ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 31/25

Talent Management: the Role of Bosses

Guido Friebel, Michael Raith

www.rfberlin.com

JULY 2025

Talent Management: the Role of Bosses

Authors

Guido Friebel, Michael Raith

Reference

JEL Codes: D2, D8, L2, M5

Keywords: middle managers, internal labor markets, human capital, training, talent hoarding, multi-divisional firm, intra-firm mobility

Recommended Citation: Guido Friebel, Michael Raith (2025): Talent Management: the Role of Bosses. RFBerlin Discussion Paper No. 31/25

Access

Disclaimer

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 175 522 4064
E-mail: info@rfberlin.com
Web: www.rfberlin.com

Talent Management: the Role of Bosses

Guido Friebel (Frankfurt) and Michael Raith (Rochester)[1]

July 14, 2025

[1]Contact: gfriebel@wiwi.uni-frankfurt.de; raith@simon.rochester.edu. This paper builds on and replaces our 2014 paper entitled "Managers, training and internal labor markets." Many thanks to the editor, Fabian Lange, and two anonymous referees, whose comments greatly helped us to improve the paper. For helpful comments on earlier versions we would also like to thank Ricardo Alonso, Guy Arie, Guillermo Caruana, Marco Celentani, Kong-Pin Chen, Mikhail Drugov, Bob Gibbons, Daniel Herbold, Richard Holden, Maike Lentner, Mike Powell, Patrick Rey, John Roberts, Dirk Sliwka, Thierry Verdier, Mike Waldman, Julie Wulf, and seminar audiences at Boston University, Carlos III, Marburg, MIT, Queen's, Rochester, UNSW Sydney, USC, the Spring 2012 NBER Organizational Economics Working Group meeting, the 2014 POeK Colloquium in Cologne, and the 2024 CESIFO workshop on reskilling in Venice.

Abstract

Managers ("bosses") are central to the development and allocation of human capital in firms because they train employees and learn about their abilities. While a multi-divisional firm wants to allocate workers to wherever they are most productive, bosses who are rewarded for their units' performance prefer to hold on to good employees, and the prospect of losing good people weakens the incentives to train them. We derive the optimal incentive contract for bosses that enables a firm to change from "silos" with only upward mobility to a "lattice" with cross-divisional mobility. Compared to silos, a lattice achieves a more efficient allocation of people to positions, but also entails agency costs that may exceed the benefits. We suggest empirical predictions about when silos or a lattice are optimal, and relate our model and its results to examples and evidence.

Keywords: middle managers, internal labor markets, human capital, training, talent hoarding, multi-divisional firm, intra-firm mobility

JEL-codes: D2, D8, L2, M5

1 Introduction

Modern firms' success depends on how talented workers are identified, developed, motivated and assigned most productively, activities that a large practitioner-focused literature has labeled "talent management" (e.g., Cappelli, 2008; Collings and Mellahi, 2009; Conaty and Charan, 2010). Practitioners and scholars alike agree that the traditional internal labor markets of the 20th century (Doeringer and Piore, 1971; Milgrom and Roberts, 1992; Baker et al., 1994) no longer suit many of today's large, multi-divisional firms. They argue that to better match people with jobs, foster human capital growth, and improve worker satisfaction and retention, firms need to do away with vertical job ladders confined to corporate "silos" (Rosen 2010), and instead implement "lattice"-like structures (Benko and Anderson 2010) with lateral mobility, such as the "internal talent markets" (Cowgill et al., 2023) that many firms have adopted. However, most firms struggle to implement even essential elements of talent management. Most workers do not feel supported in their career development (Mercer, 2021), and most firms lack formal reward systems for developing and promoting talent (i4cp, 2016). As a result, lacking career development is one of the most common causes of attrition (Work Institute, 2020, and Gallup surveys cited in Green and McClave, 2021).

We suggest a theory that zooms in on the role of managers (bosses) in internal labor markets to better understand firms' challenges with talent management. Identifying and developing talent is significantly in the hands of managers, as a rapidly growing economics literature documents.[1] However, managers' incentives often deviate from firm objectives. While firms would want to move people through the positions in which they are most productive, and have them trained accordingly, bosses face much narrower incentives as they are usually rewarded for the performance of their group, department, or project. Those rewards provide indirect incentives to invest in juniors' human capital, but they also create strong incentives for bosses to keep the best juniors on their team.

[1] For instance, Cullen and Perez-Truglia (2019), Frederiksen et al. (2020), Hoffman and Tadelis (2021), Friebel et al. (2022), Diaz et al. (2024), Sandvik et al. (2021), Minni (2023), Haegele (2024), Benson et al. (2024). For a recent survey of this research, see Hoffman and Stanton (2024). See also Gallup (2015) and the Chartered Institute of Personnel and Development (2020), Cappelli (2013), Whittaker and Marchington (2003) and Perry and Kulik (2008).

This problem has been labeled "talent hoarding," which refers to "manager behaviors aimed at preventing subordinates from pursuing jobs that will take them to work for a different manager" (Keller and Dlugos, 2024). Talent hoarding and its reverse, efforts to get rid of people without firing them, are as old as modern management:

> "Nothing does more harm than the too common practice of promoting a poor man to get rid of him, or of denying a good man promotion 'because we don't know what we'd do without him'. The promotion system must ... make difficult alike kicking upstairs and hoarding good people" (Drucker, 1954, p.154-155).

However, while talent hoarding has been discussed in the academic business literature for a while now (e.g., Mellahi and Collings, 2010), reliable evidence has been scarce.[2] Only recently, Haegele (2024) established robust evidence for talent hoarding as a corporate reality with significant negative effects on the quality of who gets promoted.

In this paper we investigate how the dual role of managers as unit leaders and as mentors leads to talent hoarding, and what it takes to prevent it. We show that the right incentive system can prevent talent hoarding and enable cross-divisional mobility of workers, which leads to a more efficient allocation of workers than a traditional internal labor market with only vertical mobility. However, there are associated agency costs of achieving this, and the benefits may or may not exceed the costs. We suggest empirical predictions about when each structure is optimal, and relate our model and its results to examples and evidence. Our paper makes a first step towards a better understanding of the obstacles to effective talent management in modern firms, and is to our knowledge the first to theoretically study lateral transfers in the internal labor market of a multi-divisional firm.

We model a firm with two divisions that is headed by a CEO (Section 3). Each division consists of a senior and a junior manager, but because our focus is on the role of bosses, the juniors are not strategic players. Our model makes the following key assumptions:

(1) Other things equal, senior managers are more productive than junior managers due to their position, in line with agency-based and knowledge-based models (Rosen, 1982; Qian,

[2] When we wrote the first version of this paper ten years ago, we could find very little evidence about talent hoarding, whereas today a simple search for "talent hoarding" yields 10 million Google results.

1994; Garicano, 2000), and consistent with empirical evidence (Lazear et al., 2015).

(2) As argued above, training has become a key responsibility of managers in many firms. Accordingly, senior managers in our model have two tasks, to run their division and to train their junior. In reality, both consist of many smaller tasks ("inputs" in the terminology of Raith, 2008) that are difficult to measure.[3] We thus assume that managers are primarily held accountable for their own unit's performance (output) and not the inputs, and may also be rewarded for the other division's output.

(3) Division output is the result of team production between the senior and the junior (Alchian and Demsetz, 1972; Meyer, 1994). That is why a senior manager who is accountable for his division benefits from having a good junior. The team nature of production both provides indirect incentives to train a junior, and is the key cause of talent hoarding; otherwise seniors would have no stake in the human capital of their juniors.

(4) For exogenous reasons, a senior manager may leave, and his position becomes vacant. Due to acquired specific human capital, the CEO prefers to fill the vacancy with a junior known to be qualified to be a senior; otherwise the CEO hires a replacement from the outside market. Hence, as in Baker et al. (1994), Bidwell (2011), Bidwell and Keller (2014), there are designated ports of entry, options for promotions from within, but also external hiring at the higher rank.

(5) Only a fraction of juniors has potential for a higher-level position. Each senior is privately informed about his junior's potential, and a junior can be promoted only based on his boss's recommendation. This assumption may or may not hold in any given firm, but is consistent with the evidence on the role of bosses for workers' career advancement (Haegele, 2024; Minni, 2023; Benson et al., 2024). In our model, recommendations take the form of cheap-talk messages from the seniors to the CEO, upon which the CEO fills a vacancy with the most productive person available.

(6) While managers can be held accountable only for outputs and not their individual activities, it is possible to tie compensation to the transfer of one's junior. Again, this may be difficult to do in some firms, as we discuss in Section 5.2, but is an important feature of

[3] See Feltham and Xie (1994), Baker (2002), and Gibbons (2005, Section 2.1) for general frameworks of performance measurement with multiple tasks.

the optimal contract in our main result.

Abstracting from a rich variety of internal labor markets in practice, we compare two different structures. In the first, which we refer to as "silos," there is only vertical mobility; that is, only the junior in the same division can be promoted to a vacant senior position.[4] The second structure, which we call a "lattice" (see e.g. Benko and Anderson, 2010), more closely resembles modern firms' efforts to support the lateral mobility of workers too. These efforts may be more centralized or decentralized (see Cowgill et al., 2024), but even in firms with decentralized "internal talent markets," workers who apply for a vacancy still often require either their boss's explicit approval (Haegele, 2024) or their boss's rating for having the potential to be promoted (Benson et al., 2024). In our model, with a lattice a senior vacancy can be filled with with a qualified junior from the other division, as long as the junior's boss recommends him.

With silos (Section 4.3), the optimal contract is simple (Proposition 1): Rewarding senior managers (only) for their own division's output incentivizes execution effort directly, and at the same time incentivizes training because a more productive junior makes the team more productive. However, output is a noisy measure of training, and in our model with wealth-constrained managers results in rents for the managers that are costly to the firm. Moreover, narrow incentives de facto create silos even if the firm supports inter-divisional mobility because senior managers who stand to lose from the departure of their junior will "hoard talent" by reporting that the junior has no potential to be promoted.

Implementing a lattice therefore requires a reward system that ensures truthful communication by the seniors, and compensates them for the possible loss of a junior to the other division. At the same time, if communication is truthful, the transfer of a junior is direct evidence of the manager's training effort, and tying rewards to a transfer is therefore cheaper than rewarding training through own-division performance. To isolate this point, we show in Proposition 2 that if the CEO could observe the juniors' potential, it would be optimal to focus training incentives maximally on a transfer-contingent bonus, and to reward own-division performance only as needed to incentivize execution effort.

[4] "The term silo is a metaphor suggesting a similarity between grain silos that segregate one type of grain from another and the segregated parts of an organization" (Rosen, 2010).

Incorporating truthtelling constraints for managers leads to our main result. The optimal incentive contract for a lattice (Proposition 3) includes a bonus for the other division's output that is contingent on the transfer of a junior—in other words, a reward based on the transferred junior's performance. Only such a bonus can elicit truthful endorsements and thus simultaneously prevent both talent hoarding and "kicking upstairs" an unqualified junior. Like in the full-information case, the expected wage cost can be lower than with silos if compensation can be significantly transfer-contingent, in which case a lattice unambiguously dominates silos. In general, however, implementing a lattice is costly for the firm because of the additional constraints, and because managers still need to be rewarded for own-division output to run their division. Silos are thus optimal if the additional agency costs of the lattice exceed its productivity gain.

