(\tau) dF_2(s_1) dF_2(s_2).$$

The term $\rho - \frac{\partial W_t^c(2x)}{\partial x}$ is strictly positive from equation (2). Since Lemmas 3 and 7 holds at period $t + 1$, $\frac{\partial}{\partial x} \iiint V_{t+1}(3x + \tau, 2s_1, 2s_2) dT(\tau) dF_2(s_1) dF_2(s_2) \geq 0$. The firm's profit function $V_t(\emptyset, \emptyset, 2x)$ is differentiable and strictly increases in x ; Lemma 3 holds¹⁵ at period t . Symmetric results hold when $z_t = (\emptyset, 2x, \emptyset)$, concluding Lemma 6 and the proof for partition 1 at period t .

Partition 2) No CEO, one old manager: First consider $z_t = (\emptyset, \emptyset, 3x)$ and $x \in \mathcal{S}_3$. The firm has a unique feasible placement offer — promote the sole manager and hire two young managers.

¹⁵If the CEO turnover probability π_t varies with the CEO's skill, then the change in CEO turnover also affects expected profits. In this case, one can define a sufficiently small upper bound on $\left| \frac{\partial \pi_t(s^c)}{\partial s^c} \right|$, i.e., $\left| \frac{\partial \pi_t(s^c)}{\partial s^c} \right| \leq \bar{\pi}$, $\bar{\pi} > 0$ sufficiently small, such that $V_t(\emptyset, \emptyset, 2x)$ strictly increases in the middleaged CEO skill. The same argument applies to the other partitions when a middleaged executive is promoted to CEO.

The executive promoted to CEO receives wage $W_t^c(3x)$ and each young manager receives W_t^m . The firm's unique optimal strategy is $\sigma_t(z_t) = (0, 1, 0, 0, W_t^c(3x), W_t^m, W_t^m)$. Promotion/layoff cutoffs and wage offers are trivial. All conditions of Lemmas 1, 2 and 5 hold. There is no quit decision, so Lemma 4 holds trivially. The old manager's expected payoff is the CEO wage, $\mu_t^2(z_t) = W_t^c(3x)$.

Period profits are $\rho(l_3 + x) - W_t^c(3x) + 2\eta_t$. The promoted old CEO will retire at period $t + 1$, leaving the position vacant. The only uncertainties are the skills of the two young managers — the firm's expected discounted future profit is $\bar{\eta}_{t+1} \equiv \iint V_{t+1}(\emptyset, 2s_1, 2s_2) dF_2(s_1) dF_2(s_2)$. The firm's profit function for partition 2 is

$$V_t(\emptyset, \emptyset, 3x) = \rho(l_3 + x) - W_t^c(3x) + 2\eta_t + \beta\bar{\eta}_{t+1}. \quad (7)$$

The old CEO will retire, so his skill x only affects the current period's payoff,

$$\frac{\partial V_t(\emptyset, \emptyset, 3x)}{\partial x} = \rho - \frac{\partial W_t^c(3x)}{\partial x} > 0,$$

where the inequality comes from equation (2). Therefore, the profit function strictly increases in x , and Lemma 3 holds. Symmetric results follow when $z_t = (\emptyset, 3x, \emptyset)$, concluding Lemma 6. Lemma 7 holds since the CEO wage function $W_t^c(3x)$ is differentiable with respect to the old executive's skill x , concluding the proof for partition 2 at period t .

Partition 3) Old CEO, one middleaged manager: First consider $z_t = (3x, \emptyset, 2y)$, $x \in \mathcal{S}_3$ and $y \in \mathcal{S}_2$. The firm has an old CEO who receives wage $W_t^c(3x)$, division 1 is vacant and division 2 has a middleaged manager. The two feasible placement offers are (i) lay off the middleaged manager or (ii) keep him. Expected profits when the firm lays off the middleaged manager and hires two young managers are

$$(i) \quad \rho(l_3 + x) - W_t^c(3x) + 2\eta_t + \beta\bar{\eta}_{t+1}. \quad (8)$$

If the firm makes employment offer $e_t = (0, 0, 0, 1, W_t^c(3x), W_t^m, W_t^2)$, we must compute the probability the middleaged manager stays given the wage offer W_t^2 . If the middleaged manager stays, then with probability one next period's state z_{t+1} will be in partition 4 — the current old CEO retires, the middleaged manager becomes old, and the young manager becomes middleaged. The skill of each executive evolves according to equation (1): the middleaged manager's skill y becomes $y + \tau$, where τ follows the distribution T ; and period $t + 1$ skill of a young manager hired at period t follows the distribution F_2 . Integrating over the possible skills of each executive, the expected payoff of the current middleaged manager when he becomes old under partition 4 is

$$\mu_{t+1}^{P4}(y) \equiv \iint \mu_{t+1}^2(\emptyset, 2s, 3y + \tau) dF_2(s) dT(\tau).$$

The function $\mu_{t+1}^{P4}(y)$ is differentiable and weakly increases in y since Lemmas 5 and 7 hold at period $t+1$. The middleaged manager's expected payoff from staying at the firm is $W_t^2 + \beta\mu_{t+1}^{P4}(y)$. Therefore, his optimal quit strategy $\varphi_t^2(z_t, e_t, q_t^2)$ is to stay if and only if this expected payoff exceeds the realized outside wage offer $q_t^2 = (1 + \beta)(\lambda y + \xi_t^{mid})$. Therefore, Lemma 4 holds.

Now compute the middleaged manager's probability of staying. By definition, $\xi_t^{mid} \sim U[0, \hat{q}^{mid}]$. Hence, the manager quits with probability one if the inside wage offer W_t^2 falls below $\underline{h}_t(y) \equiv (1 + \beta)\lambda y - \beta\mu_{t+1}^{P4}(y)$, and he stays with probability one if the wage offer exceeds $\bar{h}_t(y) \equiv (1 + \beta)(\lambda y + \hat{q}^{mid}) - \beta\mu_{t+1}^{P4}(y)$. The probability that the middleaged manager stays is

$$Pr_t^{P3}(y, W_t^2) = \begin{cases} 1 & \text{if } W_t^2 > \bar{h}_t(y), \\ \frac{W_t^2 + \beta\mu_{t+1}^{P4}(y) - (1+\beta)\lambda y}{(1+\beta)\hat{q}^{mid}} & \text{if } W_t^2 \in [\underline{h}_t(y), \bar{h}_t(y)], \\ 0 & \text{if } W_t^2 < \underline{h}_t(y). \end{cases}$$

If the firm chooses placement offer (ii) and offers wage W_t^2 , then with probability $1 - Pr_t^{P3}(y, W_t^2)$ the middleaged manager quits and the firm hires a young manager for his position, which yields the same expected profit as placement offer (i). With probability $Pr_t^{P3}(y, W_t^2)$ the middleaged manager stays, produces $l_2 + y$ during this period and becomes old in the next period. The firm's expected discounted profit if the middleaged manager stays is

$$\rho(l_3 + x) - W_t^c(3x) + l_2 + y - W_t^2 + \eta_t + \beta \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dF_2(s) dT(\tau). \quad (9)$$

Subtracting equation (8) from (9), the net discounted expected profit from keeping the middleaged manager is

$$L_t^{P3}(y, W_t^2) \equiv \left[l_2 + y - W_t^2 + \beta \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dF_2(s) dT(\tau) - \eta_t - \beta\bar{\eta}_{t+1} \right].$$

Hence, the firm's expected discounted profit from offer (ii) is

$$(ii) \quad \rho(l_3 + x) - W_t^c(3x) + 2\eta_t + \beta\bar{\eta}_{t+1} + \max_{W_t^2 \in \mathbb{R}} Pr_t^{P3}(y, W_t^2) L_t^{P3}(y, W_t^2).$$

The middleaged manager quits with probability one if $W_t^2 \leq \underline{h}_t(y)$ — such a wage offer is equivalent to firing the manager — so following placement offer (ii) we need only consider the relevant wage offers that yield a strictly positive probability of retaining the manager. For any $Pr_t^{P3}(y, W_t^2) > 0$, the firm keeps the manager — chooses offer (ii) over offer (i) — if and only if the net profit from keeping the middleaged manager is positive, i.e., $L_t^{P3}(y, W_t^2) > 0$. A middleaged manager stays with probability one if $W_t^2 \geq \bar{h}_t(y)$. Therefore, the firm never offers a wage above $\bar{h}_t(y)$. When the firm wants to keep the middleaged manager, it chooses the wage $W_t^2 \in [\underline{h}_t(y), \bar{h}_t(y)]$ that maximizes

the net expected discounted profit from placement offer (ii),

$$\frac{W_t^2 + \beta\mu_{t+1}^{P4}(y) - (1 + \beta)\lambda y}{(1 + \beta)\hat{q}^{mid}} \left[l_2 + y - W_t^2 + \beta \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dF_2(s) dT(\tau) - \eta_t - \beta\bar{\eta}_{t+1} \right]. \quad (10)$$

The net expected profit is strictly concave in W_t^2 . Therefore, for each y there exists a unique optimal wage offer $W_t^{P3*}(y)$. If \hat{q}^{mid} is sufficiently large, then it is never profitable for the firm to offer the highest wage $\bar{h}_t(y)$, and the solution is interior. In this case, the first-order condition implies

$$W_t^{P3*}(y) = \frac{1}{2} \left[l_2 + y + \beta \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dF_2(s) dT(\tau) - \eta_t - \beta\bar{\eta}_{t+1} - \beta\mu_t^{P4}(y) + (1 + \beta)\lambda y \right]. \quad (11)$$

Define $L_t^{P3*}(y) \equiv L_t^{P3}(y, W_t^{P3*}(y))$. We next show that the firm's equilibrium expected net discounted profit from keeping the middleaged manager,

$$L_t^{P3*}(y) = \frac{1}{2} \left[l_2 + (1 - \lambda)y + \beta \left[\mu_{t+1}^{P4}(y) + \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dF_2(s) dT(\tau) - \lambda y \right] - \eta_t - \beta\bar{\eta}_{t+1} \right], \quad (12)$$

strictly increases in y . It suffices to show that $\mu_{t+1}^{P4}(y) + \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dF_2(s) dT(\tau)$ increases in y at a rate of at least λ . To see this, notice that by the induction argument, both $\mu_{t+1}^{P4}(y)$ and $V_{t+1}(\emptyset, 2s, 3y + \tau)$ increase in y . When the manager is laid off or quits, his expected payoff (outside wage offer) increases at rate λ . When the firm keeps the executive, the change in the wage offer with respect to y enters the firm's profits and the manager's expected payoff with equal value but opposite signs — they cancel out. Expected output increases in y at rate $1 > \lambda$ when the executive is kept as a manager, and at rate $\rho > 1 > \lambda$ when he is promoted to CEO. Therefore, when the executive stays, the firm's profit plus manager's payoff increases at rate strictly greater than λ , and $L_t^{P3*}(y)$ strictly increases in y . Consequently, the firm defines a layoff cutoff on the manager's own skill y , as a function of the attributes of all executives: the manager's own age, the CEO's age and skill, and the fact that the other division is vacant. The firm's optimal layoff cutoff $\gamma_t^m(2, 3x, \emptyset)$ is the minimum $y \in \mathcal{S}_2$ such that $L_t^{P3*}(y)$ is non-negative — the middleaged manager is offered the managerial position if his skill y exceeds the layoff cutoff $\gamma_t^m(2, 3x, \emptyset)$, and he is laid off if his skill falls below the layoff cutoff, concluding Lemma 1. When $y > \gamma_t^m(2, 3x, \emptyset)$, the optimal wage offer gives the middleaged manager a payoff from staying in the firm that strictly increases in y ,

$$W_t^{P3*}(y) + \beta\mu_t^{P4}(y) = L_t^{P3*}(y) + (1 + \beta)\lambda y, \quad (13)$$

concluding Lemmas 2 and 5. The middleaged manager stays in the firm with probability

$$Pr_t^{P3*}(y) = \begin{cases} \frac{L_t^{P3*}(y)}{(1 + \beta)\hat{q}^{mid}} & \text{if } y > \gamma_t^m(2, 3x, \emptyset), \\ 0 & \text{if } y < \gamma_t^m(2, 3x, \emptyset). \end{cases}$$

As a side note, observe that the expected lifetime payoff increases with skill, but the current wage offer $W_t^{P3*}(y)$ might decrease: if next period's expected payoff $\mu_{t+1}^{P4}(y)$ increases in y rapidly relative to the RHS of equation (13), then the firm exploits the manager's greater future payoff by *decreasing* his current wage offer.

Define the *total surplus*¹⁶ created when the manager stays in the firm as the sum of the firm's change in profit, $l_2 + y + \beta \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dF_2(s) dT(\tau) - \eta_t - \beta \bar{\eta}_{t+1}$, plus the manager's change in payoff, $\beta \mu_t^{P4}(y) - (1 + \beta)\lambda y$. Therefore, $L_t^{P3*}(y)$ is half of the total surplus. The firm's optimal wage offer gives the manager an expected lifetime payoff from staying in the firm equal to half of the total surplus plus $(1 + \beta)\lambda y$, the deterministic component of the outside wage offer.

The firm's optimal strategy is

$$\sigma_t(3x, \emptyset, 2y) = \begin{cases} (0, 0, 0, 1, W_t^c(3x), W_t^m, W_t^{P3*}(y)) & \text{if } y > \gamma_t^m(2, 3x, \emptyset) \\ (0, 0, 0, 0, W_t^c(3x), W_t^m, W_t^m) & \text{if } y < \gamma_t^m(2, 3x, \emptyset), \end{cases}$$

where $W_t^{P3*}(y)$ is given by equation (11). Let $\Gamma_t^{P3}(y) = 1$ if $y > \gamma_t^m(2, 3x, \emptyset)$, and $\Gamma_t^{P3}(y) = 0$ otherwise. The firm's profit function for partition 3 is

$$V_t(3x, \emptyset, 2y) = \rho(l_3 + x) - W_t^c(3x) + 2\eta_t + \beta \bar{\eta}_{t+1} + \frac{\Gamma_t^{P3}(y)}{(1 + \beta)\hat{q}^{mid}} \left[L_t^{P3*}(y) \right]^2.$$

The firm's profits and the old CEO payoff are differentiable and strictly increase in x by equation (2), concluding Lemma 7. Profits are independent of y for $y \leq \gamma_t^m(2, 3x, \emptyset)$, and strictly increasing in y for $y > \gamma_t^m(2, 3x, \emptyset)$, concluding Lemma 3. When $z_t = (3x, 2y, \emptyset)$, symmetric results follow from symmetry in V_{t+1} and μ_{t+1}^{P4} , concluding Lemma 6 and the proof for partition 3 at period t .

Partition 4) No CEO, an old manager and a middleaged manager: First consider $z_t = (\emptyset, 3x, 2y)$, $x \in \mathcal{S}_3$ and $y \in \mathcal{S}_2$. The firm has four feasible placement offers: (i) promote the old manager to CEO and lay off the middleaged manager; (ii) promote the old manager and keep the middleaged manager; (iii) promote the middleaged manager and lay off the old manager; (iv) promote the middleaged manager and keep the old manager.

Conditional on promoting the old manager to CEO, expected profits from offers (i) and (ii) are the same as $V_t(3x, \emptyset, 2y)$ in partition 3. Hence, the firm's placement and wage offer, and manager's expected payoff and quit strategies are the same as in partition 3, and all lemmas hold.

We now compute optimal strategies and equilibrium payoffs conditional on promoting the middleaged manager, and show that all lemmas hold. We then derive the optimal promotion strategy.

¹⁶The surplus does not include wage W_t^2 , which is simply a transfer from the firm to the manager.

Lemma 9 *Conditional on promoting the middleaged manager to CEO, Lemmas 1-7 hold in partition 4.*

Proof. If the middleaged executive is promoted to CEO, he receives wage $W_t^c(2y)$. The firm's expected profit from placement offer (iii) is

$$(iii) \quad \rho(l_2 + y) - W_t^c(2y) + 2\eta_t + \beta\bar{v}_{t+1}(y), \quad (14)$$

where $\bar{v}_{t+1}(y)$ was defined in partition 1. When the firm chooses placement offer (iv), the old manager's optimal strategy is to quit if and only if wage offer W_t^1 is below the realized outside wage offer. Given the firm's wage offer W_t^1 , the probability that the old manager stays is

$$Pr^{old}(x, W_t^1) = \begin{cases} 1 & \text{if } W_t^1 > \lambda x + \hat{q}^{old}, \\ \frac{W_t^1 - \lambda x}{\hat{q}^{old}} & \text{if } W_t^1 \in [\lambda x, \lambda x + \hat{q}^{old}], \\ 0 & \text{if } W_t^1 < \lambda x. \end{cases} \quad (15)$$

If the firm chooses placement offer (iv) and offers wage W_t^1 , then with probability $1 - Pr^{old}(x, W_t^1)$ the old manager quits and the firm has to hire a young manager for his position, which yields the same expected profit as placement offer (iii). With probability $Pr^{old}(x, W_t^1)$ the old manager stays, produces $l_3 + x$ in period t and retires in $t + 1$. Define

$$\underline{v}_{t+1}(y) \equiv \iint \left[[1 - \pi_t] V_{t+1}(3y + \tau, \emptyset, 2s_2) + \pi_t V_{t+1}(\emptyset, \emptyset, 2s_2) \right] dT(\tau) dF_2(s_2);$$

the firm's expected profit if the old manager stays is

$$\rho(l_2 + y) - W_t^c(2y) + l_3 + x - W_t^1 + \eta_t + \beta\underline{v}_{t+1}(y). \quad (16)$$

Subtracting (14) from (16), the net profit from keeping the old manager is

$$L_t^{P4}(x, W_t^1) \equiv [l_3 + x - W_t^1 + \beta\underline{v}_{t+1}(y) - \eta_t - \beta\bar{v}_{t+1}(y)].$$

The difference in continuation values,

$$\begin{aligned} \underline{v}_{t+1}(y) - \bar{v}_{t+1}(y) &= \iiint \left\{ [1 - \pi_t] [V_{t+1}(3y + \tau, \emptyset, 2s_2) - V_{t+1}(3y + \tau, 2s_1, 2s_2)] \right. \\ &\quad \left. + \pi_t [V_{t+1}(\emptyset, \emptyset, 2s_2) - V_{t+1}(\emptyset, 2s_1, 2s_2)] \right\} dT(\tau) dF_2(s_1) dF_2(s_2), \quad (17) \end{aligned}$$

is independent of the old manager's skill x and his wage offer W_t^1 . Moreover, the term $V_{t+1}(3y + \tau, \emptyset, 2s_2) - V_{t+1}(3y + \tau, 2s_1, 2s_2)$ is independent of y : CEO wage $W_{t+1}^c(3y + \tau)$ and output $\rho(l_3 + y + \tau)$ cancel out, and the firm's employment offer and managers' quit decisions are independent of the

old CEO's skill¹⁷. Therefore, $L_t^{P4}(x, W_t^1)$ is independent of the CEO skill y ¹⁸, strictly increases in skill x , and strictly decreases in wage W_t^1 . The firm's expected profit from offer (iv) is

$$(iv) \quad \rho(l_2 + y) - W_t^c(2y) + 2\eta_t + \beta\bar{v}_{t+1}(y) + \max_{W_t^1 \in \mathbb{R}} Pr^{old}(x, W_t^1)L_t^{P4}(x, W_t^1).$$

The old manager quits with probability one if $W_t^1 \leq \lambda x$ — such a wage offer is equivalent to firing the manager — so following placement offer (iv) we need only consider the relevant wage offers that yield a strictly positive probability of retaining the manager. For any $Pr^{old}(x, W_t^1) > 0$, the firm keeps the manager — chooses offer (iv) over offer (iii) — if and only if the net profit from keeping the old manager is positive, i.e., $L_t^{P4}(x, W_t^1) > 0$. An old manager stays with probability one if $W_t^1 \geq \lambda x + \hat{q}^{old}$. Therefore, the firm never offers a wage above $\lambda x + \hat{q}^{old}$. When the firm wants to keep the old manager, it chooses the wage $W_t^1 \in [\lambda x, \lambda x + \hat{q}^{old}]$ that maximizes the net expected discounted profit from placement offer (iv),

$$\frac{W_t^1 - \lambda x}{\hat{q}^{old}} \left[l_3 + x - W_t^1 + \beta\underline{v}_{t+1}(y) - \eta_t - \beta\bar{v}_{t+1}(y) \right].$$

The net expected profit is strictly concave in W_t^1 . Therefore, for each x there exists a unique optimal wage offer $W_t^{P4*}(x)$. If \hat{q}^{old} is sufficiently large, then it is never profitable for the firm to offer the highest wage $\lambda x + \hat{q}^{old}$, and the solution is interior. In this case, the first-order condition implies

$$W_t^{P4*}(x) = \frac{1}{2} \left[l_3 + (1 + \lambda)x + \beta\underline{v}_{t+1}(y) - \eta_t - \beta\bar{v}_{t+1}(y) \right]. \quad (18)$$

The old manager's wage increases in his skill x at rate $(1 + \lambda)/2$ — the rate is strictly greater than the rate of increase in the expected outside offer (λ), but strictly less than the increase in firm's output. Define $L_t^{P4*}(x) \equiv L_t^{P4}(x, W_t^{P4*}(x))$. The equilibrium net profit from keeping the old manager strictly increases in his skill x ,

$$L_t^{P4*}(x) = \frac{1}{2} \left[l_3 + (1 - \lambda)x + \beta\underline{v}_{t+1}(y) - \eta_t - \beta\bar{v}_{t+1}(y) \right].$$

The firms' optimal layoff cutoff $\gamma_t^m(3, 2y, \emptyset)$ is the minimum $x \in \mathcal{S}_3$ such that $L_t^{P4*}(x)$ is non-negative. When $x > \gamma_t^m(3, 2y, \emptyset)$, the optimal wage offer $W_t^{P4*}(x)$ gives the old manager a payoff from staying in the firm that strictly increases in x . Once more, the firm's optimal wage offer

¹⁷This is trivially true at the terminal period, $V_N(\cdot) = 0$. In earlier periods, this independence follows from the equilibrium characterization of partitions 3 and 8.

¹⁸If the CEO turnover probability π_t is a function of the CEO skill, then $L_t^{P4}(\cdot)$ is also a function of the CEO skill. In this case, one can define a sufficiently small upper bound $\bar{\pi} > 0$ on equation (5) such that $L_t^{P4}(x, W_t^1, y)$ does not decrease too fast with the CEO skill; consequently, the firm's expected profit from offer (iv) increases in the skill of the middleaged CEO.

gives the old manager a payoff from staying equal to half of the total surplus plus the deterministic component of the outside wage offer. The old manager stays in the firm with probability

$$Pr_t^{P4*}(x) = \begin{cases} \frac{L_t^{P4*}(x)}{\hat{q}^{old}} & \text{if } x > \gamma_t^m(3, 2y, \emptyset), \\ 0 & \text{if } x < \gamma_t^m(3, 2y, \emptyset). \end{cases}$$

Let $\Gamma_t^{P4}(x) = 1$ if $x > \gamma_t^m(3, 2y, \emptyset)$, and $\Gamma_t^{P4}(x) = 0$ otherwise. Conditional on promoting the middleaged executive, the firm's expected profits are

$$v_t(2y, 3x, \emptyset) \equiv \rho(l_2 + y) - W_t^c(2y) + 2\eta_t + \beta\bar{v}_{t+1}(y) + \frac{\Gamma_t^{P4}(x)}{\hat{q}^{old}} \left[L_t^{P4*}(x) \right]^2.$$

Profits are independent of x for $x \leq \gamma_t^m(3, 2y, \emptyset)$, and strictly increasing in x for $x > \gamma_t^m(3, 2y, \emptyset)$. In partition 1 we proved that $\rho(l_2 + y) - W_t^c(2y) + 2\eta_t + \beta\bar{v}_{t+1}(y)$ is differentiable and strictly increases in y , hence so is $v_t(2y, 3x, \emptyset)$. It is straightforward to show that Lemma 7 holds.

For $z_t = (\emptyset, 2y, 3x)$, symmetric results follow from symmetry in V_{t+1} , concluding the proof. ■

Lemma 10 *In partition 4, the firm uses a monotone strategy to promote a manager to CEO.*

Proof. Define the function $d_t(3x, 2y) \equiv V_t(3x, \emptyset, 2y) - v_t(2y, 3x, \emptyset)$. It is optimal to promote the old manager to CEO if $d_t(3x, 2y) > 0$, and it is optimal to promoted the middleaged manager if $d_t(3x, 2y) < 0$. We will show that $d_t(3x, 2y)$ strictly increases in x and strictly decreases in y , and solve for the optimal CEO promotion cutoffs.

First fix the middleaged manager's skill $y \in \mathcal{S}_2$. When the old manager has skill $x < \gamma_t^m(3, 2y, \emptyset)$,

$$\frac{\partial d_t(3x, 2y)}{\partial x} = \rho - \frac{\partial W_t^c(3x)}{\partial x} > 0,$$

where the inequality follows from equation (2). When $x > \gamma_t^m(3, 2y, \emptyset)$,

$$\begin{aligned} \frac{\partial d_t(3x, 2y)}{\partial x} &= \rho - \frac{\partial W_t^c(3x)}{\partial x} - \frac{2L_t^{P4*}(x)(1-\lambda)}{\hat{q}^{old} \cdot 2}, \\ &= \rho - \frac{\partial W_t^c(3x)}{\partial x} - Pr_t^{P4*}(x)(1-\lambda). \end{aligned}$$

When $Pr_t^{P4*}(x)$ is close to one, the result $\frac{\partial d_t(3x, 2y)}{\partial x} > 0$ holds if $\rho - \frac{\partial W_t^c(3s)}{\partial s}$ is sufficiently large,

$$\rho - \frac{\partial W_t^c(a^c, s^c)}{\partial s^c} > 1 - \lambda. \quad (19)$$

Equation (19) is a natural condition: in essence it says that, ceteris paribus, a marginal increase in the CEO's skill contributes more to the firm's expected profits than a marginal increase in the skill of a manager. The result $\frac{\partial d_t(3x, 2y)}{\partial x} > 0$ holds for smaller values of $\rho - \frac{\partial W_t^c(3x)}{\partial x}$ when \hat{q}^{old} becomes

large and reduces $Pr_t^{P4^*}(x)$. That is, promotion is monotone in the old executive's skill as long as the firm is sufficiently worried about losing a talented old manager who is not promoted to CEO.

Now fix the old manager's skill $x \in \mathcal{S}_3$. When the middleaged manager has skill $y < \gamma_t^m(2, 3x, \emptyset)$,

$$\frac{\partial d_t(3x, 2y)}{\partial y} = -\frac{\partial v_t(2y, 3x, \emptyset)}{\partial y} < 0,$$

where the inequality holds since $v_t(2y, 3x, \emptyset)$ strictly increases in y . When $y > \gamma_t^m(2, 3x, \emptyset)$,

$$\begin{aligned} \frac{\partial d_t(3x, 2y)}{\partial y} &= \frac{2L_t^{P3^*}(y)}{(1+\beta)\hat{q}^{mid}} \frac{\partial L_t^{P3^*}(y)}{\partial y} - \frac{\partial v_t(2y, 3x, \emptyset)}{\partial y}, \\ &= 2Pr_t^{P3^*}(y) \frac{\partial L_t^{P3^*}(y)}{\partial y} - \frac{\partial v_t(2y, 3x, \emptyset)}{\partial y}. \end{aligned}$$

Therefore, the result $\frac{\partial d_t(3x, 2y)}{\partial y} < 0$ holds if

$$\begin{aligned} &\rho - \frac{\partial W_t^c(2y)}{\partial y} + \beta[1 - \pi_t] \frac{\partial}{\partial y} \iiint V_{t+1}(3y + \tau, 2s_1, 2s_2) dT(\tau) dF_2(s_1) dF_2(s_2) \\ &> Pr_t^{P3^*}(y) \left[(1 - \lambda) + \beta \frac{\partial}{\partial y} [\mu_{t+1}^{P4}(y) + \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dT(\tau) dF_2(s) - \lambda y] \right]. \end{aligned} \quad (20)$$

Similar to equation (19), condition (20) holds if $\rho - \frac{\partial W_t^c(2y)}{\partial y}$ is sufficiently high, that is, a marginal increase in the middleaged CEO's skill is sufficiently important to the firm's expected profits. The result $\frac{\partial d_t(3x, 2y)}{\partial x} < 0$ holds for smaller values of $\rho - \frac{\partial W_t^c(2y)}{\partial y}$ when \hat{q}^{mid} becomes large and reduces $Pr_t^{P3^*}(y)$. That is, promotion is monotone in the middleaged executive's skill as long as the firm is sufficiently worried about losing a talented middleaged manager who is not promoted to CEO.