The firm can still implement a lattice even if transfer-contingent rewards are infeasible, by unconditionally paying a large enough reward for the performance of the other division (Proposition 4). This is a costly solution, though, and therefore less likely to dominate silos.

Comparative statics predict that the difference in a firm's profit with a lattice compared to silos and, hence, the desirability of implementing a lattice, increases with the training cost and decreases with the cost of execution effort. For example, while a higher training cost reduces profit under both structures, the cost increase is greater with silos because incentivizing training solely through an own-division bonus is less efficient than the use of a transfer-contingent reward with a lattice. It follows that the more the responsibility for human capital development falls on bosses (as opposed to, for instance, formalized training), the more likely implementing a lattice is worthwhile for a firm. The profit difference is also greater the more important the senior manager is relative to the junior, as it increases the value of getting juniors with potential (as opposed to outside hires) into senior positions.

Our results help to explain both the observed practices of companies, as well as the problems of companies that fail to fully account for the role of bosses in internal labor markets (Section 5). First, our theory helps to explain the prevalence in the 20th century of internal labor markets with vertical job ladders but little lateral mobility (Doeringer and Piore, 1971; Baker et al., 1994). Silos may well be optimal, especially if running their division requires a significant share of managers' time relative to the mentoring of subordinates. Arguably, over the last decades, the importance of talent for firm performance has increased with new technologies and higher degrees of competitiveness. The relative benefit of a silo structure (lower managerial wages) hence becomes less important, and firms may find it attractive to move to a lattice.

Second, even firms that encourage cross-divisional mobility and establish internal market platforms are unable to reap the full benefits without adjusting bosses' incentives. Managers who are rewarded only for their unit's performance will hoard their best people instead of recommending them for opportunities elsewhere, as Haegele (2024) documents.

Third, simple fixes to the incentive system tend to be problematic. Shifting the weight of incentives unconditionally from unit to corporate performance works, but is expensive. Simple referral bonuses, in turn, are prone to create incentives to "kick upstairs" weaker people, which explains why such bonuses, when used, tend to be too small to matter for the allocation of managerial talent.

Fourth and most strikingly, an example of a company whose incentive contracts closely resemble the one characterized in our main result is Haidilao, a Chinese chain of hot pot restaurants. To incentivize the mentoring of protégés who will open new restaurants, rewards for restaurant managers in this chain are almost fully loaded on the performance of protégés' restaurants, with a profit share many times larger than that for the own restaurant (Zheng and Zhao, 2018; see Section 5.2 for further discussion).

Also in Section 5, we discuss alternative instruments such as job rotation or the direct measurement of bosses' mentoring, and conclude with remarks on the perspective of juniors, for whom a lattice is unambiguously preferable to silos.

Overall, our main contribution is to investigate the role of bosses as both information gatekeepers and as mentors in internal labor markets, along with their primary role of leading a unit within the organization. We show that while the allocational benefits of cross-divisional mobility over traditional job ladders are obvious, they cannot be realized without a supporting incentive system for bosses. The solution we suggest is simple in theory, but not always simple to implement in practice. Moreover, securing bosses' cooperation in facilitating cross-divisional mobility may raise or lower the wage cost, which needs to be considered alongside the benefits of greater mobility.

2 Literature

Our paper builds on a number of literatures. First, labor economics studies firms' incentives to train workers and the role of private information about workers' abilities. However, firms are treated as monolithic or run by principals (entrepreneurs or partners), and the loss of good workers to other firms is treated as the key problem. We argue that similar problems arise within firms, because it is not "the firm" but middle managers who carry out the training and as a consequence learn about workers.

A seminal contribution on the role of asymmetric information for talent management is Waldman (1984), who argues that firms know their workers better than outside firms do (and which Kahn, 2013, showed empirically). Promotions, though, are visible. Waldman shows that to minimize the poaching of good workers, optimal promotions are less frequent and are associated with larger wage increases than with symmetric learning between firms. Milgrom and Oster (1987) argue that the same information asymmetry may lead to promotion discrimination against disadvantaged workers. De Varo and Waldman (2012) predict that the signaling value of a promotion is lower for highly educated workers, and find supporting evidence in the data of Baker, Gibbs, and Holmström (1994). Firms' learning about their own workers also plays a key role in the integrative theory of Gibbons and Waldman (1999), but again it is the firm that does the learning, not individuals within the firm.[5]

Partly building on Waldman (1984), Acemoglu and Pischke (1998) show that in contrast to Becker's (1964) classical analysis but in line with practice (notably in Germany), firms have an incentive to invest in general human capital development as long as their private information about workers' abilities confers monopsony power.[6] Bar-Isaac and Leaver (2021) take this argument a step further by endogenizing the firm's information disclosure to the market. Similar to the tendency to hoard good people and "kick upstairs" bad ones in our paper, Bar-Isaac and Leaver show that a firm optimally discloses bad matches but not good ones.

[5] See Kahn and Lange (2014) for evidence of firms' imperfect learning about their workers.

[6] An alternative method, as shown by Garicano and Rayo (2017) and Fudenberg and Rayo (2019), is for training in general skills to be inefficiently slow (i.e., a long apprenticeship), which enables the firm to recoup the cost of training through the apprentice's work.

Morrison and Wilhelm (2004) emphasize, as we do, that human capital is often imparted by senior colleagues as mentors, and investigate when seniors will have the incentives to incur the individual costs of mentoring. Their focus is on partnerships, however, and the partners' equilibrium incentives to mentor stem from their ability to sell, upon retirement, their stake in the firm at a price that is sustained by the firm's reputation for promoting only high-productivity individuals to partner.

In contrast to this literature, middle managers as bosses are at the center of our theory. Middle managers have only recently come into the focus of economics. Lazear et al. (2015), Hoffman and Tadelis (2021), Friebel et al. (2022), Minni (2023), Drechsel-Grau and Holub (2024), Benson et al. (2024) are just some of the papers that have shown that middle managers play an important role in the development of workers, their decision to stay in or leave the firm, their productivity, and upward mobility. In most larger corporations, middle managers, not the principals, are the ones who bear the costs of mentoring, and who in the process acquire private information about their juniors' abilities. Bosses are likely to be concerned about losing good juniors even if the human capital in question is firm-specific (as long as it is general across divisions), and even if "poaching" occurs within the same firm, because they are accountable and often explicitly rewarded for their unit's performance. However, and in stark contrast to Waldman (1984), Acemoglu and Pischke (1998), and Bar-Isaac and Leaver (2021), in our theory, the firm's goal is to increase mobility within the firm rather than to reduce mobility across firms, and asymmetric information amplifies rather than relaxes the firm's agency costs. Similar to our focus, He and Waldman (2024) also look at asymmetric information about talent in firms, and how it can be overcome. Their setting is different from ours, though, and so are the instruments that they consider, namely job rotation and firm-wide bonuses. While He and Waldman emphasize the roles of commitment and worker visibility, in our paper managers' incentives to train workers play a key role, in addition to asymmetric information.

More broadly, our paper belongs to a literature that studies distortions in internal labor markets caused by self-interested bosses, and possible organizational solutions. This literature includes Fairburn and Malcomson (1994) and Prendergast and Topel (1996), who focus on biased evaluations by managers, as well as Carmichael (1985) and Friebel and Raith

(2004) who consider bosses' incentives to hire unthreatening but inferior subordinates. We believe that the present paper points to a more fundamental incentive problem that concerns both the production and allocation of human capital, and leads to the joint determination of both incentive contracts and the organization of internal labor markets.[7]

In contrast to much of the literature, our paper focuses on the demand side rather than the supply side of internal labor markets. That is, while leading theories such as Gibbons and Waldman (1999) focus on how to allocate employees to different possible positions, without constraints on the availability of positions, our focus is on how a firm should fill vacant positions. Slot constraints and "job vacancy chains" (that emerge when filling one vacancy creates a new vacancy elsewhere) have been studied in the industrial-relations literature; see Chase (1991) and Pinfield (1995). Economic analyses of demand-side considerations are more scarce; a prominent exception is Demougin and Siow (1994), which is otherwise very different from our paper. More recently, Ke at al. (2018) explicitly model slot constraints as a consequence of a firm's organizational structure, and analyze the consequences for the optimal management of an internal labor market with workers who are motivated by their career within the firm. Bianchi et al. (2021) study empirically how delays in older workers' retirement impedes the careers of their younger colleagues. Cowgill et al. (2024) compare both theoretically and empirically centralized vs. decentralized matching of workers to jobs within a firm. They show that firms with centralized matching processes achieve higher-productivity matches than do decentralized "talent markets," but also lead to lower worker satisfaction and growth. Neither of these papers consider bosses' incentives, which are central in ours.

Our literature review would be incomplete if we did not acknowledge the substantial work on strategic human capital and strategic human resource management (Boon et al., 2018). In their much cited book "The War for Talent," Michaels et al. (2001) pointed to "divisional

[7] Beyond the literature on internal labor markets, our paper overlaps with a literature concerned with the tension between giving managers incentives to pursue the goals of their unit, and encouraging them to coordinate and communicate with top management or other divisions. Pioneering contributions are by Levitt and Snyder (1997) and Athey and Roberts (2001); more recent work includes that of Alonso et al. (2008), Rantakari (2008, 2011), Dessein et al. (2010), Friebel and Raith (2010), and Dessein (2014).

hoarding" as a main obstacle to intra-firm mobility, but we are not aware of papers examining the strategic intricacies of the interaction between line managers and what is usually called "the firm." An exception is Keller and Dlugos (2023), who suggest (based on evidence from a health care firm) that managers who are known for developing workers also attract better workers. It is an open question under what conditions such a reputation-based system may solve the incentive problem we investigate.

3 Model

We model a profit-maximizing firm with two divisions headed by a CEO; see the organizational chart in Figure 1. Division $i = A, B$ consists of a senior manager S_i and a junior manager J_i. The senior managers are the main strategic players in our model; the juniors are not players in a game-theoretic sense (see Section 5.4 for a discussion of the juniors' perspective). The job titles are chosen for convenience; the model can more generally be interpreted as representing any two adjacent tiers of a multi-tier organization.

We will investigate whether it is optimal for the firm to make it possible for a junior to be promoted diagonally across divisions, i.e., to implement a "lattice" structure, or whether promotions will occur only within the same division, as in traditional job ladders or "silos."[8]

3.1 Production

Each division's ($i = A, B$) senior or junior manager ($k = S, J$) has a baseline productivity q_i^k that is high or low ($q_h > q_l > 0$). We assume that people acquire specific human capital as in Waldman (1999). This raises productivity by factor $\phi > 1$ if the manager has spent some time in the firm. Thus, $q_i^k \in \{\phi q_h, q_h, \phi q_l, q_l\}$. Let $\Delta q = q_h - q_l$. As we'll explain below, for juniors, Δq represents the benefit of training, whereas for seniors, it represents

[8] We focus on *promotions* to another division in order to juxtapose the market with traditional job ladders. However, our argument is more general and applies to any instances—including lateral transfers—in which a worker is more productive in a different division.