Therefore, $d_t(3x, 2y)$ strictly increases in x and strictly decreases in y . Notice that $d_t(3x, 2y)$ is continuous in x and y , and differentiable everywhere but at the layoff cutoffs. ■

Given the skill y of the middleaged manager, the optimal promotion cutoff for the old executive $\gamma_t^c(3, 2y)$ is the smallest skill $x \in \mathcal{S}_3$ such that $d_t(3x, 2y)$ is non-negative. Given the skill x of the old manager, the optimal promotion cutoff for the middleaged executive $\gamma_t^c(2, 3x)$ is the smallest skill $y \in \mathcal{S}_2$ such that $d_t(3x, 2y)$ is non-positive. The firm's optimal strategy is

$$\sigma_t(\emptyset, 3x, 2y) = \begin{cases} (1, 0, 0, 1, W_t^c(3x), W_t^m, W_t^{P3^*}(y)) & \text{if } x > \gamma_t^c(3, 2y) \text{ and } y > \gamma_t^m(2, 3x, \emptyset), \\ (1, 0, 0, 0, W_t^c(3x), W_t^m, W_t^m) & \text{if } x > \gamma_t^c(3, 2y) \text{ and } y < \gamma_t^m(2, 3x, \emptyset), \\ (0, 1, 1, 0, W_t^c(2y), W_t^{P4^*}(x), W_t^m) & \text{if } x < \gamma_t^c(3, 2y) \text{ and } x > \gamma_t^m(3, 2y, \emptyset), \\ (0, 1, 0, 0, W_t^c(2y), W_t^m, W_t^m) & \text{if } x < \gamma_t^c(3, 2y) \text{ and } x < \gamma_t^m(3, 2y, \emptyset). \end{cases}$$

The firm's profit function for partition 4 is $V_t(\emptyset, 3x, 2y) = \max\{V_t(3x, \emptyset, 2y), v_t(2y, 3x, \emptyset)\}$. Lemmas 1-5 and 7 hold. Symmetric results follow from symmetry in V_{t+1} and μ_{t+1}^{P4} . Thus, Lemma 6 holds, concluding the proof for partition 4 at period t .

Partition 5) No CEO, two middleaged managers: Consider $z_t = (\emptyset, 2x, 2y)$ and $x, y \in \mathcal{S}_2$. The firm has four feasible placement offers: (i) promote the manager of division 1 (M1) to CEO and lay off the manager of division 2 (M2); (ii) promote M1 and keep M2; (iii) promote M2 and lay off M1; (iv) promote M2 and keep M1.

Lemma 11 *Conditional on promoting M1 or M2 to CEO, Lemmas 1-7 hold in partition 5.*

Proof. First consider the promotion of M1 to CEO. When the firm lays off M2, the firm's expected profit from placement offer (i) is

$$(i) \quad \rho(l_2 + x) - W_t^c(2x) + 2\eta_t + \beta\bar{v}_{t+1}(x). \quad (21)$$

When the firm offers M2 the managerial position, that is, when the firm makes employment offer $e_t = (1, 0, 0, 1, W_t^c(2x), W_t^m, W_t^2)$, we must compute the probability that M2 stays given the wage offer W_t^2 . If M2 stays, then with probability π_t the CEO leaves the firm, and next period's state will be in partition 3. With probability $1 - \pi_t$ the CEO stays, and next period's state will be in partition 7. Integrating over the possible skills of the young manager and the possible changes in the skill of each middleaged executive, the expected payoff of M2 when he becomes old is a function of his own skill¹⁹

$$\mu_{t+1}^{P5}(y) \equiv \iiint \left[[1 - \pi_t] \mu_{t+1}^2(3x + \tau_1, 2s, 3y + \tau_2) + \pi_t \mu_{t+1}^2(\emptyset, 2s, 3y + \tau_2) \right] dT(\tau_1) dT(\tau_2) dF_2(s).$$

The function increases in y : both $\mu_{t+1}^2(3x + \tau_1, 2s, 3y + \tau_2)$ and $\mu_{t+1}^2(\emptyset, 2s, 3y + \tau_2)$ increase in y since Lemma 5 holds at period $t + 1$. M2's expected payoff from staying at the firm is $W_t^2 + \beta\mu_{t+1}^{P5}(y)$. Therefore, his optimal quit strategy $\varphi_t^2(z_t, e_t, q_t^2)$ is to stay if and only if this payoff exceeds the realized outside wage offer q_t^2 . Therefore, Lemma 4 holds.

Now compute M2's probability of staying. The manager quits with probability one if the inside wage offer W_t^2 falls below $\underline{h}_t(y) \equiv (1 + \beta)\lambda y - \beta\mu_t^{P5}(y)$, and he stays with probability one if the wage offer exceeds $\bar{h}_t(y) \equiv (1 + \beta)(\lambda y + \hat{q}_t^{mid}) - \beta\mu_t^{P5}(y)$. The probability that M2 stays is

$$Pr_t^{P5}(y, W_t^2) = \begin{cases} 1 & \text{if } W_t^2 > \bar{h}_t(y), \\ \frac{W_t^2 + \beta\mu_{t+1}^{P5}(y) - (1 + \beta)\lambda y}{(1 + \beta)\hat{q}_t^{mid}} & \text{if } W_t^2 \in [\underline{h}_t(y), \bar{h}_t(y)], \\ 0 & \text{if } W_t^2 < \underline{h}_t(y). \end{cases}$$

If the firm chooses placement offer (ii) and offers wage W_t^2 , then with probability $1 - Pr_t^{P5}(y, W_t^2)$ the middleaged manager quits and the firm hires a young manager for his position, which yields

¹⁹M2's expected payoff is not a function of the promoted CEO's skill as explained in footnote 17. M2's expected payoff varies with x when the CEO turnover probability π_t is a function of the middleaged CEO skill x .

the same expected profit as placement offer (i). With probability $Pr_t^{P5}(y, W_t^2)$ the middleaged manager stays, produces $l_2 + y$ during this period and becomes old in the next period. Define

$$\tilde{v}_{t+1}(x, y) \equiv \iiint \left\{ [1 - \pi_t] V_{t+1}(3x + \tau_1, 2s, 3y + \tau_2) + \pi_t V_{t+1}(\emptyset, 2s, 3y + \tau_2) \right\} dT(\tau_1) dT(\tau_2) dF_2(s);$$

the firm's expected discounted profit if M2 stays is

$$\rho(l_2 + x) - W_t^c(2x) + l_2 + y - W_t^2 + \eta_t + \beta \tilde{v}_{t+1}(x, y). \quad (22)$$

Subtracting equation (21) from (22), the net discounted profit from keeping M2 is

$$L_t^{P5}(y, W_t^2) \equiv l_2 + y - W_t^2 + \beta \tilde{v}_{t+1}(x, y) - \eta_t - \beta \bar{v}_{t+1}(x).$$

Similar to equation (17), the difference $\tilde{v}_{t+1}(x, y) - \bar{v}_{t+1}(x)$ is independent of x and weakly increasing in y . The expected discounted profit from offer (ii) is

$$(ii) \quad \rho(l_2 + x) - W_t^c(2x) + 2\eta_t + \beta \bar{v}_{t+1}(x) + \max_{W_t^2 \in \mathbb{R}} Pr_t^{P5}(y, W_t^2) L_t^{P5}(y, W_t^2).$$

M2 quits with probability one if $W_t^2 \leq \underline{h}_t(y)$ — such a wage offer is equivalent to firing the manager — so following placement offer (ii) we only consider the relevant wage offers that yield a strictly positive probability of retaining the manager. For any $Pr_t^{P5}(y, W_t^2) > 0$, the firm keeps the manager — chooses offer (ii) over offer (i) — if and only if the net profit from keeping the middleaged manager is positive, i.e., if $L_t^{P5}(y, W_t^2) > 0$. M2 stays with probability one if $W_t^2 \geq \bar{h}_t(y)$. Therefore, the firm never offers a wage above $\bar{h}_t(y)$. When the firm wants to keep M2, it chooses the wage $W_t^2 \in [\underline{h}_t(y), \bar{h}_t(y)]$ that maximizes the net expected discounted profit from placement offer (ii),

$$\frac{W_t^2 + \beta \mu_{t+1}^{P5}(y) - (1 + \beta)\lambda y}{(1 + \beta)\hat{q}^{mid}} \left[l_2 + y - W_t^2 + \beta \tilde{v}_{t+1}(x, y) - \eta_t - \beta \bar{v}_{t+1}(x) \right]. \quad (23)$$

The net expected profit is strictly concave in W_t^2 . Therefore, for each y there exists a unique optimal wage offer $W_t^{P5*}(y)$. If \hat{q}^{mid} is sufficiently large, then it is never profitable for the firm to offer the highest wage $\bar{h}_t(y)$, and the solution is interior. In this case, the first-order condition implies

$$W_t^{P5*}(y) = \frac{1}{2} \left[l_2 + y + \beta \tilde{v}_{t+1}(x, y) - \eta_t - \beta \bar{v}_{t+1}(x) - \beta \mu_{t+1}^{P5}(y) + (1 + \beta)\lambda y \right]. \quad (24)$$

We next show that the firm's equilibrium expected net discounted profit from keeping M2,

$$L_t^{P5*}(y) \equiv \frac{1}{2} \left[l_2 + (1 - \lambda)y + \beta \left[\mu_{t+1}^{P5}(y) + \tilde{v}_{t+1}(x, y) - \lambda y \right] - \eta_t - \beta \bar{v}_{t+1}(x) \right],$$

strictly increases in y . It suffices to show that $\mu_{t+1}^{P5}(y) + \tilde{v}_{t+1}(x, y)$ increases in y at a rate of at least λ . This holds by the argument used in partition 3. Consequently, the firm's optimal layoff

cutoff $\gamma_t^m(2, 2x, \emptyset)$ is the minimum $y \in \mathcal{S}_2$ such that $L_t^{P5^*}(y)$ is non-negative, concluding Lemma 1. When $y > \gamma_t^m(2, 2x, \emptyset)$, the optimal wage offer gives the middleaged manager a payoff from staying in the firm that strictly increases in y ,

$$W_t^{P5^*}(y) + \beta\mu_t^{P5}(y) = L_t^{P5^*}(y) + (1 + \beta)\lambda y,$$

concluding Lemmas 2 and 5. Once more, the firm's optimal wage offer gives M2 a payoff from staying equal to half of the total surplus plus the deterministic component of the outside wage offer. A middleaged manager stays with probability

$$Pr_t^{P5^*}(y) = \begin{cases} \frac{L_t^{P5^*}(y)}{(1+\beta)\hat{q}^{mid}} & \text{if } y > \gamma_t^m(2, 2x, \emptyset), \\ 0 & \text{if } y < \gamma_t^m(2, 2x, \emptyset). \end{cases}$$

Let $\Gamma_t^{P5}(y) = 1$ if $y > \gamma_t^m(2, 2x, \emptyset)$, and $\Gamma_t^{P5}(y) = 0$ otherwise. Conditional on promoting M1, the firm's expected discounted profits are

$$v_t(2x, \emptyset, 2y) = \rho(l_2 + x) - W_t^c(2x) + 2\eta_t + \beta\bar{v}_{t+1}(x) + \frac{\Gamma_t^{P5}(y)}{(1 + \beta)\hat{q}^{mid}} \left[L_t^{P5^*}(y) \right]^2.$$

Profits are independent of y for $y \leq \gamma_t^m(2, 2x, \emptyset)$, and strictly increasing in y for $y > \gamma_t^m(2, 2x, \emptyset)$. In partition 1 we proved that $\rho(l_2 + x) - W_t^c(2x) + 2\eta_t + \beta\bar{v}_{t+1}(x)$ strictly increases in x , hence Lemma 3 holds.

Conditional on promoting M2 to CEO, analogous results follow from symmetry in V_{t+1} and μ_{t+1}^d , concluding the proof. ■

Lemma 12 *In partition 5, the firm promotes the more-skilled manager to CEO.*

Proof. Define the function $d_t(2x, 2y) \equiv v_t(2x, \emptyset, 2y) - v_t(2y, 2x, \emptyset)$. By symmetry in period $t + 1$ profits and payoffs, $d_t(2x, 2x) = 0$. It is optimal to promote M1 if $d_t(2x, 2y) > 0$, and it is optimal to promote M2 if $d_t(2x, 2y) < 0$. We will show that $d_t(2x, 2y)$ strictly increases in x and strictly decreases in y .

Fix $y \in \mathcal{S}_2$. When $x < \gamma_t^m(2, 2y, \emptyset)$, we have $\frac{\partial v_t(2y, 2x, \emptyset)}{\partial x} = 0$ and

$$\frac{\partial d_t(2x, 2y)}{\partial x} = \frac{\partial v_t(2x, \emptyset, 2y)}{\partial x} > 0,$$

where the inequality follows from Lemma 11: profits strictly increase in the CEO's skill. When $x > \gamma_t^m(2, 2y, \emptyset)$,

$$\frac{\partial d_t(2x, 2y)}{\partial x} = \rho - \frac{\partial W_t^c(2x)}{\partial x} + \beta \frac{\partial \bar{v}_{t+1}(x)}{\partial x} - \frac{2L_t^{P5^*}(x)}{(1 + \beta)\hat{q}^{mid}} \frac{\partial L_t^{P5^*}(x)}{\partial x},$$

and we need to show that

$$\rho - \frac{\partial W_t^c(2x)}{\partial x} + \beta \frac{\partial \bar{v}_{t+1}(x)}{\partial x} > Pr_t^{P5^*}(x) \frac{2\partial L_t^{P5^*}(x)}{\partial x}. \quad (25)$$

Similar to equation (20), condition (25) holds if $\rho - \frac{\partial W_t^c(2x)}{\partial x}$ is sufficiently high, that is, a marginal increase in the middleaged CEO's skill is sufficiently important to the firm's expected profits. The result $\frac{\partial d_t(2x, 2y)}{\partial x} > 0$ holds for smaller values of $\rho - \frac{\partial W_t^c(2x)}{\partial x}$ if \hat{q}^{mid} is larger, as this reduces $Pr_t^{P5^*}(x)$. That is, promotion is monotone in the middleaged executive's skill as long as the firm is sufficiently worried about losing a talented middleaged manager who is not promoted to CEO.

Symmetry in continuation payoffs implies $\frac{\partial d_t(2x, 2y)}{\partial y} < 0$, concluding the proof. ■

The firm's optimal strategy is

$$\sigma_t(\emptyset, 2x, 2y) = \begin{cases} (1, 0, 0, 1, W_t^c(2x), W_t^m, W_t^{P5^*}(y)) & \text{if } x > y \text{ and } y > \gamma_t^m(2, 2x, \emptyset), \\ (1, 0, 0, 0, W_t^c(2x), W_t^m, W_t^m) & \text{if } x > y \text{ and } y < \gamma_t^m(2, 2x, \emptyset), \\ (0, 1, 1, 0, W_t^c(2y), W_t^{P5^*}(x), W_t^m) & \text{if } x < y \text{ and } x > \gamma_t^m(2, 2y, \emptyset), \\ (0, 1, 0, 0, W_t^c(2y), W_t^m, W_t^m) & \text{if } x < y \text{ and } x < \gamma_t^m(2, 2y, \emptyset). \end{cases}$$

Expected profits are $V_t(\emptyset, 2x, 2y) = \max\{v_t(2x, \emptyset, 2y), v_t(2y, 2x, \emptyset)\}$. Lemmas 1 to 7 hold, concluding the proof.

Partition 6) No CEO, two old managers: Consider $z_t = (\emptyset, 3x, 3y)$ and $x, y \in \mathcal{S}_3$. The firm has four feasible placement offers: (i) promote the manager of division 1 (M1) to CEO and lay off the manager of division 2 (M2); (ii) promote M1 and keep M2; (iii) promote M2 and lay off M1; (iv) promote M2 and keep M1.

Lemma 13 *Conditional on promoting M1 or M2 to CEO, Lemmas 1-7 hold in partition 6.*

Proof. First consider the promotion of M1 to CEO. When the firm lays off M2, the firm's expected profit from placement offer (i) is

$$(i) \quad \rho(l_3 + x) - W_t^c(3x) + 2\eta_t + \beta\bar{\eta}_{t+1}. \quad (26)$$

When the firm offers M2 the managerial position, that is, when the firm makes employment offer $e_t = (1, 0, 0, 1, W_t^c(3x), W_t^m, W_t^2)$, the probability that M2 stays given the wage offer W_t^2 is $Pr^{old}(y, W_t^2)$, defined in equation (15).

If the firm chooses placement offer (ii) and offers wage W_t^2 , then with probability $1 - Pr^{old}(y, W_t^2)$ the old manager M2 quits and the firm hires a young manager for his position, which yields the

same expected profit as placement offer (i). With probability $Pr^{old}(y, W_t^2)$, M2 stays, produces $l_3 + y$ during this period and then retires. The expected discounted profit if M2 stays is

$$\rho(l_3 + x) - W_t^c(3x) + l_3 + y - W_t^2 + \eta_t + \beta \int V_{t+1}(\emptyset, 2s, \emptyset) dF_2(s). \quad (27)$$

Subtracting equation (26) from (27), the net discounted profit from keeping M2 is

$$L_t^{P6}(y, W_t^2) \equiv l_3 + y - W_t^2 + \beta \int V_{t+1}(\emptyset, 2s, \emptyset) dF_2(s) - \eta_t - \beta \bar{\eta}_{t+1}.$$

The expected discounted profit from offer (ii) is

$$(ii) \quad \rho(l_3 + x) - W_t^c(3x) + 2\eta_t + \beta \bar{\eta}_{t+1} + \max_{W_t^2 \in \mathbb{R}} Pr^{old}(y, W_t^2) L_t^{P6}(y, W_t^2).$$

M2 quits with probability one if $W_t^2 \leq \lambda y$ — such a wage offer is equivalent to firing the manager — so following placement offer (ii) we only consider the relevant wage offers that yield a strictly positive probability of retaining the manager. For any $Pr_t^{P6}(y, W_t^2) > 0$, the firm keeps the manager — chooses offer (ii) over offer (i) — if and only if the net profit from keeping the old manager is positive, i.e., $L_t^{P6}(y, W_t^2) > 0$. M2 stays with probability one if $W_t^2 \geq \lambda y + \hat{q}^{old}$. Therefore, the firm never offers a wage above $\lambda y + \hat{q}^{old}$. When the firm wants to keep M2, it chooses the wage $W_t^2 \in [\lambda y, \lambda y + \hat{q}^{old}]$ that maximizes the net expected discounted profit from placement offer (ii),

$$\frac{W_t^2 - \lambda y}{\hat{q}^{old}} \left[l_3 + y - W_t^2 + \beta \int V_{t+1}(\emptyset, 2s, \emptyset) dF_2(s) - \eta_t - \beta \bar{\eta}_{t+1} \right].$$

The net expected profit is strictly concave in W_t^2 . Therefore, for each y there exists a unique optimal wage offer $W_t^{P6*}(y)$. If \hat{q}^{old} is sufficiently large, then it is never profitable for the firm to offer the highest wage $\lambda y + \hat{q}^{old}$, and the solution is interior. Then the first order-condition implies

$$W_t^{P6*}(y) = \frac{1}{2} \left[l_3 + (1 + \lambda)y + \beta \int V_{t+1}(\emptyset, 2s, \emptyset) dF_2(s) - \eta_t - \beta \bar{\eta}_{t+1} \right]. \quad (28)$$

The firm's equilibrium expected net discounted profit from keeping M2 strictly increases in y ,

$$L_t^{P6*}(y) \equiv \frac{1}{2} \left[l_3 + (1 - \lambda)y + \beta \int V_{t+1}(\emptyset, 2s, \emptyset) dF_2(s) - \eta_t - \beta \bar{\eta}_{t+1} \right]. \quad (29)$$

Consequently, the firms' optimal layoff cutoff $\gamma_t^m(3, 3x, \emptyset)$ is the minimum $y \in \mathcal{S}_3$ such that $L_t^{P6*}(y)$ is non-negative, concluding Lemma 1. When $y > \gamma_t^m(3, 3x, \emptyset)$, the optimal wage offer $W_t^{P6*}(y)$ gives the old manager a payoff from staying in the firm that strictly increases in y , concluding Lemmas 2 and 5. Once more, the firm's optimal wage offer gives the old manager a payoff from staying equal to half of the total surplus plus the deterministic component λy of the outside wage offer. An old manager stays with probability

$$Pr_t^{P6*}(y) = \begin{cases} \frac{L_t^{P6*}(y)}{\hat{q}^{old}} & \text{if } y > \gamma_t^m(3, 3x, \emptyset), \\ 0 & \text{if } y < \gamma_t^m(3, 3x, \emptyset). \end{cases}$$

Let $\Gamma_t^{P6}(y) = 1$ if $y > \gamma_t^m(3, 3x, \emptyset)$, and $\Gamma_t^{P6}(y) = 0$ otherwise. Conditional on promoting M1, the firm's expected profits are

$$v_t(3x, \emptyset, 3y) = \rho(l_3 + x) - W_t^c(3x) + 2\eta_t + \beta\bar{\eta}_{t+1} + \frac{\Gamma_t^{P6}(y)}{\hat{q}^{old}} \left[L_t^{P6*}(y) \right]^2.$$

Equation (2) implies $v_t(3x, \emptyset, 3y)$ strictly increases in x . Profits are independent of y for $y \leq \gamma_t^m(3, 3x, \emptyset)$, and strictly increasing in y for $y > \gamma_t^m(3, 3x, \emptyset)$. Therefore, Lemma 3 hold. It is straightforward to show that Lemma 7 holds.

Conditional on promoting M2 to CEO, symmetric results follow from symmetry in V_{t+1} , concluding the proof. ■

Lemma 14 *In partition 6, the firm promotes the more-skilled manager to CEO.*

Proof. Define the function $d_t(3x, 3y) \equiv v_t(3x, \emptyset, 3y) - v_t(3y, 3x, \emptyset)$. Trivially, $d_t(3x, 3x) = 0$. It is optimal to promote M1 if $d_t(3x, 3y) > 0$, and it is optimal to promote M2 if $d_t(3x, 3y) < 0$. We will show that $d_t(3x, 3y)$ strictly increases in x and strictly decreases in y .

Fix $y \in \mathcal{S}_3$. When $x < \gamma_t^m(3, 3y, \emptyset)$,

$$\frac{\partial d_t(3x, 3y)}{\partial x} = \rho - \frac{\partial W_t^c(3x)}{\partial x} > 0,$$

where the inequality follows from equation (2). When $x > \gamma_t^m(3, 3y, \emptyset)$,

$$\begin{aligned} \frac{\partial d_t(3x, 3y)}{\partial x} &= \rho - \frac{\partial W_t^c(3x)}{\partial x} - \frac{2L_t^{P6*}(x)(1-\lambda)}{\hat{q}^{old} \cdot 2}, \\ &= \rho - \frac{\partial W_t^c(3x)}{\partial x} - Pr_t^{P6*}(x)(1-\lambda). \end{aligned}$$

The result $\frac{\partial d_t(3x, 3y)}{\partial x} > 0$ then follows if $\rho - \frac{\partial W_t^c(3x)}{\partial x}$ is high enough that equation (19) holds, i.e., a marginal increase in the CEO's skill contributes more to the firm's expected profits than a marginal increase in the skill of a manager. The result holds for smaller values of $\rho - \frac{\partial W_t^c(3x)}{\partial x}$ if \hat{q}^{old} is larger, as this reduces $Pr_t^{P6*}(x)$. That is, promotion is monotone in the old executive's skill as long as the firm is sufficiently worried about losing a talented old manager who is not promoted to CEO.

Symmetry implies $\frac{\partial d_t(3x, 3y)}{\partial y} < 0$, concluding the proof. Notice that $d_t(3x, 3y)$ is continuous in x and y , and differentiable everywhere but at the layoff cutoffs ■

The firm's optimal strategy is

$$\sigma_t(\emptyset, 3x, 3y) = \begin{cases} (1, 0, 0, 1, W_t^c(3x), W_t^m, W_t^{P6*}(y)) & \text{if } x > y \text{ and } y > \gamma_t^m(3, 3x, \emptyset), \\ (1, 0, 0, 0, W_t^c(3x), W_t^m, W_t^m) & \text{if } x > y \text{ and } y < \gamma_t^m(3, 3x, \emptyset), \\ (0, 1, 1, 0, W_t^c(3y), W_t^{P6*}(x), W_t^m) & \text{if } x < y \text{ and } x > \gamma_t^m(3, 3y, \emptyset), \\ (0, 1, 0, 0, W_t^c(3y), W_t^m, W_t^m) & \text{if } x < y \text{ and } x < \gamma_t^m(3, 3y, \emptyset). \end{cases}$$

Expected profits are $V_t(\emptyset, 3x, 3y) = \max\{v_t(3x, \emptyset, 3y), v_t(3y, 3x, \emptyset)\}$. M1 is promoted to CEO and receives $W_t^c(3x)$ if $x > y$; he chooses between the firm's wage offer $W_t^{P6^*}(x)$ and the outside wage offer $\lambda x + \xi_t^1$ if $x < y$ and $x > \gamma_t^m(3, 3y, \emptyset)$; and M1 is laid off if $x < y$ and $x < \gamma_t^m(3, 3y, \emptyset)$, which yields an expected payoff $\lambda x + \frac{\hat{q}^{old}}{2}$. Therefore, M1's expected period payoff is

$$\mu_t^1(\emptyset, 3x, 3y) = \begin{cases} W_t^c(3x) & \text{if } x > y, \\ \int \max\{W_t^{P6^*}(x), \lambda x + \xi_t^1\} \frac{1}{\hat{q}^{old}} d\xi_t^1 & \text{if } x < y \text{ and } x > \gamma_t^m(3, 3y, \emptyset), \\ \lambda x + \frac{\hat{q}^{old}}{2} & \text{if } x < y \text{ and } x < \gamma_t^m(3, 3y, \emptyset). \end{cases}$$

In the zero probability event that $x = y$, we assume the firm promotes each manager with equal probability, keeping the manager not promoted if his skill is above the layoff cutoff. M2's expected payoff is symmetric, $\mu_t^2(\emptyset, 3x, 3y) = \mu_t^1(\emptyset, 3y, 3x)$. Lemmas 1 to 7 hold, concluding the proof.

Partition 7) Old CEO, an old manager and a middleaged manager: First consider $z_t = (3z, 3x, 2y)$, $z, x \in \mathcal{S}_3$ and $y \in \mathcal{S}_2$. The firm has four feasible placement offers: (i) lay off both managers; (ii) keep the old manager and lay off the middleaged manager; (iii) keep the middleaged manager and lay off the old manager; (iv) keep both managers.

The expected discounted profit from placement offer (i) is

$$(i) \quad \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta\bar{\eta}_{t+1}.$$

Exploiting the results from partition 6, the expected discounted profit from placement offer (ii),

$$(ii) \quad \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta\bar{\eta}_{t+1} + \frac{1}{\hat{q}^{old}} \left[L_t^{P6^*}(x) \right]^2,$$

is higher than (i) if and only if $x > \gamma_t^m(3, 3z, \emptyset)$, where $\gamma_t^m(3, 3z, \emptyset)$ and $L_t^{P6^*}(x)$ were defined in partition 6. The optimal wage offer is $W_t^{P6^*}(x)$.

Exploiting the results from partition 3, the expected discounted profit from placement offer (iii),

$$(iii) \quad \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta\bar{\eta}_{t+1} + \frac{1}{(1 + \beta)\hat{q}^{mid}} \left[L_t^{P3^*}(y) \right]^2,$$

is higher than (i) if and only if $y > \gamma_t^m(2, 3z, \emptyset)$, where $\gamma_t^m(2, 3z, \emptyset)$ and $L_t^{P3^*}(y)$ were defined in partition 3. The optimal wage offer is $W_t^{P3^*}(y)$.