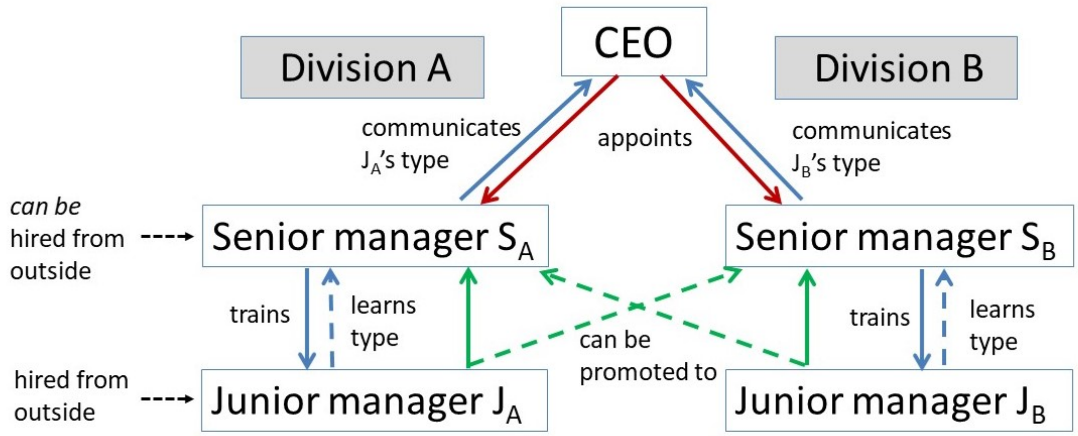

Figure 1: Structure of the firm

the difference between a qualified and an unqualified manager. Specific human capital, in turn, is the reason for the firm to prefer insiders to outsiders when filling vacancies.

The productivity of division i's senior-junior team is additive in both managers' productivities:

$$t_i(q_i^S, q_i^J) = \kappa_S q_i^S + \kappa_J q_i^J,$$

where we assume $\kappa_S > \kappa_J > 0$ to reflect the greater impact of a senior manager on division performance. A manager's baseline productivity and his position in the organization are thus complementary, which creates a benefit for the firm from promoting good juniors, as we will see shortly. The team may have quite different compositions of senior and junior managers. It is useful to simplify notation and denote specific values of team productivity $t_i(q_i^S, q_i^J)$ by the shorthand t_{sj}. Both senior or junior managers can be of productivity H,h,L, or l depending on whether $q_i^k = \phi q_h, q_h, \phi q_l$, or q_l. For example, t_{hL} denotes the productivity of a division with a high-productivity senior manager without specific human capital and a low-productivity junior manager with specific human capital.

Division i's output is a random variable $Y_i \in \{0, 1\}$ whose expected value $\Pr(Y_i = 1)$ depends on the team productivity and on S_i's execution effort $x_i \in \{0, 1\}$, and is given by

$$y_i(t_{sj}, x_i) = t_{sj} + \theta x_i \tag{1}$$

for $\theta > 0$. We assume $t_{HH} + \theta \leq 1$, which guarantees $y_i \leq 1$ and thus that $\Pr(Y_i = 1) = y_i$ is well-defined. (More generally, we will use capital letters for random variables and lowercase letters for their expected values.)

3.2 Hiring, training, attrition, and promotions

The timing of our model is as follows.

1. The CEO appoints managers S_A and S_B by offering them incentive contracts that the managers accept or reject. In equilibrium, contracts are designed such that the managers accept. Each senior manager has productivity q_h and reservation utility $\underline{U}$.

2. Each senior S_i $(i = A, B)$ hires a junior J_i from a pool of *ex ante* identical agents, who initially have productivity q_l.[9]

3. At cost τe_i for $\tau > 0$, S_i can invest in training effort $e_i \in \{0, 1\}$. If and only if $e_i = 1$, J_i's productivity increases from q_l to q_h. Through training, S_i also teaches J_i what it takes to be a senior manager. Being a senior manager, however, also requires innate talent, which J_i has with probability μ. Thus, training always increases a junior's productivity in a junior position, but it increases the productivity as a future senior only for a subset of juniors. If S_i invests in training, he acquires private information about J_i's "potential" to be a senior manager.[10]

4. Both S_i and J_i acquire specific human capital, which raises their productivity by factor ϕ.

[9] The assumptions that a firm can always hire a good senior manager without specific human capital, while juniors initially are of low quality and remain so unless they receive training, can be relaxed. They reflect, however, a common tradeoff between external and internal appointments: Outsiders can be hired from a liquid and informative market, but lack specific human capital, whereas insiders may be more productive, but need to be trained first. See Bidwell (2011) and Bidwell and Keller (2014).

[10] Benson et al. (2024) investigate the distinction between performance and potential that many firms implement using "9-box grids." In their data, 23% of men receive the highest performance rating but only 5% the highest potential rating. Minni (2023) finds that senior managers have considerable impact on juniors' careers because of the private information they have about juniors' suitability for positions elsewhere in the firm.

5. With probability $1-\sigma$, each senior quits for exogenous reasons (thus σ is the probability of staying with the firm); S_A or S_B quitting are independent events.

   If a senior quits, he can be replaced either internally by one of the juniors, or from the external market. Each senior communicates to the CEO whether his junior has the potential to be a senior manger; this communication is cheap talk. While a departing senior reveals his junior's type truthfully because there is no reason to misrepresent it, a senior who stays with the firm may have an incentive to lie. He may "hoard" a good junior or "kick upstairs" a bad one, depending on the incentives in place. Juniors, however, are unable to ascertain or to prove their own suitability for a higher-level position.[11]

   If a junior is promoted, he is replaced by a rookie junior with productivity q_l and without specific human capital, just like at stage 1. The CEO fills any senior vacancy to maximize the firm's profit at the time; she cannot commit to any hiring or promotion policy upfront.

6. Each senior S_i (who is either the original manager or a newly appointed one) exerts execution effort $x_i \in \{0, 1\}$ at cost ξx_i, for $\xi > 0$. The cost parameters τ and ξ represent the demands on a senior manager's time to train their junior and to run their division, respectively. As we will see, these parameters are key drivers of the firm's optimal choice between silos and a lattice.

7. Division outputs are realized and wages are paid.

We separate training (stage 3) and execution effort (stage 6) because the benefits of training tend to accrue with greater delay than those of many other managerial actions. Technically, however, there are no periods and no discounting; our model is a static multi-stage game.

Replacing a senior who leaves has the following consequences. Promoting a junior with potential results in having a senior with productivity ϕq_h. Hiring a senior from the outside

[11] These assumptions are consistent with Haegele's (2024) evidence that managers hoard their best workers, thus stalling the workers' mobility, except when the managers themselves move elsewhere in the organization. These patterns emerge because workers often cannot advance in their organization without the explicit support of their bosses, see also Minni (2023).

market results in having a senior with ability q_h (like in stage 1), without specific human capital. If a trained junior without potential is promoted, he retains his specific human capital but has a low ability as a senior, thus productivity ϕq_l.

To rank the two replacement options we make two assumptions: First, the firm prefers to promote *internally* a junior with potential (and to fill the resulting junior vacancy with a rookie) than to hire an outsider while keeping the high-productivity junior in his position. That is, t_{Hl} exceeds t_{hH}, which can be written as

$$\text{(INT)} \qquad \underbrace{\kappa_S(\phi - 1)q_h}_{y_G} > \underbrace{\kappa_J(\phi q_h - q_l)}_{y_L}.$$

The left-hand side of (INT) is the gain y_G of appointing as senior a good insider with specific human capital over an outsider without. The right-hand side is the loss y_L from replacing a high-productivity junior with specific human capital with a rookie. Because the production function is additive, the gain and loss from promoting a junior from the *other* division instead of an outsider are the same.

Second, the firm prefers to hire from the *outside* than to promote a junior whose potential is unknown to the CEO. Assuming that a senior invested in training, but without communication about his junior, the probability from the CEO's perspective that a junior has potential is μ. Hiring a senior from outside then leaves in place a junior with productivity ϕq_h, whereas promoting a junior is followed by hiring a rookie. We thus assume

$$\text{(OUT)} \qquad t_{hH} > \mu t_{Hl} + (1 - \mu)t_{Ll}. \tag{2}$$

Since $t_{Hl} > t_{hH}$ per (INT), (OUT) places an upper bound on μ. If μ exceeded that bound, the firm would always fill vacancies internally and would not need to rely on senior managers' information to identify good candidates.

3.3 Wage contracts

Managers are risk-neutral and are protected by limited liability; we assume that S_i's compensation w_i must be non-negative. A senior accepts to work for the firm if the expected equilibrium utility exceeds his reservation utility $\underline{U}$. A senior who leaves at stage 3 also receives $\underline{U}$. To simplify the exposition, we assume that the limited-liability constraint is binding and the participation constraint is not. For these two assumptions to hold without constraining the values of other parameters, we set $\underline{U} = 0$, in which case $w_i \geq 0$ and the incentive constraints derived below imply non-negative utility, see (9) in Section 4.2.

In many organizations, measurement problems (hidden actions and hidden information) are central. They are also central in our theory, in which we assume that execution effort x_i and training effort e_i cannot be observed. Because production is team production, we assume that no individual performance measures for the team members are available. The only available performance measures are hence the division outputs Y_i, $i = A, B$, which are verifiable. In our view, these assumptions capture the essential logic of the causes of talent hoarding, and possible remedies.

The firm offers wages W_i that consist of a fixed salary α, a bonus β for the own division's output, and a bonus γ for the other division's output. This specification rules out contracts that depend nonlinearly on (Y_A, Y_B).

However, we do allow for wages that are contingent on whether the own junior is transferred to the other division, given that the promotion of a junior to another division is a verifiable event. We build on the standard argument in the internal labor markets literature that wages can be tied to positions and therefore presume that wages can also be conditioned on changes in positions. Here, the wages of one manager (the senior) can be made contingent on another manager to be transferred to a different division (the junior). Thus,

$$W_i = \begin{cases} \alpha_n + \beta_n Y_i + \gamma_n Y_{-i} & \text{if } J_i \text{ stays in division } i \text{ (is \underline{n}ot transferred)} \\ \alpha_t + \beta_t Y_i + \gamma_t Y_{-i} & \text{if } J_i \text{ is \underline{t}ransferred (promoted) to division } -i \end{cases}$$

with symmetric coefficients $\zeta = (\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t)$ for each senior. If S_i's wage is not designed to be contingent on the transfer of J_i, then we will use the simpler notation $\zeta = (\alpha, \beta, \gamma)$.

If a senior leaves and is replaced, we assume that the newly-appointed senior manager works under the same incentive contract as an incoming senior at stage 1 of the game. However, our results do not depend on the details of how replacement managers are paid.[12]

[12] Strictly speaking, since our model is static, replacement managers need not be given any training

3.4 First-best and problem statement

We assume that it is optimal from the firm's perspective for the seniors to invest in both training and in execution effort, both under first-best (that is, under full information and if training and execution effort are contractible) and second-best conditions. Without training, no junior would ever have the potential for a senior position, and an internal labor market could not exist—senior positions would always be filled with outside hires. In turn, it will become clear in our analysis what would happen if the firm did not incentivize execution effort. To preview the answer, it is the presence of multiple tasks (training and execution) with constraints on performance measurement that makes it difficult for firms to establish cross-divisional mobility of workers. The parameter conditions that ensure the optimality of inducing training and execution effort amount to upper bounds on the effort costs τ and ξ. We refrain from stating the full conditions formally, but provide more detail in the proof of Proposition 1.