If the firm makes placement offer (iv), we must compute the probability each manager stays given wage offers (W_t^1, W_t^2) . As in partition 4, the old manager stays with probability $Pr^{old}(x, W_t^1)$. The expected discounted payoff of a middleaged manager who stays in the firm is the following. With probability $Pr^{old}(x, W_t^1)$ the old manager also stays; in this case, the middleaged manager will be the sole inside executive in the next period and will receive the CEO promotion. With

probability $1 - Pr^{old}(x, W_t^1)$ the old manager quits and is replaced by a young manager; in this case, next period's state will be in partition 4, which yields an expected payoff $\mu_{t+1}^{P4}(y)$ to the middleaged manager. Define

$$\begin{aligned}\Delta W_{t+1}(y) &\equiv \int W_{t+1}^c(3y+\tau)dT(\tau) - \mu_{t+1}^{P4}(y), \\ \underline{h}_t(y, W_t^1, x) &\equiv (1 + \beta)\lambda y - \beta\mu_{t+1}^{P4}(y) - Pr^{old}(x, W_t^1)\beta\Delta W_{t+1}(y), \\ \bar{h}_t(y, W_t^1, x) &\equiv (1 + \beta)(\lambda y + \hat{q}^{mid}) - \beta\mu_{t+1}^{P4}(y) - Pr^{old}(x, W_t^1)\beta\Delta W_{t+1}(y).\end{aligned}$$

The middleaged manager stays with probability

$$Pr_t^{P7}(y, W_t^2, x, W_t^1) = \begin{cases} 1 & \text{if } W_t^2 > \bar{h}_t(y, W_t^1, x) \\ \frac{W_t^2 + \beta\mu_{t+1}^{P4}(y) + Pr^{old}(x, W_t^1)\beta\Delta W_{t+1}(y) - (1+\beta)\lambda y}{(1+\beta)\hat{q}^{mid}} & \text{if } W_t^2 \in [\underline{h}_t(y, W_t^1, x), \bar{h}_t(y, W_t^1, x)] \\ 0 & \text{if } W_t^2 < \underline{h}_t(y, W_t^1, x) \end{cases}$$

Notice that $Pr_t^{P7}(y, W_t^2, x, W_t^1)$ is weakly increasing in W_t^1 — if the old manager receives a higher wage offer, he is more likely to stay. Therefore, the middleaged manager is more likely to be promoted to CEO and receive the CEO wage premium $\Delta W_{t+1}(y)$.

The firm's expected discounted profit from placement offer (iv) is

$$\begin{aligned}(iv) \quad &\max_{(W_t^1, W_t^2) \in \mathbb{R}^2} \rho(l_3 + z) - W_t^c(3z) + [1 - Pr^{old}(x, W_t^1)] [1 - Pr_t^{P7}(y, W_t^2, x, W_t^1)] [2\eta_t + \beta\bar{\eta}_{t+1}] \\ &+ Pr^{old}(x, W_t^1) [1 - Pr_t^{P7}(y, W_t^2, x, W_t^1)] \left[l_3 + x - W_t^1 + \eta_t + \beta \int V_{t+1}(\emptyset, \emptyset, 2s) dF_2(s) \right] \\ &+ [1 - Pr^{old}(x, W_t^1)] Pr_t^{P7}(y, W_t^2, x, W_t^1) \left[l_2 + y - W_t^2 + \eta_t + \beta \iint V_{t+1}(\emptyset, 2s, 3y+\tau) dT(\tau) dF_2(s) \right] \\ &+ Pr^{old}(x, W_t^1) Pr_t^{P7}(y, W_t^2, x, W_t^1) \left[l_3 + x - W_t^1 + l_2 + y - W_t^2 + \beta \int V_{t+1}(\emptyset, \emptyset, 3y+\tau) dT(\tau) \right].\end{aligned}$$

Define

$$\begin{aligned}v_{t+1}^{P7}(y) &\equiv \iiint \left[V_{t+1}(\emptyset, \emptyset, 3y+\tau) + V_{t+1}(\emptyset, 2s_1, 2s_2) \right. \\ &\quad \left. - V_{t+1}(\emptyset, 2s_1, 3y+\tau) - V_{t+1}(\emptyset, \emptyset, 2s_2) \right] dT(\tau) dF_2(s_1) dF_2(s_2),\end{aligned}$$

and rewrite

$$\begin{aligned}(iv) \quad &\max_{(W_t^1, W_t^2) \in \mathbb{R}^2} \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta\bar{\eta}_{t+1} + Pr^{old}(x, W_t^1) Pr_t^{P7}(y, W_t^2, x, W_t^1) \beta v_{t+1}^{P7}(y) \\ &+ Pr_t^{P7}(y, W_t^2, x, W_t^1) \left[l_2 + y - W_t^2 + \beta \iint V_{t+1}(\emptyset, 2s, 3y+\tau) dT(\tau) dF_2(s) - \eta_t - \beta\bar{\eta}_{t+1} \right] \\ &+ Pr^{old}(x, W_t^1) \left[l_3 + x - W_t^1 + \beta \int V_{t+1}(\emptyset, \emptyset, 2s) dF_2(s) - \eta_t - \beta\bar{\eta}_{t+1} \right].\end{aligned}$$

Following placement offer (iv), we only consider the relevant wage offers that yield a strictly positive probability of retaining the managers, $W_t^1 > \lambda x$ and $W_t^2 > \underline{h}_t(y, W_t^1, x)$. If \hat{q}^{old} and \hat{q}^{mid} are sufficiently high, these retention probabilities are strictly less than one — solutions are interior and it is routine to compute the Hessian matrix

$$H_y = \begin{bmatrix} -\frac{2}{\hat{q}^{old}} \left[1 - \frac{\beta^2 v_{t+1}^{P7}(y) \Delta W_{t+1}(y)}{(1+\beta) \hat{q}^{mid} \hat{q}^{old}} \right] & \frac{\beta \left[v_{t+1}^{P7}(y) - \Delta W_{t+1}(y) \right]}{(1+\beta) \hat{q}^{mid} \hat{q}^{old}} \\ \frac{\beta \left[v_{t+1}^{P7}(y) - \Delta W_{t+1}(y) \right]}{(1+\beta) \hat{q}^{mid} \hat{q}^{old}} & -\frac{2}{(1+\beta) \hat{q}^{mid}} \end{bmatrix}.$$

Define

$$h_t(y) \equiv 1 - \frac{\beta^2}{4(1+\beta) \hat{q}^{mid} \hat{q}^{old}} \left[v_{t+1}^{P7}(y) + \Delta W_{t+1}(y) \right]^2,$$

and compute the determinant

$$|H_y| = \frac{4}{(1+\beta) \hat{q}^{mid} \hat{q}^{old}} h_t(y).$$

If $(1+\beta) \hat{q}^{mid} \hat{q}^{old}$ is sufficiently large, then $\frac{\partial^2(iv)}{\partial(W_t^1)^2} < 0$, $\frac{\partial^2(iv)}{\partial(W_t^2)^2} < 0$, $h_t(y) \in (0, 1]$, and $|H_t| > 0$. In this case, the Hessian matrix is negative definite and the maximization problem has the following unique solution

$$\begin{aligned} W_t^{P7old*}(x, y) &= \frac{L_t^{P6*}(x)}{h_t(y)} + \frac{L_t^{P3*}(y)}{(1+\beta) \hat{q}^{mid}} \frac{\beta [v_{t+1}^{P7}(y) + \Delta W_{t+1}(y)]}{2h_t(y)} + \lambda x, \\ W_t^{P7mid*}(y, x) &= \frac{L_t^{P3*}(y)}{h_t(y)} \left[1 - \frac{\beta^2 \Delta W_{t+1}(y) [v_{t+1}^{P7}(y) + \Delta W_{t+1}(y)]}{2(1+\beta) \hat{q}^{mid} \hat{q}^{old}} \right] + \frac{L_t^{P6*}(x)}{\hat{q}^{old}} \frac{\beta [v_{t+1}^{P7}(y) - \Delta W_{t+1}(y)]}{2h_t(y)} \\ &\quad - \beta \mu_{t+1}^{P4}(y) + (1+\beta) \lambda y, \end{aligned}$$

where $L_t^{P3*}(y)$ and $L_t^{P6*}(x)$ were defined in equations (12) and (29).

Under placement offer (iv), given optimal wage offers $W_t^{P7old*}(x, y)$ and $W_t^{P7mid*}(y, x)$, the old manager stays with probability

$$Pr_t^{P7old*}(x, y) = \frac{L_t^{P6*}(x)}{\hat{q}^{old} h_t(y)} + \frac{L_t^{P3*}(y)}{(1+\beta) \hat{q}^{mid}} \frac{\beta [v_{t+1}^{P7}(y) + \Delta W_{t+1}(y)]}{2 \hat{q}^{old} h_t(y)},$$

and the middleaged manager stays with probability

$$Pr_t^{P7mid*}(y, x) = \frac{L_t^{P3*}(y)}{(1+\beta) \hat{q}^{mid} h_t(y)} + \frac{L_t^{P6*}(x)}{\hat{q}^{old}} \frac{\beta [v_{t+1}^{P7}(y) + \Delta W_{t+1}(y)]}{2(1+\beta) \hat{q}^{mid} h_t(y)}.$$

The firm's expected discounted profit from placement offer (iv) is

$$\begin{aligned} (iv) \quad \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta \bar{\eta}_{t+1} &+ \frac{[L_t^{P6*}(x)]^2}{\hat{q}^{old} h_t(y)} + \frac{[L_t^{P3*}(y)]^2}{(1+\beta) \hat{q}^{mid} h_t(y)} \\ &+ \frac{L_t^{P6*}(x)}{\hat{q}^{old}} \frac{L_t^{P3*}(y)}{(1+\beta) \hat{q}^{mid}} \frac{\beta [v_{t+1}^{P7}(y) + \Delta W_{t+1}(y)]}{h_t(y)}. \end{aligned}$$

Lemma 15 *In partition 7, the firm's placement offer is monotone in the skill of each manager.*

Proof. The firm's expected discounted profit from placement offer (i) is independent of x and y . We have shown that the firm's profit from placement offer (ii) strictly increases in x and is independent of y , while (iii) strictly increases in y and is independent of x . It remains to show that when placement offer (iv) yields the highest expected profit, then the profit from placement offer (iv) strictly increases in x and y .

Fix the skill of the middleaged manager, $y \in \mathcal{S}_2$. Let $W_t^{P7old*}(x, y)$ and $W_t^{P7mid*}(y, x)$ be the optimal wage offers following placement offer (iv) when skills are x and y . Increase the old manager's skill to $x' = x + \epsilon$, $\epsilon > 0$, and consider (non-optimal) wage offers $W_t^{1'} = W_t^{P7old*}(x, y) + \lambda\epsilon$ and $W_t^{2'} = W_t^{P7mid*}(y, x)$. Both managers stay with the same probability as before, and profits increase by $Pr_t^{P7old*}(x, y)(1 - \lambda)\epsilon > 0$. Therefore, profits are even higher when the firm chooses wages optimally: the profit from placement offer (iv) strictly increases in x .

An analogous argument holds for the middleaged manager. Fix the skill x of the old manager, and let $W_t^{P7old*}(x, y)$ and $W_t^{P7mid*}(y, x)$ be the optimal wage offers following placement offer (iv) when skills are x and y . Increase the middleaged manager's skill to $y' = y + \epsilon$, $\epsilon > 0$, and consider (non-optimal) wage offers $W_t^{1'} = W_t^{P7old*}(x, y)$ and

$$\begin{aligned} W_t^{2'} &= W_t^{P7mid*}(y, x) + (1 + \beta)\lambda y - \beta Pr_t^{P7old*}(x, y) \int [W_{t+1}^c(3y' + \tau) - W_{t+1}^c(3y + \tau)] dT(\tau) \\ &\quad - \beta [1 - Pr_t^{P7old*}(x, y)] [\mu_{t+1}^{P4}(y') - \mu_{t+1}^{P4}(y)]. \end{aligned}$$

The new wage offers keep managers' quit probabilities the same, while profits increase by

$$\begin{aligned} &Pr_t^{P7mid*}(y, x) \left\{ (1 - \lambda)\epsilon + \beta Pr_t^{P7old*}(x, y) [W_{t+1}^c(3y') - W_{t+1}^c(3y) + (\rho - 1)\epsilon] \right. \\ &+ \left. \beta [1 - Pr_t^{P7old*}(x, y)] [\mu_{t+1}^{P4}(y') - \mu_{t+1}^{P4}(y) + \int [V_{t+1}(\emptyset, 2s, 3y') - V_{t+1}(\emptyset, 2s, 3y)] dF(s) - \lambda\epsilon] \right\}. \end{aligned}$$

Recall that in partition 3 we proved that $\mu_{t+1}^{P4}(y) + \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dT(\tau) dF_2(s) - \lambda y$ increases in y . Consequently, profits are even higher when the firm chooses wages optimally: the profit from placement offer (iv) strictly increases in y .

Therefore, the firm's optimal placement offers are as follows. Given the middleaged manager's skill y , the firm's layoff cutoff $\gamma_t^m(3, 3z, 2y)$ for the old manager is the minimum $x \in \mathcal{S}$ such that expected profit from *either* placement offer (ii) or (iv) is greater than the expected profit from *both* offers (i) and (iii). Given the old manager's skill x , the firm's layoff cutoff $\gamma_t^m(2, 3z, 3x)$ for the

midleaged manager is the minimum $y \in \mathcal{S}$ such that the expected profit from *either* placement offer (iii) or (iv) is greater than expected profit from *both* offers (i) and (ii). Symmetric results hold when we exchange managers' divisions, since continuation payoffs are symmetric. ■

Lemma 16 *In partition 7, a manager's expected discounted lifetime payoff from staying in the firm is increasing in his own skill.*

Proof. If \hat{q}^{mid} and \hat{q}^{old} are sufficiently large, then it is routine to show that, for each fixed placement offer (i)-(iv), the firm's optimal wage offers give to each manager an expected lifetime payoff from staying in the firm that increases in the manager's own skill. Moreover, placement offers are monotone in a manager's own skill. It remains to show that each manager's payoff is monotone in his own skill when the firm switches between optimal placement offers.

The firm's expected discounted profits from placement offers (i) and (iii) are independent of the old manager's skill (he is laid off). The firm's profits from placement offers (ii) and (iv) and its wage offer to the old manager are increasing in his skill x . Therefore, as the old manager's skill increases, either the firm switches from offers (i) or (iii) to (ii) or (iv), which strictly increases the old manager's payoff from staying in the firm, or the firm switches between offers (ii) and (iv).

Therefore, it only remains to show that the old manager's payoff does not decrease when the firm optimally switches between placement offers (ii) and (iv). Since the firm's expected profit from each placement offer increases continuously in the old manager's skill, the firm only switches between optimal offers (ii) and (iv) when they have the same expected profit — which only happens when offer (iv) keeps the midleaged manager with probability zero. In this case, offer (iv) gives the old manager the same expected payoff as offer (ii), and the change does not decrease the manager's expected payoff.

The same argument extends the monotonicity result to the midleaged manager's payoff. ■

Partition 8) Old CEO, two midleaged managers: Consider $z_t = (3z, 2x, 2y)$, $z \in \mathcal{S}_3$ and $x, y \in \mathcal{S}_2$. The firm has four feasible placement offers: (i) lay off both managers; (ii) keep the manager of division 1 (M1) and lay off the manager of division 2 (M2); (iii) keep M2 and lay off M1; (iv) keep both managers.

The expected discounted profit from placement offer (i) is

$$(i) \quad \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta\bar{\eta}_{t+1}.$$

Exploiting the results from partition 3, the expected discounted profit from placement offer (ii),

$$(ii) \quad \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta\bar{\eta}_{t+1} + \frac{1}{(1 + \beta)\hat{q}^{mid}} \left[L_t^{P3*}(x) \right]^2,$$

is higher than (i) if and only if $x > \gamma_t^m(2, 3z, \emptyset)$, where $\gamma_t^m(2, 3z, \emptyset)$ and $L_t^{P3*}(x)$ were defined in partition 3. The optimal wage offer is $W_t^{P3*}(x)$. The expected profit from placement offer (iii),

$$(iii) \quad \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta\bar{\eta}_{t+1} + \frac{1}{(1 + \beta)\hat{q}^{mid}} \left[L_t^{P3*}(y) \right]^2,$$

is higher than (i) if and only if $y > \gamma_t^m(2, 3z, \emptyset)$. The firm's expected profit from placement offer (ii) is higher than (iii) if and only if $x > y$.

Now we compute managers' expected payoffs from staying in the firm and their equilibrium quit probabilities when the firm chooses placement offer (iv) and wages (W_t^1, W_t^2) . If the manager of division d stays and the other manager quits, next period's state will be in partition 4, so that manager d 's expected discounted payoff is $W_t^d + \beta\mu_{t+1}^{P4}(s^d)$. If both managers stay, next period's state will be in partition 6, so that manager d 's expected discounted payoff is $W_t^d + \beta \iint \mu_{t+1}^1(\emptyset, 3s^d + \tau_1, 3s^{d-} + \tau_2) dT(\tau_1) dT(\tau_2)$ — recall that $\mu_{t+1}^1(\emptyset, 3s^d, 3s^{d-}) = \mu_{t+1}^2(\emptyset, 3s^{-d}, 3s^d)$. Let $\Delta W_{t+1}(s^d, s^{-d})$ be the change in the next period's expected payoff of a middleaged manager with skill s^d from staying in the firm if the other middleaged manager with skill s^{-d} also stays,

$$\Delta W_{t+1}(s^d, s^{-d}) = \iint \mu_{t+1}^1(\emptyset, 3s^d + \tau_1, 3s^{d-} + \tau_2) dT(\tau_1) dT(\tau_2) - \mu_{t+1}^{P4}(s^d).$$

Given probability $Pr_t^{P8}(y, W_t^2, x, W_t^1)$ that M2 stays, M1 stays if and only if

$$(1 + \beta)(\xi_t^1 + \lambda x) < W_t^1 + \beta\mu_{t+1}^{P4}(x) + Pr_t^{P8}(y, W_t^2, x, W_t^1)\beta\Delta W_{t+1}(x, y).$$

A similar inequality characterizes the quit decision of M2. Therefore, we can define the managers' equilibrium strategies as cutoffs (ξ_t^{1*}, ξ_t^{2*}) on the random component of the outside wage offer. These cutoffs imply probabilities of staying in the firm of $Pr_t^{P8}(x, W_t^1, y, W_t^2) = \frac{\xi_t^{1*}}{\hat{q}^{mid}}$ for M1, and $Pr_t^{P8}(y, W_t^2, x, W_t^1) = \frac{\xi_t^{2*}}{\hat{q}^{mid}}$ for M2.

Following placement offer (iv), we only consider the relevant wage offers that yield a strictly positive probability of retaining the managers. If \hat{q}^{mid} is sufficiently high, these retention probabilities are strictly less than one — solutions are interior — and the system of equations

$$\begin{aligned} (1 + \beta)(\xi_t^{1*} + \lambda x) &= W_t^1 + \beta\mu_{t+1}^{P4}(x) + \frac{\xi_t^{2*}}{\hat{q}^{mid}}\beta\Delta W_{t+1}(x, y), \\ (1 + \beta)(\xi_t^{2*} + \lambda y) &= W_t^2 + \beta\mu_{t+1}^{P4}(y) + \frac{\xi_t^{1*}}{\hat{q}^{mid}}\beta\Delta W_{t+1}(y, x), \end{aligned}$$

has a unique solution

$$\begin{aligned}\xi_t^{1*} &= \left\{ \frac{W_t^1 + \beta\mu_{t+1}^{P4}(x)}{1 + \beta} - \lambda x + \beta \frac{\Delta W_{t+1}(x, y)}{(1 + \beta)\hat{q}^{mid}} \left[\frac{W_t^2 + \beta\mu_{t+1}^{P4}(y)}{1 + \beta} - \lambda y \right] \right\} / h_t(x, y), \\ \xi_t^{2*} &= \left\{ \frac{W_t^2 + \beta\mu_{t+1}^{P4}(y)}{1 + \beta} - \lambda y + \beta \frac{\Delta W_{t+1}(y, x)}{(1 + \beta)\hat{q}^{mid}} \left[\frac{W_t^1 + \beta\mu_{t+1}^{P4}(x)}{1 + \beta} - \lambda x \right] \right\} / h_t(x, y),\end{aligned}$$

where

$$h_t(x, y) \equiv 1 - \frac{\beta^2 \Delta W_{t+1}(x, y) \Delta W_{t+1}(y, x)}{[(1 + \beta)\hat{q}^{mid}]^2}.$$

The firm's expected discounted profit from placement offer (iv) is

$$\begin{aligned}(iv) \quad & \max_{(W_t^1, W_t^2) \in \mathbb{R}^2} \rho(l_3 + z) - W_t^c(3z) + [1 - Pr_t^{P8}(x, W_t^1, y, W_t^2)] [1 - Pr_t^{P8}(y, W_t^2, x, W_t^1)] [2\eta_t + \beta\bar{\eta}_{t+1}] \\ & + Pr_t^{P8}(x, W_t^1, y, W_t^2) [1 - Pr_t^{P8}(y, W_t^2, x, W_t^1)] \left[l_2 + x - W_t^1 + \eta_t + \beta \iint V_{t+1}(\emptyset, 3x + \tau, 2s) dT(\tau) dF_2(s) \right] \\ & + [1 - Pr_t^{P8}(x, W_t^1, y, W_t^2)] Pr_t^{P8}(y, W_t^2, x, W_t^1) \left[l_2 + y - W_t^2 + \eta_t + \beta \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dT(\tau) dF_2(s) \right] \\ & + Pr_t^{P8}(x, W_t^1, y, W_t^2) Pr_t^{P8}(y, W_t^2, x, W_t^1) \left[l_2 + x - W_t^1 + l_2 + y - W_t^2 + \beta \iint V_{t+1}(\emptyset, 3x + \tau_1, 3y + \tau_2) dT(\tau_1) dT(\tau_2) \right].\end{aligned}$$

Define

$$\begin{aligned}v_{t+1}^{P8}(x, y) &\equiv \iiint \left[V_{t+1}(\emptyset, 3x + \tau_1, 3y + \tau_2) + V_{t+1}(\emptyset, 2s_1, 2s_2) \right. \\ &\quad \left. - V_{t+1}(\emptyset, 2s_1, 3y + \tau_2) - V_{t+1}(\emptyset, 3x + \tau_1, 2s_2) \right] dT(\tau_1) dT(\tau_2) dF_2(s_1) dF_2(s_2),\end{aligned}$$

and rewrite

$$\begin{aligned}(iv) \quad & \max_{(W_t^1, W_t^2) \in \mathbb{R}^2} \rho(l_3 + z) - W_t^c(3z) + 2\eta_t + \beta\bar{\eta}_{t+1} + Pr_t^{P8}(x, W_t^1, y, W_t^2) Pr_t^{P8}(y, W_t^2, x, W_t^1) \beta v_{t+1}^{P8}(x, y) \\ & + Pr_t^{P8}(y, W_t^2, x, W_t^1) \left[l_2 + y - W_t^2 + \beta \iint V_{t+1}(\emptyset, 2s, 3y + \tau) dT(\tau) dF_2(s) - \eta_t - \beta\bar{\eta}_{t+1} \right] \\ & + Pr_t^{P8}(x, W_t^1, y, W_t^2) \left[l_2 + x - W_t^1 + \beta \iint V_{t+1}(\emptyset, 3x + \tau, 2s) dT(\tau) dF_2(s) - \eta_t - \beta\bar{\eta}_{t+1} \right].\end{aligned}$$

From the derivatives of (iv) with respect to wages, define

$$\begin{aligned}M_1 &\equiv \left[1 - \frac{\beta^2 \Delta W_{t+1}(y, x) v_{t+1}^{P8}(x, y)}{[(1 + \beta)\hat{q}^{mid}]^2 h_t(x, y)} \right], \\ M_2 &\equiv \left[1 - \frac{\beta^2 \Delta W_{t+1}(x, y) v_{t+1}^{P8}(x, y)}{[(1 + \beta)\hat{q}^{mid}]^2 h_t(x, y)} \right], \\ M_{12} &\equiv \frac{1}{2(1 + \beta)\hat{q}^{mid} h_t(x, y)} \beta \left[2v_{t+1}^{P8}(x, y) - h_t(x, y) \left[v_{t+1}^{P8}(x, y) + \Delta W_{t+1}(x, y) + \Delta W_{t+1}(y, x) \right] \right], \\ M_h &\equiv M_1 M_2 - [M_{12}]^2,\end{aligned}$$

and compute the Hessian matrix

$$H = \begin{bmatrix} \frac{-2M_1}{(1+\beta)\hat{q}^{mid}h_t(x,y)} & \frac{2M_{12}}{(1+\beta)\hat{q}^{mid}h_t(x,y)} \\ \frac{2M_{12}}{(1+\beta)\hat{q}^{mid}h_t(x,y)} & \frac{-2M_2}{(1+\beta)\hat{q}^{mid}h_t(x,y)} \end{bmatrix}$$

If \hat{q}^{mid} is sufficiently large, then $\frac{\partial^2(iv)}{\partial(W_t^1)^2} < 0$, $\frac{\partial^2(iv)}{\partial(W_t^2)^2} < 0$, and

$$|H| = \frac{1}{[(1+\beta)\hat{q}^{mid}h_t(x,y)]^2} \left[4 - \frac{\beta^2 [v_{t+1}^{P8}(x,y) + \Delta W_{t+1}(x,y) + \Delta W_{t+1}(y,x)]^2}{[(1+\beta)\hat{q}^{mid}]^2} \right] > 0.$$

The Hessian matrix is negative definite and the maximization problem has a unique solution, given by the following wage offer $W_t^{P8*}(x,y)$ to M1 and $W_t^{P8*}(y,x)$ to M2,

$$\begin{aligned} W_t^{P8*}(x,y) &= \frac{L_t^{P3*}(x)}{M_h} \left[M_1 + \frac{\beta \Delta W_{t+1}(x,y)}{(1+\beta)\hat{q}^{mid}} M_{12} \right] + \frac{L_t^{P3*}(y)}{M_h} \left[M_{12} + \frac{\beta \Delta W_{t+1}(y,x)}{(1+\beta)\hat{q}^{mid}} M_1 \right] \\ &\quad - \beta \mu_{t+1}^{P4}(x) + (1+\beta)\lambda x, \\ W_t^{P8*}(y,x) &= \frac{L_t^{P3*}(y)}{M_h} \left[M_2 + \frac{\beta \Delta W_{t+1}(y,x)}{(1+\beta)\hat{q}^{mid}} M_{12} \right] + \frac{L_t^{P3*}(x)}{M_h} \left[M_{12} + \frac{\beta \Delta W_{t+1}(x,y)}{(1+\beta)\hat{q}^{mid}} M_2 \right] \\ &\quad - \beta \mu_{t+1}^{P4}(y) + (1+\beta)\lambda y, \end{aligned}$$

where $L_t^{P3*}(s^d)$ was defined in (12). It is routine to use $W_t^{P8*}(x,y)$ and $W_t^{P8*}(y,x)$ to compute the firm's equilibrium expected discounted profit from placement offer (iv).

Lemma 17 *In partition 8, the firm's placement offer is monotone in the skill of each manager.*

Proof. The proof is analogous to the proof of Lemma 15. Let $W_t^{P8*}(x,y)$ and $W_t^{P8*}(y,x)$ be the optimal wage offers when managers have skills s^d and s^{-d} . Now increase manager d 's skill to $s^{d'} > s^d$ and consider (non-optimal) wage offers $W_t^{1'}$ and $W_t^{2'}$ such that managers stay with the same probabilities as before. One can show that this new wage offer increases the firm's expected discounted profits, which increase by even more when the firm chooses wages optimally.

The firm's optimal placement offers are as follows. Given M2's skill y , the firm's layoff cutoff $\gamma_t^m(2, 3z, 2y)$ for M1 is the minimum $x \in \mathcal{S}_2$ such that expected profit from *either* placement offer (ii) or (iv) is greater than the maximum expected profit from offers (i) and (iii). Given M1's skill x , the firm's layoff cutoff $\gamma_t^m(2, 3z, 2x)$ for M2 is the minimum $y \in \mathcal{S}_2$ such that the expected profit from *either* placement offer (iii) or (iv) is greater than the maximum expected profit from offers (i) and (ii). Symmetric results hold when we exchange managers' divisions, since continuation payoffs are symmetric. ■

Lemma 18 *In partition 8, a manager's expected discounted lifetime payoff from staying in the firm is increasing in his own skill.*

Proof. The proof is analogous to the proof of Lemma 16. ■

A.5 Proof of Proposition 1

Monotonicity is not a necessary condition for Proposition 1. However, the proof for the non-monotonic equilibrium case requires solving the model recursively, and defining the corresponding strategies and expected payoffs. Therefore, we will consider parameters such that Theorem 2 holds and use the notation and functions defined in the Theorem's proof.