Conditions (INT) and (OUT) stated above then imply that in the first best, the CEO will fill a vacancy ideally with a junior with potential from the same division or from the other division, and otherwise—that is, if neither junior has potential—by hiring from outside.[13] Since $\underline{U} = 0$, each senior then needs to be paid the costs of training and execution:

$$w^{FB} = \tau + \xi.^{14} \tag{3}$$

In turn, if the firm paid a constant wage in the presence of the hidden-action and hidden-information problems described above, then the seniors would not exert either training or execution effort, and there would be no internal labor market.

incentives. In an ongoing organization, though, new managers need to train workers too, which is why it is reasonable to assume that they inherit the same wage contracts.

[13] Because the divisions are symmetric, it is never optimal to transfer a manager laterally to the other division. In contrast, if for instance division B were more important, it might be preferable to transfer a proven S_A to division B instead of promoting a junior.

[14] For this expression to hold, it is immaterial whether the payment is unconditional and a senior can be required to stay, or instead the senior might leave, forgoes his wage payment, and the wage is paid to his successor.

The objectives of our analysis below are twofold. First, we will characterize the optimal contract that, aside from inducing training and execution effort, also induces truthful communication from a senior manager to the CEO about his junior's potential, which is required for condition (INT) to hold and thus for cross-divisional transfers of juniors to be feasible. To save on the associated agency costs, however, it may be more profitable for the firm to forgo cross-divisional transfers and only promote juniors within the same division, as in traditional job ladders. Accordingly, a second objective of our analysis is to examine the firm's optimal design of its internal labor market as a choice between a "lattice" with cross-divisional transfers, and "silos" without. We denote this choice by $L \in \{0, 1\}$, where $L = 1$ refers to a lattice and $L = 0$ to silos.

The CEO cannot commit to this choice directly; she instead fills any vacant senior position at stage 5 of the game to maximize profit given her information about juniors. As we will see, however, the choice of the incentive contract for the senior managers in effect pins down the choice between $L = 1$ and $L = 0$.

See Table 1 for a list of the model's parameters, including some notation that we will introduce in Section 4.

4 Analysis

In this section, we first set up the firm's contracting problem (Sections 4.1 and 4.2). The silo case (Section 4.3) is analytically straightforward but leads to an important insight about talent hoarding and the centrality of bosses' incentives for talent management in firms. We examine the lattice case in Section 4.4. There, we first solve for the optimal full-information contract, which delivers a key part of the intuition for our main result, the optimal transfer-contingent contract subject to all relevant constraints. We conclude by deriving the optimal contract under the restriction that incentives cannot be transfer-contingent. In Section 4.5, we show how the firm's optimal choice between silos and a lattice depends on different parameters of the model.

For our analysis, we can exploit the symmetry of the model to focus on *one* division only, say division A. To do so while keeping the bookkeeping clean, define division A's profit, Π_A,

Symbol(s)	Description
q_l	Productivity of a junior before training
q_h	Productivity of a junior after training (not accounting for specific human capital); productivity of a senior hired from outside
Δq	$q_h - q_l$
ϕ	Factor by which specific human capital increases productivity
κ_S, κ_J	Position-specific productivity coefficients for a senior, junior
Y_i, yi	Realized and expected output of division i, where $Y_i \in \{0,1\}$ and $y_i \in [0,1]$
t_{sj}	Productivity of a team, where $s, j \in (H[\phi q_h], h[q_h], L[\phi q_l], l[q_l])$
θ	Productivity of execution effort
e_i	Training effort, $\in \{0,1\}$
τ	Cost of training
x_i	Execution effort, $\in \{0,1\}$
ξ	Cost of execution effort
μ	Probability that a trained junior has the potential to be a senior manager
σ	Probability that each senior stays with the firm beyond stage 5 of the game
$\underline{U}$	Senior managers' reservation utility
W_i, w_i	Firm's realized and expected wage cost of S_i
Π_i, π_i	Firm's realized and expected profit of division i
$\alpha_n, \beta_n, \gamma_n$	Constant payment, own-division bonus, and other-division bonus for a senior whose junior is [n]ot transferred to the other division
$\alpha_t, \beta_t, \gamma_t$	Constant payment, own-division bonus, and other-division bonus for a senior whose junior is [t]ransferred to the other division
ζ	Complete contract: $\zeta = (\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t)$
L	$L = 1$: lattice; $L = 0$: silos
v	$v = (1-\sigma)(1-\mu)$ probability of a vacancy that cannot be filled by the division's own junior

Table 1: Model parameters

as output in division A (Y_A), minus the compensation W_A paid to S_A *or her replacement*. Thus, Π_A includes possible payments to S_A based on division B's output, but does not include payments made to S_B based on division A's output. Below, we will often drop the division subscript A with the understanding that we continue to focus on one division only.

A lattice differs from silos in two states of the world. In the first, there is the opportunity to transfer a junior with potential to another division: S_A stays (probability σ), S_B leaves (probability $1-\sigma$), J_B has no potential to be a senior (probability $1-\mu$ if $e_A = 1$), but J_A does have potential (probability μ if $e_B = 1$). Here, if the CEO knew the juniors' types, S_B would be replaced by J_A, and J_A by a rookie (if his quality is known). This would be an overall productivity gain, but would lead to a productivity loss y_L in division A. It is this event that plays a key role for S_A's incentive constraints derived below. The second state of the world that differs from silos is the reverse: S_A leaves, J_A has no potential to be a senior, S_B stays in the firm, and J_B does have potential to be a senior. Then S_A is replaced by J_B, leading to a productivity gain y_G over hiring a new senior from outside as would be the case with silos. For a more compact notation, denote by $v = (1 - \sigma)(1 - \mu)$ the probability of a vacant S-position for which the same-division junior is not qualified. Then the ex-ante probability of each of the two states explained is $\sigma\mu v$.

4.1 Equilibrium payoffs

The firm's expected profit from either division, π, equals expected output y minus S_A's expected wage w. We derive each in turn.

Output: With silos, and assuming that $e_A = 1$ in equilibrium, denote division A's expected productivity by t^0. It is given by

$$t^0 = \sigma t_{HH} + (1 - \sigma)[\mu t_{Hl} + (1 - \mu)t_{hH}]. \tag{4}$$

The first term represents the case in which S_A stays with the firm, which occurs with probability σ. Training increases J_A's productivity from q_l to q_h, and both S_A and J_A acquire specific human capital, resulting in team productivity t_{HH}. The second term represents the case in which S_A quits. If J_A has the potential to be a senior (with probability μ), then J_A is promoted (with productivity ϕq_h as a senior) and is then replaced by a rookie, resulting in team productivity t_{Hl}. If J_A has no potential, then a new S_A is hired from outside (with productivity q_h), which leads to a t_{hH} team. Assuming $x_A = 1$, expected output then is $y^0 = t^0 + \theta$.

As explained above, a lattice differs from silos in two states, one in which J_A is promoted to S_B, which causes productivity loss y_L in division A relative to the silo case, and another in which S_A leaves and is replaced by J_B instead of an outside hire, which leads to a productivity gain y_G. Adding the effects of these two events leads to expected team productivity

$$t^L = t^0 + \sigma\mu v(y_G - y_L), \tag{5}$$

and expected output then is $y^L = t^L + \theta$.

Wage cost: With silos, each division produces expected output y^0, and without any transfers of juniors, S_A's expected wage from the firm's perspective is $w^0 = \alpha + (\beta + \gamma)y^0$, where α is S_A's salary and β and γ are the bonuses for own- and other-division output, respectively.

With a lattice, a key expression in S_A's wage, which also appears in the incentive constraints derived below, is the difference in S_A's pay if J_A is promoted to S_B, relative to the silo case in which a senior vacancy in division B is filled from outside. Assuming execution effort $x_A = x_B = 1$ in equilibrium, it is given by

$$g = \alpha_t + \beta_t(t_{Hl} + \theta) + \gamma_t(t_{HH} + \theta) - \alpha_n - \beta_n(t_{HH} + \theta) - \gamma_n(t_{hH} + \theta). \tag{6}$$

In (6), different contract parameters apply depending on whether or not J_A is transferred ($\alpha_t, \beta_t, \gamma_t$ vs. $\alpha_n, \beta_n, \gamma_n$). If J_A is promoted, S_A needs to hire a rookie, whereas division B gets a senior with specific human capital, hence the team productivities t_{Hl} for division A and t_{HH} for B. If J_A stays in division A, a new S_B has to be hired from outside, hence the team productivities t_{HH} for division A and t_{hH} for B.

We can now write S_A's expected wage as

$$w^L = w^0 + \sigma\mu v[g + (\beta_n y_G - \gamma_n y_L)]. \tag{7}$$

In (7), the second term describes how a lattice differs from silos. The first term in square brackets, g, is the case just discussed, where J_A is transferred to division B. The second term is the reverse situation, where S_A leaves and is replaced by J_B. The new S_A inherits the no-transfer terms of the contract, and the impact on the firm's wage bill for division A is β_n times the productivity gain y_G in division A, minus the γ_n times the loss y_L in division B.

4.2 Incentive constraints and the firm's contracting problem

The firm needs to provide incentives for the senior managers to invest in both training and execution effort. Moreover, if the firm chooses to implement a lattice, a senior manager must have an incentive to truthfully communicate his junior's potential in case of a vacancy in the other division.

Execution incentives: S_A's execution effort x_A at stage 6 of the game increases his own division's expected output y_A additively by θx_A, for which he is paid βY_A (recall $Y_A \in \{0,1\}$) in the silo case, and in the lattice case $\beta_n Y_A$ if J_A is not transferred and $\beta_t Y_A$ if J_A is transferred. Given the execution effort cost ξ, it is optimal for S_A to choose $x_A = 1$ in each case if and only if

$$\beta, \beta_n, \beta_t \geq \frac{\xi}{\theta} \equiv \beta^x. \tag{8}$$

Training incentives: Manager S_A's expected utility at stage 1 of the game is his wage minus the cost of training and execution effort, taking into account that with probability $1-\sigma$ he will quit (and then receive $\underline{U}=0$). Assuming $e_B = x_A = x_B = 1$, it is given by

$$v_A(e_A, e_B) = \sigma\{\alpha_n + \beta_n[e_A t_{HH} + (1-e_A)t_{HL} + \theta] + \gamma_n y^0 - \xi\} + L\sigma\mu v g e_A - \tau e_A, \tag{9}$$

with g as defined in (6), and with $L \in \{0,1\}$ depending on whether the firm chooses silos or a lattice. The first term in (9) corresponds to the silo (no-transfer) case. An S_A hired at stage 1 receives the payments $(\alpha_n, \beta_n, \gamma_n)$ only if he stays (with probability σ). Division A's expected output y_A depends on his training effort e_A, whereas, assuming $e_B = x_B = 1$, B's output y_B is simply y^0. As in the wage equation (7), the term $\sigma\mu v g$ is the difference in S_A's wage if J_A is transferred (promoted to S_B), whereas (unlike in (7)) the reverse case is irrelevant to S_A. The costs of training τ and execution effort ξ enter differently in (9) because S_A invests training effort before he might quit, but invests execution effort only if he stays with the firm. Differentiating (9) with respect to e_A, we obtain that S_A invests in training at stage 3 of the game if

$$f = \sigma\beta_n\kappa_J\phi\Delta q + L\sigma\mu v g - \tau \geq 0, \tag{10}$$

where the first term represents $\partial(t_{HH} - t_{HL})/\partial e_A$.

Truthtelling incentives: With a lattice, the CEO requires information from the seniors about their juniors, because per conditions (INT) and (OUT), she will promote a junior if and only if she knows that he has potential. As discussed in Section 3.2, a senior who stays could gain from lying, whereas no such incentive exists for a departing senior. To express the relevant truthtelling constraints, we can focus on the case in which there is a vacancy in division B that could potentially be filled by J_A.

If J_A has potential and S_A reports that to the CEO, then S_A stands to lose a productive junior. Assuming $x_A = x_B = 1$, the resulting difference in pay compared to hoarding J_A is g as defined in (6), and thus the truthtelling constraint to prevent hoarding of a good junior is

$$g \geq 0. \tag{11}$$

If J_A does not have potential (but still has productivity ϕq_h as junior), then promoting J_A to S_B is doubly inefficient, because J_A would be a bad senior manager and because a rookie would need to replace him. Nevertheless, "kicking upstairs" a junior without potential could still benefit S_A personally, depending on the incentives provided. If S_A reports truthfully, then J_A is not transferred, and S_A's wage is $\alpha_n + \beta_n(t_{HH} + \theta) + \gamma_n(t_{hH} + \theta)$, where the last term reflects the fact that in division B, a senior is hired from outside. If S_A lies and reports having a junior with potential who is then promoted, then his resulting wage is $\alpha_t + \beta_t(t_{Hl} + \theta) + \gamma_t(t_{LH} + \theta)$. Thus, assuming $x_A = x_B = 1$, the truthtelling constraint to prevent kicking upstairs a junior without potential is

$$h = \alpha_n + \beta_n(t_{HH} + \theta) + \gamma_n(t_{hH} + \theta) - \alpha_t - \beta_t(t_{Hl} + \theta) - \gamma_t(t_{LH} + \theta) \geq 0. \tag{12}$$

The firm's talent management problem: Is it best for the firm to implement a lattice ($L = 1$), or are silos ($L = 0$) optimal because the truthtelling constraints (11) and (12) are not required? Denote expected (per-silo) profits for each case by $\pi^0 = y^0 - w^0$ and $\pi^L = y^L - w^L$, respectively, define the expected profit for endogenous L as $\pi = \pi^0 + L(\pi^L - \pi^0)$, and recall that $\zeta = (\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t)$ denotes the vector of contract parameters. Then the firm's optimization problem can be written as:

$$\max_{L \in \{0,1\}, \zeta} \pi \text{ s.t. } \zeta \geq 0;\ \beta_n, \beta_t \geq \beta^x;\ f \geq 0;\ Lg \geq 0;\ Lh \geq 0. \tag{13}$$

In (13), the constraint $\zeta \geq 0$ follows from limited liability. The next two constraints are effort incentive constraints for execution and training, respectively, and need to be satisfied under either incentive system. Finally, the constraints $Lg \geq 0$ and $Lh \geq 0$ state that the truthtelling constraints $g, h \geq 0$ are required if and only if a lattice is implemented.

4.3 Silos: narrow incentives and talent hoarding

We first analyze the silos case, in which case a contract is characterized by $\zeta = (\alpha, \beta, \gamma)$. Without cross-divisional transfers, there is no link between the two divisions, and therefore no reason to pay managers based on the other division's performance. Moreover, with limited liability and $\underline{U} = 0$, it is optimal to set $\alpha = 0$. All that is needed is an own-division bonus large enough to incentivize both training and execution effort.

Proposition 1 *Conditional on $L = 0$ (the firm implements silos), the minimal own-division bonus β that induces $e_i = 1$ is given by*

$$\beta^S = \frac{\tau}{\sigma \kappa_J \phi \Delta q}. \tag{14}$$

An optimal silo contract is given by $\alpha = \gamma = 0$ and $\beta^0 = \max\{\beta^S, \beta^x\}$. The expected wage cost for each manager is $\beta^0 y^0 > w^{FB}$.

All proofs are in the Appendix. The expression for β^S is intuitive. As in any model with binary effort, effort is easier to induce—and thus the required bonus is smaller—the smaller the cost of (training) effort τ and the larger the value of effort $\kappa_J \phi \Delta q$. In addition, effort is easier to induce the larger σ (the probability that S_A stays) is, because the senior benefits from a good junior only if he stays with the firm. The firm's wage cost exceeds the first-best level (3), and S_A thus earns a rent, because division output is a noisy measure of S_A's investment in training.

Importantly, the link between managerial incentives and talent management goes both ways. As shown, simple, narrow incentives are optimal when the firm chooses to forgo lateral mobility and relies on vertical promotions only. However, the converse also holds: A firm that provides narrow incentives de facto ends up with silos even if ostensibly it supports lateral mobility. Recall that we have assumed that the CEO retains the decision rights over how to fill positions. Suppose then, with the silo contract of Proposition 1 in place, that the CEO considers promoting J_A to fill an S_B-vacancy, and asks S_A about J_A's potential. S_A will then report that J_A has no potential, irrespective of whether that's the case. That is, even if J_A has the potential to be a senior, then S_A will hoard J_A by lying to the CEO. That's because with the contract of Proposition 1, the wage difference g in (11), using $\beta_n = \beta_t = \beta$

etc., reduces to $-\beta y_L < 0$, reflecting S_A's loss of a good junior without offsetting benefit. Without truthful information about J_A, however, the CEO would never promote him.

It follows, then, that traditional internal labor markets with vertical job ladders and without lateral mobility are an equilibrium consequence of providing narrow incentives to managers. They will persist even if firms officially promote lateral mobility, as long as bosses' incentives are unchanged, and as long as the firm has to rely on bosses' recommendations about whom to promote. See Section 5.1 for further discussion.

4.4 Lattice: Talent management with cross-divisional transfers

Implementing a lattice ($L = 1$) requires a combination of broader incentives than with silos, as well as transfer-contingent pay. Before we state our main result, it is useful to consider the full-information case first.

Full-information contract: Suppose that transfer-contingent wages are feasible, but that the CEO has full information, so that the truthtelling conditions (11) and (12) are not required.

Proposition 2 *An optimal full-information contract ζ is given by $\alpha_n = \alpha_t = 0$, $\beta_n = \beta_t = \beta^x$, $\gamma_n = 0$, and*

$$\gamma_t = \frac{\frac{\kappa_J \phi \Delta q}{\mu v}(\beta^S - \beta^x) + y_L \beta^x}{t_{HH} + \theta}, \tag{15}$$

where β^S is the optimal silo bonus according to Proposition 1. That is, the optimal own-division bonus equals the minimum required to ensure execution effort, and a bonus for the other division's output is paid if and only if J_A is transferred (promoted).

The intuition for this result carries over to our main result below. First, it is optimal to set both the fixed payment α_n and the other-division bonus when there is no transfer γ_n to zero as these payments both increase the wage and *reduce* training incentives. More precisely, these two payments have no effect on S_A's incentive to train J_A to have a good junior on his team. However, being payments conditional on the non-transfer of J_A, they reduce S_A's net reward g in (6) from the transfer of a good J_A to division B.

Second, the transfer-contingent bonuses α_t, β_t, and γ_t are all equivalent in terms of their impact on the firm's profit and on training incentives, and β_t additionally has to satisfy (8) to incentivize execution effort. The optimal contract is therefore not unique, but one solution is to set $\alpha_t = 0$ and $\beta_t = \beta^x$, their minimal feasible values.

Third and most importantly, it is more efficient to incentivize training by rewarding the transfer of a junior than by rewarding own-division performance. With full information, only a J_A with potential is ever transferred, which makes this event a noiseless signal of S_A's training effort. Because the managers are risk neutral, it is therefore optimal to focus training incentives on this event. By contrast, incentivizing training through β_n not only leaves a rent to S_A like in the silo case, but in fact provides worse incentives than in the silo case because S_A may lose J_A to the other division. The optimal solution involves setting β_n as small as possible, which is $\beta_n = \beta^x$, and then to set γ_t to satisfy the training incentive constraint (10), which leads to (15). This solution is shown in Figure 2, in which the feasible set of contracts is the shaded area defined by $\beta_n \geq \beta^x$ and $f \geq 0$. The result that it is cheaper to incentivize training through γ_t than through β_n is reflected in iso-profit lines in β_n-γ_t space that are steeper than the training incentive constraint $f \geq 0$.

Indeed, if the execution effort cost ξ equals zero, the optimal own-division bonuses are $\beta_n = \beta_t = 0$, and the firm can attain the first-best by paying a (large) other-division bonus $\gamma_t = \frac{\tau}{\sigma\mu v(t_{HH}+\theta)}$ only in the event of a transfer. The wage according to (7) then reduces to τ, which in this case equals the first-best wage, leaving no rent to S_A.

The opposite conclusion holds if the execution cost ξ is large, that is, if running a division is costly relative to training a junior. For instance, if $\xi = \theta\beta^S$, and consequently $\beta^x = \beta^S$ is the lower bound for both β_n and β_t, then having to pay an other-division bonus $\gamma_t > 0$ in addition to satisfy the training incentive constraint unambiguously raises the wage cost above that of silos.

Optimal contract: We can now state the solution to (13) under the constraint $L = 1$, which is the cost-minimizing contract that implements a lattice when the division managers have private information about their juniors.

Proposition 3 *An optimal lattice contract is characterized by $\alpha_n = \alpha_t = \gamma_n = 0$ and $\beta_t = \beta^x$. For these values, define $(\tilde{\beta}, \tilde{\gamma})$ as the solution in β_n, γ_t to the equations $f = h = 0$. Then the optimal values of β_n, γ_t are as follows:*

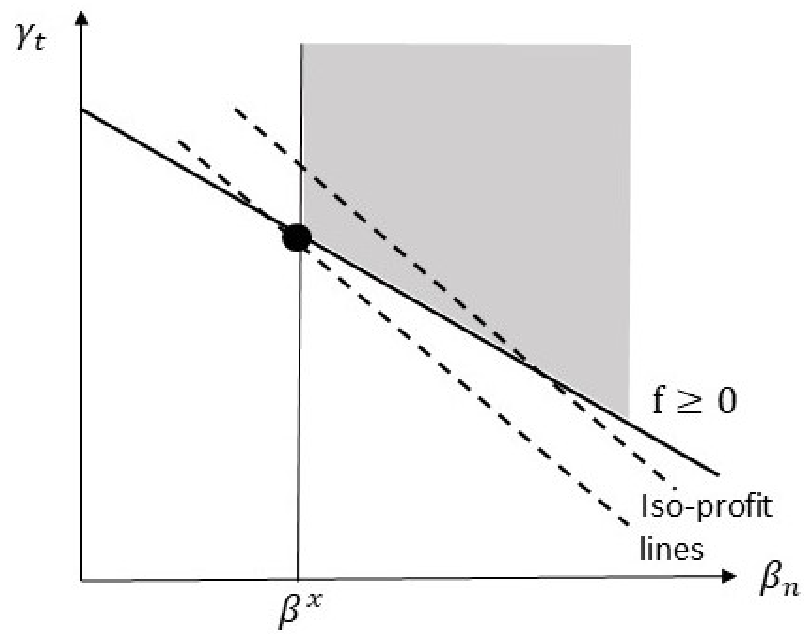

Figure 2: Illustration of Proposition 2

(1) If $\beta^x \leq \tilde{\beta}$, then $\beta_n = \tilde{\beta}$ and $\gamma_t = \tilde{\gamma}$. That is, if $\beta^x = \xi/\theta$ is low, i.e., the cost of execution effort is low or its value high, then β_n and γ_t solve the training incentive constraint and the no-kicking-upstairs constraint. The no-hoarding constraint is slack.