Fix any period $t + 1 < N - 1$ and solve the model recursively up to period $t + 1$, following the proof of Theorem 2. At period $t + 1$, it is routine to show that the constant component of the CEO wage, \overline{W}_{t+1}^c , does not affect managers' quit decisions or the firm's employment offer. The constant enters manager's expected payoff $\mu_{t+1}^{P4}(s)$ and the firm's expected discounted profit $V_{t+1}(\cdot)$, affecting decisions at period t ; the constant does not affect differences in the firm's profit function, $V_{t+1}(z_t) - V_{t+1}(z'_t)$. At period t , we will consider \overline{W}_{t+1}^c sufficiently high that the expected next period payoff of a middleaged manager competing for the CEO promotion is sufficiently high²⁰.

R1. Consider an old manager with skill s^{old} and a talented middleaged manager with skill s^{mid} , who has a sufficiently high chance to win internal promotion to CEO when facing a young manager with unknown talent. The firm promotes the old manager to CEO unless s^{mid} is sufficiently greater than s^{old} , even in the absence of learn-by-doing and when middleaged managers tend to receive better outside wage offers, $\hat{q}^{old} < \hat{q}^{mid}$.

The sole relevant partition is partition 4, where the firm must choose whether to promote to CEO a middleaged manager or an old manager. We will show that if both managers have the same high skill x , then the firm strictly prefers to promote the old manager to CEO. That is, following the results from Lemma 10, we need to show that $d_t(3x, 2x) > 0$. By definition,

$$\begin{aligned} d_t(3x, 2x) &= V_t(3x, \emptyset, 2x) - v_t(2x, 3x, \emptyset) \\ &= \rho(l_3 + x) - W_t^c(3x) + 2\eta_t + \beta\bar{\eta}_{t+1} + \frac{\Gamma_t^{P3}(x)}{(1 + \beta)\hat{q}^{mid}} \left[L_t^{P3*}(x) \right]^2 \\ &\quad - \rho(l_2 + x) + W_t^c(2x) - 2\eta_t - \beta\bar{v}_{t+1}(x) - \frac{\Gamma_t^{P4}(x)}{\hat{q}^{old}} \left[L_t^{P4*}(x) \right]^2. \end{aligned}$$

²⁰We exploit the recursive nature of our model to adjust \overline{W}_{t+1}^c . We find numerically that the results hold when \overline{W}^c is constant over time.

After some algebra,

$$\begin{aligned}
d_t(3x, 2x) &= \rho(l_3 - l_2) - [W_t^c(3x) - W_t^c(2x)] \\
&+ \beta(1 - \pi_t) \iiint \left[V_{t+1}(\emptyset, 2s_1, 2s_2) - V_{t+1}(3x + \tau, 2s_1, 2s_2) \right] dT(\tau) dF_2(s_1) dF_2(s_2) \\
&+ \frac{\Gamma_t^{P3}(x)}{(1 + \beta)\hat{q}^{mid}} \left[L_t^{P3^*}(x) \right]^2 - \frac{\Gamma_t^{P4}(x)}{\hat{q}^{old}} \left[L_t^{P4^*}(x) \right]^2.
\end{aligned}$$

Recall that $\Gamma_t^{P3}(x) = 1$ if $L_t^{P3^*}(x) > 0$. Therefore, it is sufficient to show that $L_t^{P3^*}(x)$ is sufficiently high so to offset the negative terms in $d_t(3x, 2x)$. By definition,

$$\begin{aligned}
L_t^{P3^*}(x) &= \frac{1}{2} \left[l_2 + (1 - \lambda)x - \eta_t + \beta[\mu_{t+1}^{P4}(x) - \lambda x] \right. \\
&+ \left. \beta \iiint \left[V_{t+1}(\emptyset, 2s_1, 3x + \tau) - V_{t+1}(\emptyset, 2s_1, 2s_2) \right] dT(\tau) dF_2(s_1) dF_2(s_2) \right], \\
L_t^{P4^*}(x) &= \frac{1}{2} \left[l_3 + (1 - \lambda)x - \eta_t + \beta[\underline{v}_{t+1}(x) - \bar{v}_{t+1}(x)] \right].
\end{aligned}$$

When the middleaged manager has a strictly positive probability of being promoted to CEO in the following period, his expected payoff $\mu_{t+1}^{P4}(x)$ includes next period's CEO wage $W_{t+1}^c(3x + \tau) = \bar{W}_{t+1}^c + w_{t+1}^c(3x + \tau)$. Notice that all terms for the firm's next period expected profit appear in $d_t(3x, 2x)$ as differences, not as levels, hence \bar{W}_{t+1}^c drops out. Therefore, the result holds if \bar{W}_{t+1}^c and the probability of promotion are sufficiently high. That is, if the probability of promotion and the CEO wage are sufficiently high, then the firm can profitably exploit this by keeping as manager the middleaged executive, who is competing for the CEO promotion and considers the CEO wage in his expected payoff. The old manager no longer considers the CEO wage in his expected payoff since he will retire before having another chance of internal promotion. Notice that as the CEO wage prize \bar{W}_{t+1}^c becomes larger, the result extends to middleaged managers with smaller probability of future promotion, i.e., middleaged managers with lower skills.

R2. *A talented middleaged manager is less likely to leave the firm if the CEO is old.*

Consider a middleaged manager in division 1 with skill $y \in \mathcal{S}_2$. Under partition 3 the CEO is old, and the middleaged manager's expected discounted lifetime payoff from staying in the firm is

$$\begin{aligned}
W_t^{P3^*}(y) + \beta\mu_{t+1}^{P4}(y) &= \frac{1}{2} \left[l_2 + y + \beta \iint V_{t+1}(\emptyset, 3y + \tau, 2s) dT(\tau) dF_2(s) \right. \\
&- \left. \eta_t - \beta\bar{\eta}_{t+1} + \beta\mu_{t+1}^{P4}(y) + (1 + \beta)\lambda y \right].
\end{aligned}$$

The talented middleaged manager considers his chance of future promotion to CEO and the CEO wage $W_{t+1}^c(3y + \tau) = \bar{W}_{t+1}^c + w_{t+1}^c(3y + \tau)$ when computing his expected payoff $\mu_{t+1}^{P4}(y)$.

Under partition 5 when the middleaged manager M2 with skill x is promoted to CEO, the middleaged manager M1's expected discounted lifetime payoff from staying in the firm is

$$\begin{aligned} & W_t^{P5*}(y) + \beta\mu_{t+1}^{P5}(y) \\ = & \frac{1}{2} \left[l_2 + y + \beta \iiint \left[[1 - \pi_t] \mu_{t+1}^1(3x + \tau_1, 3y + \tau_2, 2s) + \pi_t \mu_{t+1}^{P4}(y) \right] dT(\tau_1) dT(\tau_2) dF_2(s) \right. \\ & \left. + \beta \tilde{v}_{t+1}(x, y) - \eta_t - \beta \bar{v}_{t+1}(x) + (1 + \beta) \lambda y \right]. \end{aligned}$$

With probability $1 - \pi_t > 0$ the CEO stays and M1 will not be promoted — the CEO wage does not enter the manager's next period expected payoff $\mu_{t+1}^1(3x + \tau_1, 3y + \tau_2, 2s)$.

We need to show that $W_t^{P3*}(y) + \beta\mu_{t+1}^{P4}(y) > W_t^{P5*}(y) + \beta\mu_{t+1}^{P5}(y)$, that is,

$$\begin{aligned} & \iint V_{t+1}(\emptyset, 3y + \tau, 2s) dT(\tau) dF_2(s) - \bar{\eta}_{t+1} - \tilde{v}_{t+1}(x, y) + \bar{v}_{t+1}(x) \\ + & \left[1 - \pi_t \right] \left[\mu_{t+1}^{P4}(y) - \iiint \mu_{t+1}^1(3x + \tau_1, 3y + \tau_2, 2s) dT(\tau_1) dT(\tau_2) dF_2(s) \right] > 0. \end{aligned}$$

Notice that all terms for the firm's next period expected profit appear as differences, not as levels, hence \bar{W}_{t+1}^c drops out. Therefore, the result holds if \bar{W}_{t+1}^c is high enough that the payoff difference $\mu_{t+1}^{P4}(y) - \iiint \mu_{t+1}^1(3x + \tau_1, 3y + \tau_2, 2s) dT(\tau_1) dT(\tau_2) dF_2(s)$ is sufficiently high. That is, the talented middleaged manager has a strictly higher probability of future promotion if the current CEO is old than if the CEO is middleaged. Therefore, if the CEO wage is sufficiently high, then the middleaged manager is more likely to stay if the CEO is old.

If the CEO wage \bar{W}_{t+1}^c is sufficiently high, then the result extends to a majority of middleaged managers — those with sufficiently high probabilities of promotion. Numerically we find that the result holds for a majority of manager even when \bar{W}_{t+1}^c is low, i.e., the notion of “sufficiently talented” holds for most managers.

R3. *An old manager is more likely to leave the firm if the CEO is old.*

Consider an old manager in division 1 with skill $y \in \mathcal{S}_3$. Under partition 6 when the old manager M2 is promoted to CEO, the old manager M1's payoff from staying in the firm is

$$W_t^{P6*}(y) = \frac{1}{2} \left[l_3 + (1 + \lambda)y + \beta \int V_{t+1}(\emptyset, \emptyset, 2s) dF_2(s) - \eta_t - \beta \bar{\eta}_{t+1} \right].$$

Under partition 4 when the middleaged manager M2 with skill x is promoted to CEO, the old manager M1's payoff from staying in the firm is

$$W_t^{P4*}(y) = \frac{1}{2} \left[l_3 + (1 + \lambda)y + \beta v_{t+1}(x) - \eta_t - \beta \bar{v}_{t+1}(x) \right].$$

We need to show that $W_t^{P4*}(y) - W_t^{P6*}(y) > 0$, that is,

$$\begin{aligned} & \iiint [V_{t+1}(3x+\tau, \emptyset, 2s_2) - V_{t+1}(3x+\tau, 2s_1, 2s_2) \\ & + V_{t+1}(\emptyset, 2s_1, 2s_2) - V_{t+1}(\emptyset, \emptyset, 2s_2)] dT(\tau) dF_2(s_1) dF_2(s_2) > 0. \end{aligned}$$

Using the results from partitions 3 and 8, the integral of $V_{t+1}(3x+\tau, \emptyset, 2s_2) - V_{t+1}(3x+\tau, 2s_1, 2s_2)$ is negative — it measures the option value of having one extra manager of unknown talent when the CEO position is not vacant,

$$\begin{aligned} V_{t+1}(3x+\tau, \emptyset, 2s_2) - V_{t+1}(3x+\tau, 2s_1, 2s_2) &= Pr_t^{P3*}(s_2) L_t^{P3*}(s_2) - Pr_t^{P8*}(s_1, s_2) L_t^{P8*}(s_1, s_2) \\ &- Pr_t^{P8*}(s_2, s_1) L_t^{P8*}(s_2, s_1) - Pr_t^{P8*}(s_1, s_2) Pr_t^{P8*}(s_2, s_1) v_{t+1}^{P8}(s_1, s_2). \end{aligned}$$

Notice that the CEO's contribution to profits $\rho(l_3 + x + \tau) - W_t^c(3x + \tau)$ drops out. Let $s^{max} = \max\{s_1, s_2\}$, $s^{min} = \min\{s_1, s_2\}$. Using the results from partitions 1 and 5, the integral of $V_{t+1}(\emptyset, 2s_1, 2s_2) - V_{t+1}(\emptyset, \emptyset, 2s_2)$ is positive — it measures the option value of having one extra manager of unknown talent when the CEO position is vacant,

$$\begin{aligned} V_{t+1}(\emptyset, 2s_1, 2s_2) - V_{t+1}(\emptyset, \emptyset, 2s_2) &= \rho(s^{max} - s_2) - [W_t^c(s^{max}) - W_t^c(s_2)] \\ &+ \beta [\bar{v}_{t+1}(s^{max}) - \bar{v}_{t+1}(s^2)] + Pr_t^{P5*}(s^{min}) L_t^{P5*}(s^{min}). \end{aligned}$$

If the CEO skill's marginal contribution to profit $\rho - \frac{\partial W_t^c(a^c, s^c)}{\partial s^c}$ is sufficiently large, then the positive term $V_{t+1}(\emptyset, 2s_1, 2s_2) - V_{t+1}(\emptyset, \emptyset, 2s_2)$ dominates the negative term $V_{t+1}(3x + \tau, \emptyset, 2s_2) - V_{t+1}(3x + \tau, 2s_1, 2s_2)$: the option value of a second manager is higher when the CEO position is vacant, because he is a potential candidate for the promotion. Moreover, the negative term $V_{t+1}(3x + \tau, \emptyset, 2s_2) - V_{t+1}(3x + \tau, 2s_1, 2s_2)$ goes to zero as the probability that a middleaged manager who is not promoted to CEO stays becomes sufficiently small, that is, if the upper bound \hat{q}^{mid} on the outside wage offer is sufficiently high.

R4. *Given any wage offer, a middleaged manager is more likely to stay if the firm retains another manager who is either (i) older and closer to retirement, or (ii) middleaged and sufficiently less talented.*

Consider a middleaged manager in division 1 and fix his skill $y \in \mathcal{S}_2$. Under partitions 7 and 8, we will show that M1's expected lifetime payoff from staying in the firm is higher when the firm asks the manager of division 2 to stay, than when the firm lays off M2.

First consider any fixed wage offer to M1, \tilde{W}_t^1 . Under either partition 7 or 8, if the firm lays off M2 or M2 quits, then M1's expected discounted lifetime payoff from staying is $\tilde{W}_t^1 + \beta \mu_{t+1}^{P4}(y)$.