(2) If $\beta^x \in (\tilde{\beta}, \beta^S]$, then $\beta_n = \beta^x$, and γ_t solves $f = 0$. That is, if ξ/θ is in an intermediate range, then both truthtelling constraints are slack, and the optimal β_n and γ_t correspond to the full-information case.

(3) If $\beta^x > \beta^S$, then $\beta_n = \beta^x$, and γ_t solves $g = 0$. That is, if ξ/θ is large, the own-division bonus ensures execution effort, the no-hoarding constraint is binding, and the training incentive constraint is slack.

The key intuition of Proposition 2 continues to hold: It is more efficient to incentivize training with a transfer-contingent reward than with an own-division bonus, because the event of a transfer is perfect information about S_A's investment in training. However, with privately informed senior managers, the constraints (11) and (12) are required to establish truthful communication. As the proof shows, out of the transfer-contingent bonuses α_t, β_t, and γ_t, only $\gamma_t > 0$ can satisfy both truthtelling constraints. Intuitively, only a reward tied to a

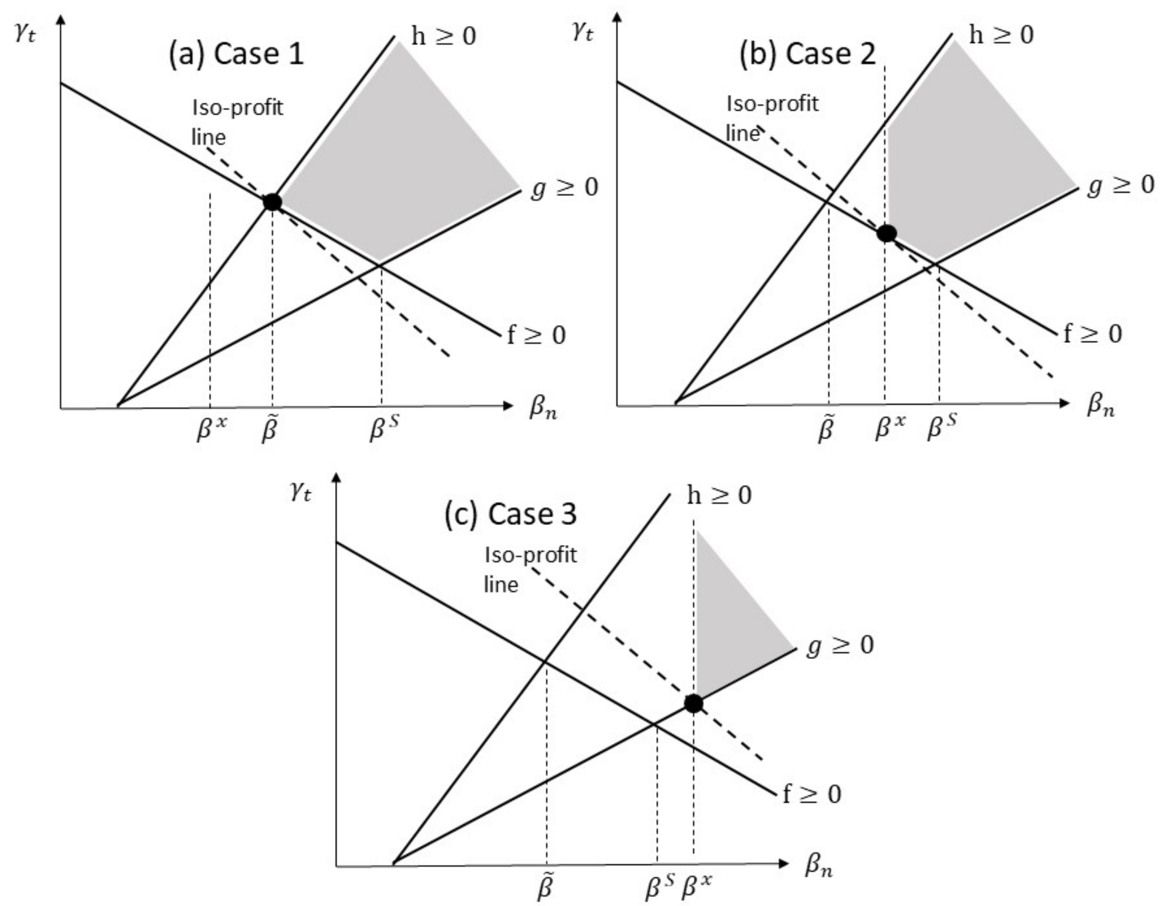

Figure 3: Illustration of Proposition 3

promoted J_A's performance as S_B can simultaneously encourage S_A to recommend J_A if he has potential, while preventing S_A from recommending J_A if he does not have potential. By contrast, an unconditional referral fee $\alpha_t > 0$ could tempt S_A to falsely claim that J_A has potential. It is therefore strictly optimal to set $\alpha_t = 0$ and $\beta_t = \beta^x$.

What remains is to determine the optimal values of β_n and γ_t, and the resulting cases are shown in Figure 3. Like Figure 2, each panel in Figure 3 shows the training incentive constraint and a representative iso-profit line. Both are downward-sloping, with the iso-profit line being steeper as explained for Proposition 2 above.

For S_A to truthfully reveal that J_A has potential requires a minimal γ_t that increases with β_n (the $g \geq 0$-line). For S_A to truthfully reveal that J_A does not have potential requires a maximal γ_t that also increases with β_n (the $h \geq 0$-line). The training incentive constraint

(10) intersects with the no-hoarding constraint (11) at β^S, because if $g = 0$, then (10) reduces to the silo case.

Panel (a) shows the optimal contract when β^x is smaller (Case 1 of Proposition 3). Like in the full-information case, it is optimal to incentivize training as much as possible through γ_t, but that option is constrained by the truthtelling constraint (12). That is, S_A must care enough about his own division (through β_n) and not too much about division B (rewarded through γ_t), to prevent him from kicking upstairs a J_A without potential.

For an intermediate range of β^x (Case 2, panel b), the optimal solution coincides with the full-information contract: The execution effort constraint $\beta_n \geq \beta^x$ is binding, γ_t is chosen to satisfy the training incentive constraint, and both truthtelling constraints are slack.

Finally, if β^x exceeds β^S (Case 3, panel c), then, as in the silo case, setting $\beta_n = \beta^x$ more than suffices to incentivize training, but a sufficiently large γ_t is still needed to satisfy (11), i.e., to prevent talent hoarding.

Simple output-based contracts: Transfer-contingent compensation like in Proposition 3 adds a layer of complexity that may be difficult to implement in some firms; see our discussion below in Section 5.2. Can the firm implement a lattice without transfer-contingent rewards? Yes, but it's likely to be costly, as our next result shows.

Proposition 4 *Assume that only output-based contracts are feasible: $w_i = \alpha + \beta y_i + \gamma y_j$. Then if*

$$\kappa_J \phi \Delta q \geq \mu v(y_G + y_L), \tag{16}$$

that is, if the benefit of increasing J_A's productivity exceeds the marginal value of J_A as replacement for S_B, then the optimal contract that induces training and execution effort is given by $\alpha = 0$, $\beta = \beta^0 = \max\{\beta^S, \beta^x\}$ and $\gamma = \frac{y_L}{y_G}\beta$. The firm's expected wage bill is then higher than with silos.

If (16) does not hold, then the optimal contract is given by $\alpha = 0$ and $\beta = \max\{\beta^x, \hat{\beta}\}$, where $\hat{\beta}$ is part of the solution to $f = 0$ and $h = 0$ in β, γ. Finally, the optimal γ solves $f = 0$ for $\beta = \max\{\beta^x, \hat{\beta}\}$.

Whether or not (16) holds, the wage cost is strictly higher than with the optimal transfer-contingent contract according to Proposition 3.

Condition (16) establishes an upper bound on the probability μv that J_A would be promoted to S_B, conditional on S_A staying with the firm. It ensures that the firm's main benefit from investing in training J_A is to have a productive junior in division A, and not the marginal benefit of developing a second internal candidate (besides J_B) to replace a departing S_B. The optimal contract then incentivizes training through the own-division bonus β, and pays $\gamma > 0$ only to satisfy the no-hoarding constraint (11). Moreover, since satisfying (11) with equality means $g = 0$ (the no-hoarding condition is binding), then (10) reduces to the silo case, and the optimal β equals the silo bonus, unless β^x is larger. The expected wage cost is higher than with silos, both because of the increase in productivity if $L = 1$ (since the cost is proportional to expected output), and because the firm pays $\gamma > 0$ on top of the silo bonus β^S.

Condition (16) may not hold if for J_A, the probability μv of a senior vacancy in the other division is larger. Then, reminiscent of the logic of Proposition 2, it is optimal to incentivize training through γ. The optimal β is then either the minimal value β^x required for execution effort, or (only if $\phi q_l > q_h$) possibly a higher value required to satisfy the no-kicking-upstairs constraint (12) in addition to the training incentive constraint.

The result that a transfer-contingent contract always strictly dominates is straightforward: There is no value in paying a senior for the other division's performance unless it's to hold him accountable for his promoted junior's performance. It is therefore strictly optimal, starting from a contract with $\beta_n = \beta_t = \beta$ and $\gamma_n = \gamma_t = \gamma$, to reduce γ_n and raise β_n, which keeps all constraints satisfied while strictly saving money.

4.5 Silos or lattice?

The key message of our paper is that although a lattice leads to more efficient personnel allocations than do traditional silos, implementing one requires the cooperation of bosses as the gatekeepers in an internal labor market. In our model, this requirement takes the form of incentive constraints that may raise the firm's cost of paying senior managers above the silo level. As we have seen, however, tying a senior's wages to the promotion of a junior can also *lower* the wage cost relative to silos because in equilibrium, a promotion is a perfect signal of the senior's training investment.

Three cases are therefore possible: First, a lattice is unambiguously preferred to silos if the wage cost is lower with a lattice. Second, a lattice may still be preferred to silos if the wage cost is higher, but the productivity gain outweighs the cost increase. Finally, however, the incentive costs of implementing a lattice may be too large compared to the productivity gain, in which case the firm would optimally choose silos.

Our final result shows that the firm's optimal choice between silos and a lattice is systematically driven by the effort cost parameters and the position-specific productivity coefficients for each manager:

Proposition 5 *The difference $\Delta\pi = \pi^L - \pi^0$ between the firm's profit with a lattice and with silos is (a)) weakly increasing in the training cost τ, (b) strictly decreasing in the cost of execution effort ξ, (c) strictly increasing in the importance of a senior manager κ_S, and (d) strictly decreasing in the importance of a junior manager κ_J if condition (16) holds.*

A larger training cost τ favors a lattice because, as explained for Proposition 2, training is cheaper to incentivize with a lattice than with silos (even if the additional truthtelling constraints add to the lattice wage costs). Since the silo bonus β^S is proportional to τ (and $\beta_n \leq \beta^S$), it follows that τ increases the wage cost more with silos than with a lattice. A larger execution effort cost ξ tightens the execution effort constraint (8) and requires a larger β_t and possibly β_n. It thus reduces the firm's ability to load incentives on γ_t through which the rents paid to the managers can be saved.

A higher productivity of a senior manager κ_S raises the allocative benefit y_G of promoting an insider over hiring from outside. It also shifts the wage cost in favor of a lattice. Finally, a higher productivity of a junior manager κ_J has the opposite effect, except that sufficient condition (16) is required because the "no kicking upstairs" truthtelling constraint (12), which is binding in case 1 of Proposition 3, is more difficult to satisfy with a larger κ_J.