In partition 7, when M2 is old with skill x and the firm chooses to keep him by offering wage W_t^2 , the old manager stays with probability²¹ $Pr^{old}(x, W_t^2) > 0$. If both managers stay, M1 will be the only inside manager in period $t + 1$, and will be promoted to CEO with probability one. Therefore, M1's expected payoff is $\tilde{W}_t^1 + \beta\mu_{t+1}^{P4}(y) + Pr^{old}(x, W_t^2)\beta[\int W_{t+1}^c(3y+\tau)dT(\tau) - \mu_{t+1}^{P4}(y)]$. Since the CEO wage is sufficiently high, $\int W_{t+1}^c(3y+\tau)dT(\tau) \geq \mu_{t+1}^{P4}(y)$ — the inequality is strict if M1 is not promoted with probability one under partition 4. Therefore, M1's expected payoff from staying is higher when the firm keeps the old manager M2. A similar result holds in partition 8, when the firm keeps a middleaged manager M2 with skill $x \leq y - \underline{T} - \bar{T}$: M1 is promoted with probability one if the sufficiently less talented M2 stays. The result extends to higher skill levels $x \in (y - \underline{T} - \bar{T}, y)$ if the probability that M2's skill will be higher than M1's skill at period $t + 1$ is sufficiently small.

Now consider flexible wages. The middleaged manager's expected payoff from staying in the firm under partition 3 is

$$\frac{1}{2} \left[l_2 + (1 - \lambda)y + \beta \left[\mu_{t+1}^{P4}(y) + \int V_{t+1}(\emptyset, 2s, 3y) dF_2(s) - \lambda y \right] - \eta_t - \beta \bar{\eta}_{t+1} \right] + (1 + \beta)\lambda y,$$

while his payoff under partition 7, when the firm keeps both the middleaged manager and the old manager, is

$$\begin{aligned} & \frac{1}{2} \left[l_2 + (1 - \lambda)y + \beta \left[\mu_{t+1}^{P4}(y) + \int V_{t+1}(\emptyset, 2s, 3y) dF(s) - \lambda y \right] - \eta_t - \beta \bar{\eta}_{t+1} \right] + (1 + \beta)\lambda y \\ & + \frac{1}{2} Pr_t^{P7old*}(x, y) \beta \left[v_{t+1}^{P7}(y) + W_{t+1}^c(3y) - \mu_{t+1}^{P4}(y) \right]. \end{aligned}$$

Since $Pr_t^{P7old*}(x, y) > 0$, it is sufficient to show that $v_{t+1}^{P7}(y) + W_{t+1}^c(3y) - \mu_{t+1}^{P4}(y) > 0$.

From partition 7, recall that the firm's expected future profit function

$$\begin{aligned} v_{t+1}^{P7}(y) & \equiv \iiint \left[V_{t+1}(\emptyset, 3y + \tau, \emptyset) - V_{t+1}(\emptyset, 3y + \tau, 2s_2) \right. \\ & \quad \left. + V_{t+1}(\emptyset, 2s_1, 2s_2) - V_{t+1}(\emptyset, 2s_1, \emptyset) \right] dT(\tau) dF_2(s_1) dF_2(s_2), \end{aligned}$$

consists of differences in the firm's expected profit, hence it is independent of \bar{W}_{t+1}^c . For any M1 expected to be promoted to CEO next period with probability less than one under partition 4, if the constant \bar{W}_{t+1}^c of CEO wage is sufficiently high then $W_{t+1}^c(3y) - \mu_{t+1}^{P4}(y)$ is sufficiently high and the result $v_{t+1}^{P7}(y) + W_{t+1}^c(3y) - \mu_{t+1}^{P4}(y) > 0$ holds. For any M1 promoted with probability one, $W_{t+1}^c(3y) - \mu_{t+1}^{P4}(y) = 0$ and $V_{t+1}(\emptyset, 3y, 2s_2) = V_{t+1}(3y, \emptyset, 2s_2)$, hence

$$\begin{aligned} v_{t+1}^{P7}(y) & \equiv \iiint \left[V_{t+1}(\emptyset, 3y + \tau, \emptyset) - V_{t+1}(3y + \tau, \emptyset, 2s_2) \right. \\ & \quad \left. + V_{t+1}(\emptyset, 2s_1, 2s_2) - V_{t+1}(\emptyset, 2s_1, \emptyset) \right] dT(\tau) dF_2(s_1) dF_2(s_2), \end{aligned}$$

²¹If M2 stays with probability zero, such wage offer is equivalent to laying M2 off.

In this case, we follow the proof of R3 above to show that $v_{t+1}^{P7}(y) > 0$. The CEO skill's marginal contribution to profit $\rho - \frac{\partial W_t^c(a^c, s^c)}{\partial s^c}$ drops out in the negative term $V_{t+1}(\emptyset, 3y, \emptyset) - V_{t+1}(3y, \emptyset, 2s_2)$ since the CEO is the same in both cases. The term $V_{t+1}(\emptyset, 2s_1, 2s_2) - V_{t+1}(\emptyset, 2s_1, \emptyset)$ will be sufficiently positive if the CEO skill's marginal contribution to profit $\rho - \frac{\partial W_t^c(a^c, s^c)}{\partial s^c}$ is sufficiently large — the firm gains *both* by choosing which manager to promote to CEO, and by choosing whether to keep or not the manager not promoted. Moreover, the negative term $V_{t+1}(\emptyset, 3y, \emptyset) - V_{t+1}(3y, \emptyset, 2s_2)$ goes to zero as the probability that a middleaged manager who is not promoted to CEO stays becomes sufficiently small, that is, if the upper bound \hat{q}^{mid} on the outside wage offer is sufficiently high.

R5. *The firm's optimal placement/wage offers to a manager are complex, potentially non-monotonic functions of the attributes of all other executives.*

The example in Section 3 proves the result.

B Extended Model

In this section we describe a full-fledged version of our model, in which the firm can costly search for outside experienced executives, and show how the model becomes intractably complex. We then show how our simplified, tractable model retains the key incentive considerations of the full model.

B.1 Search for Outside Experienced Executives

Each period, the firm can hire an Executive Search Agency (ESA) to search for one experienced outside executive who is a suitable candidate for the CEO position, and up to two suitable candidates for the managerial positions. The ESA charges C^{CEO} for its CEO search and C^m for each managerial search. During the search, the ESA draws the candidates from a known probability distribution over three characteristics: age $a^j \in \{2, 3\}$, expected skill \tilde{s}^j , and demanded wage W_t^j , where $j \in \{CEO, 1, 2\}$ indicates the CEO, the first and the second managerial candidates, respectively. This probability distribution represents the pool of executives working in other firms. After the search, the ESA conveys the candidate's characteristics $\{a^j, \tilde{s}^j, W_t^j\}$ to the firm. The executive's true skill is $s^j = \tilde{s}^j + v^j$, where v^j is an independent zero-mean random variable; the firm can only learn the executive's true skill after he works for one period at the firm. In addition to the search cost, the firm incurs in an adjustment cost C^a if it hires the outside CEO, which represents the output forgone in the process of the new outside CEO becoming familiar with the

inner working of the firm and its market²².

B.2 Personnel Management Technology

At the beginning of the period, the firm and its inside managers observe the characteristics (age and skill) of all inside executives.

CEO Selection Process: The firm first engages in the process of selecting a CEO. The firm can either hire the ESA to draw an outside CEO candidate at the beginning of the CEO selection process, before making offers to inside executives, or after a promotion offer to an inside executive was rejected, and the candidate quit the firm. However, the firm can only search for an outside CEO candidate once. The CEO selection process ends whenever an executive accepts firm's offer.

When the firm has a CEO from the previous period, it can either lay the CEO off or make a wage offer to him. The current CEO then receives an outside employment (wage) offer, which is private information, and decides whether to stay at the firm or to quit and take the outside wage offer.

When the firm has no CEO (the CEO retired, quit or was laid off), it can either hire an outside executive as CEO (if the firm searched), or make a promotion/wage offer to an inside manager. The outside executive always accepts the offer at his demanded wage W_t^{CEO} . An inside manager who received the CEO promotion offer then receives an outside wage offer, which is private information, and decides whether to accept the promotion or to quit and take the outside wage offer. The process continues until an executive accepts the CEO promotion.

Managerial Selection Process: After the CEO selection process is concluded, the firm can hire the ESA to draw up to two outside managerial candidates. The firm can only search for outside managerial candidates once every period. After the search, the firm announces which inside managers will be laid off, and makes wage offers to the inside managers that it wants to retain. Each inside manager receives an outside employment offer. Managers simultaneously decide whether to stay in the firm or to quit and take the outside wage offer. The firm then fills in the vacant managerial positions hiring one or both outside experienced executives, if the search was conducted, or hiring outside young executives.

Hired outside executives draw their true skill, production takes place, profits are realized and wages are paid. Old executives retire, remaining executives get older.

²²Ang and Nagel (2007) find that a significant portion of the performance difference between external CEO hires and inside CEO hires is associated with external hire's relatively poor performance in their first year, presumably due to this learning/adjustment process.

B.3 Analysis

Dimension of the Model: At the beginning of the period, (a) the CEO position might be vacant or filled by an old executive; (b) the potential outside CEO candidate might be middleaged, old or nonexistent (if the firm does not hire the ESA to search); (c) each managerial position might be vacant, filled by a middleaged or an old executive; (d) each potential outside managerial candidate might be middleaged, old or nonexistent. There are **216** possible combinations of age-profiles of executives ($216 = 2 \times 3 \times 6 \times 6$). In addition, one has to take into account the skill dimension of each of the (potentially) six executives involved, and the possible multiple rounds of the CEO selection process. Firm-wide and outside-executives distribution of ages and skills enter incentives and strategies non-trivially, as they influence the probability distribution over future actions and states. The complexity of both the firm’s action space and the state-space is daunting: fully solving and characterizing the equilibrium would result in a paper of unpractical length.

Key Incentives: The above framework has two key features that are intrinsic to the executive labor market: (a) it is costly to search for outside experienced executives, and (b) there is extensive uncertainty about the outcome. As a result, when searching for an outside CEO candidate, the firm knows that, with some positive probability, the costly search will not uncover a talented candidate at a profitable wage. Therefore, maintaining a pool of talented inside executives is valued: talented managers not only increase output today, but they are also potential candidates for the CEO position in the future; the firm internalizes that nurturing inside managers is important. The firm also internalizes the option value embedded in its on-the-job skill assessment of managers. The firm can hire an outside manager to learn his true skill, keeping executives that turn out to be talented, and laying off these who turn out to be untalented — to hire and evaluate a new executive.

In addition, the firm’s inside executives know that the firm might find or not a “good” outside candidate, a potential competitor for the CEO promotion. Inside managers take this into account when evaluating their own probability of CEO promotion, and hence their expected payoff from staying in the firm.

Our simplifying assumption that search costs are prohibitively high greatly improves the tractability of the model, while preserving both of these intrinsic incentive considerations. The assumption implies that all external hires are young executives. With uncertainty about their talent, the firm faces a tradeoff between firing an experienced executive with known talent, but shorter employment horizon, and hiring an outside worker with unknown talent, but longer employment horizon. The young executive could be highly talented, a suitable candidate for the CEO position, or a lemon.

This framework emphasizes firm's endogenous value for (a) nurturing inside talented managers, and (b) on-the-job skill assessment of young executives.

Hiring a young outside manager also affects inside experienced managers who are competing for the CEO promotion. In their quit decisions, they must consider the probability that the new manager will be highly talented and win the tournament. Our framework retains the relevant uncertainty about the CEO promotion, even in periods where both inside managers are experienced with known talent. An experienced manager can optimally quit if he receives a better outside employment offer. When a manager quits, he is replaced by a young executive, who can turn out to be a better CEO candidate than an experienced manager who chose to stay at the firm. Hence, inside managers always consider in their quit decision the probability that the firm will hire a young manager, who may be highly talented and win the next CEO promotion.

References

- [1] **Ang, James S. and Gregory Leo Nagel** "Transferable Skills vs. Homecourt Advantage: A Comparison of Real Asset Performance by Externally and Internally Hired CEOs," November 2007. Available at SSRN: <http://ssrn.com/abstract=1028714>
- [2] **Bai, Chong-En and Yijiang Wang** "Uncertainty in Labor Productivity and Specific Human Capital Investment," *Journal of Labor Economics*, July 2003, *21*(3), pp. 651-675.
- [3] **Belzil, Christian and Michael Bognanno** "The Promotion Dynamics of American Executives," IZA Discussion Paper No. 1003, January 2004.
- [4] **Bernhardt, Dan** "Strategic Promotion and Compensation," *Review of Economic Studies*, April 1995, *62*(2), pp. 315-339.
- [5] **Bernhardt, Dan and David Scoones** "Promotion, Turnover, and Preemptive Wage Offers," *American Economic Review*, September 1993, *83*(4), pp. 771-791.
- [6] **Bognanno, Michael L.** "Corporate Tournaments," *Journal of Labor Economics*, April 2001, *19*(2), pp. 290-315.
- [7] **Carmichael, Lorne** "Firm-Specific Human Capital and Promotion Ladders," *Bell Journal of Economics*, Spring 1983, *14*(1), pp. 251-258.
- [8] **Demougin, Dominique and Aloysius Siow** "Careers in Ongoing Hierarchies," *American Economic Review*, December 1994, *84*(5), pp. 1261-1277.

- [9] **Gibbons, Robert and Michael Waldman** “A Theory of Wage and Promotion Dynamics Inside Firms,” *Quarterly Journal of Economics*, November 1999, *114*(4), pp. 1321-1358.
- [10] **Hayes, Rachel M., Paul Oyer and Scott Schaefer** “Co-Worker Complementarity and the Stability of Top Management Teams,” *Journal of Law, Economics, & Organization*, April 2006, *22*(1), pp. 184-212.
- [11] **Lazear, Edward P. and Sherwin Rosen** “Rank-Order Tournaments as Optimum Labor Contracts,” *Journal of Political Economy*, October 1981, *89*(5), pp. 841-864.
- [12] **Prendergast, Canice** “The Role of Promotion in Inducing Specific Human Capital Acquisition,” *Quarterly Journal of Economics*, May 1993, *108*(2), pp. 523-534.
- [13] **Rosen, Sherwin** “Prizes and Incentives in Elimination Tournaments,” *American Economic Review*, September 1986, *76*(4), pp. 701-715.
- [14] **Waldman, Michael** “Job Assignments, Signalling, and Efficiency,” *RAND Journal of Economics*, Summer 1984, *15*(2), pp. 255-267.
- [15] **Zábojník, Ján and Dan Bernhardt** “Corporate Tournaments, Human Capital Acquisition, and the Firm Size-Wage Relation,” *Review of Economic Studies*, July 2001, *68*(3), pp. 693-716.