The effects of other parameters on $\Delta\pi$ are ambiguous, especially when the effects on π^L and π^0 individually are unambiguous. For example, a larger specific human capital factor ϕ increases the allocative benefit $y_G - y_L$ of a lattice, but also reduces the agency costs of silos by more than the costs of a lattice. Similarly, a larger value of execution effort θ favors a lattice by saving on own-division performance pay when there is no transfer ($\beta_n(t^0 + \theta)$ vs.

$\beta^S(t^0+\theta))$, but also (like for κ_J just discussed) tightens constraint (12).

5 Discussion

Our paper makes two contributions. The first is to develop a theoretical framework to investigate the problem of talent hoarding that has received substantial interest in strategic management and is of great importance for practitioners, but has not been treated theoretically in economics. The second is to suggest an incentive-contracting solution that prevents talent hoarding while preserving incentives for training. This solution, however, may entail agency costs that exceed the benefits of cross-divisional mobility. It follows that for some firms, traditional vertical job ladders may be optimal for incentive reasons.

Hard evidence on reward systems for talent management is hard to come by, but below we discuss one example and how it fits the model. We also discuss other ways to deal with the problem, and conclude with the perspective of workers (in our model, the junior managers), which our formal analysis has omitted. First, though, we discuss what happens when a firm wants to implement a lattice without taking into consideration the managerial incentive issues that are at the core of our theory.

5.1 Building a lattice without supporting incentives?

Haegele (2024) presents the first evidence of talent hoarding in economics, based on data from a large German manufacturing firm. Haegele's focus firm shares features of many other companies, suggesting that talent hoarding may be widespread, consistent with the management literature on the topic, see Keller and Dlugos (2023).

Vacancies in Haegele's focus firm are posted on a job portal through which interested employees can apply. But while the firm encourages internal applications, in practice mobility is limited because employees refrain from applying for positions in other units for fear of retaliation from their bosses. In the light of our model, these observations, which according to Keller and Dlugos (2023) are common, come as no surprise. When a firm's managers are rewarded for their team's performance but not for talent development, managers prefer to keep good people on their team. This is the situation we discussed in Section 4.3: Even if a firm officially supports cross-divisional mobility, in equilibrium no junior with potential is ever recommended for a transfer as long as their bosses are rewarded for own-division performance only, because the truthtelling constraint (11) is violated. Incidentally, under-investment in training is *not* a problem in this case, because managers who can hide good people don't have to worry about losing them to other units.

A very different example is Tencent, a Chinese tech giant with 100,000 employees in six divisions, which has has been operating a job transfer program since 2011 but without any adjustment of rewards for bosses. When it emerged over time that bosses' resistance to team members moving elsewhere was both common and strong—including retaliatory poor performance ratings and pressure to finalize additional work—, the company decided to enable workers to apply for new positions without the knowledge of their boss. That works, of course, only when workers can make a case for themselves (including through interviews) and do not require their boss's personal recommendation, unlike in our model and in Haegele's (2024) focus firm. Even so, and like in our model, Tencent does not allow workers to apply for a promotion in a different division; they must transfer laterally first before being considered for a vertical promotion. For all details on this example, see Qing et al. (2021). For further discussion of "internal talent markets" like Tencent's, see Cowgill et al. (2023, 2024).

5.2 Accountability and rewards for talent development in practice

There is a general consensus that successful talent management requires accountability and supporting reward systems for managers (i4cp, 2016). Putting that idea into practice is easier for some firms than for others. The report of i4cp (2016) singles out engineering company Fluor as an example of a company that has increased managers' accountability for talent development "by tying executive performance and compensation to how well managers develop their direct reports, even if it means moving their best people into positions with other divisions." Some consulting firms are known to put emphasis on talent development when deciding on promotion cases, for instance, NERA and Capgemini.

A key feature of the contract of Proposition 3 is that managers are not just rewarded for the transfer of their mentees, but are also held accountable for the mentees' performance in their new position, to prevent managers from "kicking upstairs" less qualified people. This is easiest to do for firms whose units are sufficiently similar to allow for firm-wide standardized systems of performance measurement, such as in retail chains.

At juice bar chain Joe & the Juice, for instance, employees complete standardized training modules but also receive a significant part of their training through mentoring by their bosses. Their IT system

> "not only tracked employees' grades on modules, but also trained them. This...made it easy for country managers to see how well a regional manager's former employees performed once they were promoted away from the region, a factor that the country manager took into account when evaluating the regional managers." (Rouen and Srinivasan, 2019).

An example of a company whose incentive system is strikingly close to the optimal contract of Proposition 3 is the Chinese hot pot restaurant chain Haidilao, which has 1300 restaurants worldwide (for all details of this example, see Zheng and Zhao, 2018, and Campbell et al., 2023). In 2011, the chain's growth was inhibited by the lack of qualified restaurant managers, who could learn the requisite skills only on the job, as protégés of existing managers. CEO Zhang Yong recognized, however, that managers had little incentive to mentor people who would not only leave in due course, but who would become competitors of the mentor.

Zhang put in place a unique reward system designed to incentivize managers to train protégés and to establish "families" of restaurants run by fully-trained protégés of a manager. Specifically, managers could choose an incentive plan that paid 0.4% of the own restaurant's profit, 3.1% of profits of restaurants opened by the manager's protégés', and 1.5% of second-generation protégés' restaurants. Alternatively, they could choose a plan that just paid 2.8% of the own restaurant's profit (Zheng and Zhao, 2018). Although the company lets managers choose between these two plans, it is clear that the first one is relevant only if the manager already has a "child" restaurant, whereas the simpler plan dominates for managers whose first protégé is still training, or who do not intend to train new managers.

Given the managers' dominant choices, Haidilao's incentive scheme in effect emulates the contract of Proposition 3: First, rewards are contingent on the establishment of a family, i.e. the transfer of a junior. In particular, a bonus based on another restaurant's performance is paid only when a protégé has opened a restaurant ($\gamma_n = 0$, $\gamma_t > 0$). Second, the establishment of a protégé's restaurant and the subsequent change in incentives not just adds a bonus, but *shifts* incentives from own-restaurant to protégé-restaurant performance, much like in Proposition 3 when ξ is small: $3.1\% > 2.8\% > 0.4\%$ means, in the model's notation, $\gamma_t > \beta_n \gg \beta_t$, consistent with Proposition 3.

What works for chains with largely homogeneous units is arguably harder to implement in firms with heterogeneous divisions and functions. One obstacle is that the impact of a promoted manager's ability on her division's performance is spread out in time and difficult to measure, a problem that goes beyond our static setting. Indeed, the delay between employees' actions and measurable outcomes is one of the reasons why internal labor markets exist in the first place (see Milgrom and Roberts, 1992, pp. 363-364). In a dynamic setting, if J_A is promoted to S_B at time τ, the referring manager S_A would need to receive the bonuses β_t and γ_t during some window $[\tau + m, \tau + n]$. Determining the appropriate window may be difficult: Rewards spread out over a long time might fail to incentive managers whose time horizons are shorter, whereas paying high rewards during a short window might create incentives *for the firm* to renege on its obligations by manipulating the event of a transfer. For instance, if a manager refers a good employee for a vacant position, hoping to receive a high reward, top management could claim that the employee is unqualified but transfer him later, ostensibly for reasons unrelated to the manager's referral.

Another potential obstacle is the internal accounting for bonuses paid for another unit's performance. Clearly, division B would not want to pay a bonus to manager S_A out of its own budget. However, division A, too, would be reluctant to pay S_A a bonus based on a different division whose performance and performance measurement it cannot control. In practice, a common solution to internalize positive externalities between divisions is to pay managers based on *corporate* performance (Bushman et al., 1995; Alok and Gopalan, 2018). That is equivalent to paying an other-division bonus such as γ in our model, with one caveat: The wage $w = \beta y_A + \gamma y_B$ in our model can be expressed as $w = (\beta - \gamma)y_A + \gamma(y_A + y_B)$, with γ applied to corporate performance. But if the remaining own-division bonus $\beta - \gamma$ has to be nonnegative, we obtain $\gamma \leq \beta$ as an additional constraint. This constraint is likely to be satisfied in Proposition 3 if execution effort is costly and hence β_n, β_t are relatively large, but may be binding if execution effort is low-cost relative to training effort.

But such implementation and calibration challenges are common to the design of explicit incentives in general, and don't preclude their use. Short of specifying a mentee's "output," for instance, firms could devise scorecard measures of managers' training and development efforts that incorporate data on numbers of promoted mentees, salaries in the new positions, and tenure in those positions. Those measures are still outcome-based and are thus quite different from mentees' subjective evaluations of their bosses' career development efforts, which we discuss next. Furthermore, the example of Haidilao shows that paying managers baseed on corporate performance is not the only way to reward externalities across units, and that bonuses that reward externalities more highly than own performance exist in practice.

5.3 Alternative Solutions

We confined our formal analysis to the design of incentive contracts of bosses. In reality, of course, firms' efforts to organize their internal labor markets are more complex than that. In this section, we discuss potential and actual alternative solutions.

Direct monitoring of training effort: All solutions discussed so far provide outcome-based incentives for training, that is, rewards based on outcomes for mentees, such as promotions. Some firms choose to measure and reward training effort directly, such as instrument maker Agilent, whose top managers receive a third of their compensation based on HR development efforts (Conaty and Charan, 2010, p. 160). In the tech firm of Hoffman and Tadelis (2021), managers with high people scores are rewarded by promotions. One of the dimensions of the overall management score is whether a manager supports the career of the employee. One way to measure training effort is through 360-degree reviews, in which subordinates are asked about their boss's training efforts. However, such reviews are costly—at MFS Investment Management, they take up significant time of C-level executives (Hall and Lim, 2002)—and, of course, subjective and hence subject to many sources of error.

Job rotation: General Electric, Novartis, SAS Institute and many other companies have programs in which junior and mid-career managers go through different assignments across functions and divisions. The primary objective of job rotation is usually to develop junior employees' skills and to prepare them for higher-level positions (e.g., Conaty and Charan, 2010, 228-229).[15] One might think that job rotation also solves several problems emphasized in this paper: Employees can develop their skills through the rotation program rather than having to rely on training by their boss; bosses can no longer "hoard" good people; and talented employees become more visible to senior managers in different units, reducing the adverse-selection problem created by bosses' private information.

However, job rotation does not diminish the importance of training by managers; it only spreads the training responsibility across multiple managers. While alleviating the adverse-selection problem, it makes the moral-hazard problem worse: Managers have even less incentive to train someone who will leave their unit not with some probability, but for sure! Another downside of job rotation, though not captured in our model, is that while employees acquire general managerial skills, they may lack the time and motivation to acquire division- (or industry-)specific expertise. After many years of emphasizing general managerial skills, General Electric decided in 2012 to decided to keep senior managers in their divisions longer to help them acquire the expertise needed to compete in their industries (Linebaugh, 2012). To conclude this point, job rotation has its advantages but has costs as well, and is therefore a good solution for some firms but not others.

Other leadership development initiatives: Other initiatives, too, simultaneously foster human capital development and generate firm-wide information about talented employees. General Electric's Management Development Institute in Crotonville, NY (which GE sold in 2024), used to run leadership courses and "workout" sessions in which high-potential employees are encouraged to publicly challenge their bosses' views (Martin and Schmidt, 2010). Likewise, at Maersk Group, the shipping conglomerate, top executives hold "people strategy sessions" in which they review the top 120 positions in the company and the people in those positions to determine whether people are optimally matched to jobs (Groysberg and Abbott, 2013).

Many other firms have adopted similar initiatives, cf. Conaty and Charan (2010) and

[15] Minni (2023) shows how at one of the world's largest consumer product firms the rotation of managers, intended to broaden the managers' skills, also impacts the careers of those they manage.

Benko and Anderson (2010). One could argue that as a result, the extent of private information about top talent may be only small. Efforts for everyone to get to know everyone else are costly, however, and can realistically cover only the top tier of managers, perhaps 100-200 people. It follows that even though private information may not be an issue at the very top of a company, moving down the ranks it will eventually begin to matter (recall our remark in Section 3 that the tiers in our model may represent any adjacent tiers in a larger hierarchy).

In conclusion, job rotation and other leadership development initiatives have clear benefits but have disadvantages or costs as well. The general argument of our paper remains: achieving cross-divisional mobility requires costly solutions that address the key role that bosses (at least traditionally) play as mentors and as holders of private information.

5.4 Cross-divisional Mobility from the Workers' Perspective

Our analysis has focused squarely on the role of middle managers for the production and allocation of human capital formation in firms, and has treated the workers as entirely passive. We have argued that the managers' role is critical to why silos have been so prevalent throughout business history, and why for many firms, transitioning from silos to a market is not that easy.

That said, firms' efforts to improve mobility in internal labor markets are significantly driven by the workers' interests as well, as Bryan et al. (2006) point out:

> "Many a frustrated manager has searched in vain for the right person for a particular job, knowing that he or she works somewhere in the company. And many talented people have had the experience of getting stuck in a dead-end corner of the company, never finding the right experiences and challenges to grow, and, finally, bailing out (Bryan et al., 2006)."

Cowgill et al. (2023) discuss how firms can set up "internal talent markets" to overcome this problem, and Cowgill et al. (2024) examine the tradeoffs firms face between satisfying worker preferences and their desire for growth on one hand, and the firm's preference for high-productivity matches on the other.

An early example of conflicting interests is the case of Johnson & Johnson, whose divisions had traditionally been very independent. Pearson and Hurstak (1992) describe the negative consequences of J&J's silo structure from the workers' perspective: "Many junior executives found it tough to move up when young presidents stood in the way, and tougher still to jump over to a separate company [within J&J]." As early as the 1990s, CEO Ralph Larsen took measures to facilitate cross-company mobility in order to remedy J&J's "chronic problem of career-pathing." In line with our theory, however, these measures turned out to be unpopular with many senior managers, given the absence of adjustments to *their* incentives. Decades later, the example of Tencent discussed in Section 5.2 reflects the same tension. The company's internal transfer program offers its workers "infinite possibilities" for career development, and its managers understand the benefits for the company, but they also lack incentives to support the mobility of *their* people.

Aside from the allocational advantage explored in this paper, therefore, an internal market can also help retain talented workers and motivate them to invest in their own human capital, especially in skills required at higher levels. Our formal analysis does not cover this angle partly because the point is straightforward. The puzzle is why internal markets are hard to get to work, and we have argued that bosses are the key to that puzzle.

6 Conclusion

Traditional vertical job ladders in firms—the subject of a large economics literature—have recently been giving way to active "talent management" aimed at optimally matching people with positions. This includes efforts to increase the cross-divisional mobility of employees, i.e., to transform silos to a lattice. Our paper is the first to examine these efforts from an economics perspective.

We identify what we believe to be a major obstacle for firms' talent management: the incentives of bosses, who play a key role for training and allocating talent. These managers are conventionally focused on their own unit's performance, rather than on the firm's. A firm's efforts to increase cross-divisional mobility then undermines bosses' incentives to invest in training and causes them to strategically use their private information either by "hoarding" good or "kicking upstairs" bad employees.

The model captures the origins of these agency problems in modern firms. Because production takes place in teams, it is impossible to observe individual productivities. Also, training effort or its outcomes cannot be directly measured. In this setting, we show that the right contract can both provide incentives for training and prevent talent hoarding, but it may not be profit-maximizing for a firm to implement such a contract. Our results thus not only shed light on the challenges faced by companies transitioning to talent markets, they also highlight the rationale of having traditional internal labor markets that have often been criticized as silos. While silos lead to inefficient matches of people to positions, they create staightforward incentives for managers to train their employees.

The most important practical implication of our analysis is that establishing greater (cross-divisional) mobility for junior managers is not simply a matter of opening up new career paths, but also requires changes to the incentives provided to higher-level managers or other supporting practices such as job rotation, monitoring of training effort, and other development initiatives. Optimal incentives that account for both moral hazard and adverse selection at the manager level require giving managers a stake in the success of their transferred mentees. That may not always be easy to do, but is likely to become more feasible as firms continue to refine their performance metrics. Importantly, we show that transfer-contingent rewards may in fact lower rather than raise the agency costs for the firm.

Our setting offers a concrete and close-up view of the "selective-intervention puzzle" investigated in transaction-cost economics. Here, the question is what kinds of organizational costs counterbalance the benefit of bringing two business units under the umbrella of a common hierarchy. In our context, that benefit is to establish a market for talent between two otherwise independent divisions. Abstract discussions of the puzzle have emphasized agency problems resulting from integration that are unrelated to its benefits (Williamson, 1985, pp. 135-138; Tadelis and Williamson, 2013). Our results imply—contrary to these arguments that selective intervention does not necessarily weaken managers' incentives. Depending on how much of a stake in the benefits of the intervention can be given to the managers, it may strengthen incentives. This is in line with Marino and Zabojnik (2004) and Friebel and Raith (2010).

Our model thus stresses what we believe to be the problem in many organizations: the agents that are affected by integration are also central to its benefits. Furthermore, by putting more structure on the rationale for integration, we show that whether integration is efficient depends on what exactly can be measured and incentivized. If senior managers' rewards can be sufficiently tied to productivity-increasing transfers, then incentives with integration (a market) can in fact be stronger than without (silos), in which case integration is unambiguously optimal.

In studying the internal labor market of a multi-divisional firm, our paper departs from the literature and extends the reach of economic analysis to questions of importance to today's large companies (Conaty and Charan, 2010). It also reinforces the insight that the production and allocation of human capital in firms is not simply in the hands of "the firm" but significantly in the hands of its managers, whose interests may not align with their firm's. We hope that our theoretical results provide insights about the consequences for firm policies and advance the agenda proposed by Gibbons (2013) for organizational economics to "focus on what managers actually do," in the spirit of Cyert and March's (1963) insight that "managers devote much more time and energy to the problems of managing their coalition than they do to the problems of dealing with the outside world" (p. 205-6).

Appendix: Proofs

Proof of Proposition 1: Per (8), inducing $x_A = 1$ requires $\beta \geq \beta^x$. And given $L = 0$, (10) reduces to

$$\sigma\beta\kappa_J\phi\Delta q - \tau \geq 0, \tag{17}$$

leading to the stated expression for β^S. Thus, the smallest β that implements $e_A = x_A = 1$ is $\beta^0 = \max\{\beta^S, \beta^x\}$, and the optimality of $\alpha = \gamma = 0$ was discussed in the text.

The stage-1 expected division productivity is t^0, and because $\beta^0 Y_A$ is paid at the end of the game irrespective of whether the original S_A stays or leaves, the expected wage cost is $\beta^0(t^0 + \theta) = \beta^0 y^0$. To see that this cost strictly exceeds the first-best wage w^{FB}, note that since $\beta = \max\{\beta^S, \beta^x\}$, we have

$$\beta y^0 \geq \beta^S t^0 + \beta^x \theta.$$

The second term equals ξ, and for the first, we have $t^0 > \sigma t_{HH} > t_{hH} > \kappa_J \phi q_h$ and therefore $\beta^S t^0 > \frac{\tau}{\sigma} \frac{q_h}{\Delta q} > \frac{\tau}{\sigma}$. It follows that $\beta y^0 > \tau + \xi = w^{FB}$. ■

Recall from Sections 3.4 that we restrict attention to parameters for which it is optimal for the firm to incentivize both training and execution effort. To provide more detail on this assumption, consider the case $\beta^x \leq \beta^S$. Here, the execution incentive constraint is slack under Proposition 1, which means there are two alternatives that do not provide training incentives: The first is a contract that incentivizes execution but not training, and the optimal such contract is $(\alpha, \beta, \gamma) = (0, \beta^x, 0)$. Without training, the ex-ante expected productivity of a division is $\hat{t}^0 = \sigma t_{HL} + (1 - \sigma) t_{hL}$, where the second term reflects the fact that a manager who leaves is always replaced by an outsider because the junior hasn't been trained. The firm's expected profit in this case is $(1 - \beta^x)(\hat{t}^0 + \theta)$.

The second option is a contract that incentivizes neither training nor execution effort, and with $\underline{U}$ normalized to 0, a manager can be hired at wage zero, and the firm's expected profit in this case is simply $\hat{t}^0$. As an example, a set of parameters that meets all of our assumptions is $q_h = 0.4$, $q_l = 0.15$, $\kappa_S = 0.7$, $\kappa_J = 0.35$, $\phi = 1.7$, $\mu = 0.6$, $\sigma = 0.8$, $\tau = 0.01$, $\xi = 0.01$, $\theta = 0.25$. With these numbers, the firm's per-division profit with silos is 0.85. A contract that incentivizes only execution yields a profit of 0.745, and a contract that incentivizes neither training nor execution yields a profit of 0.53. A silo contract per Proposition 1 is therefore the most profitable of the three. (As it turns out, however, a lattice is still more profitable in this case. The example parameter set falls under case (1) of Proposition 3, and the resulting expected profit is 0.853 with a lattice instead of 0.850 with silos.)

More generally, increases in the training cost τ or the execution cost ξ weakly increase the firm's agency costs without affecting silo or lattice output, and without affecting the profit under contracts that do not incentivize training or execution. It is therefore straightforward that the optimality of incentivizing both training and execution amounts to upper bounds on τ and ξ.

Proof of Proposition 2: Provided that the contract implements $e_i = x_i = 1$ for $i = A, B$, the resulting expected output y^L does not depend on the contract ζ. The full-information version of (13) is therefore equivalent to $\min_\zeta w^L$ subject to the limited-liability constraint $\zeta \geq 0$, the execution incentive constraint (8), and the training incentive constraint (10).

1. It is optimal to set $\alpha_n = \gamma_n = 0$, because both variables increase the wage cost while reducing training incentives. That is,

$$
\begin{aligned}
\frac{\partial w^L}{\partial \alpha_n} &= 1 - \sigma\mu v > 0, \\
\frac{\partial w^L}{\partial \gamma_n} &= t^0 + \theta - \sigma\mu v(t_{hH} + \theta + y_L) \\
&= \sigma t_{HH} + (1-\sigma)\mu t_{Hl} + v t_{hH} + \theta - \sigma\mu v(t_{hH} + \theta + t_{HH} - t_{Hl}) \\
&= \sigma(1-\mu v)t_{HH} + v(1-\sigma\mu)t_{hH} + (1-\sigma)\mu[1+\sigma(1-\mu)]t_{Hl} + (1-\sigma\mu v)\theta > 0, \\
\frac{\partial f}{\partial \alpha_n} &= -\sigma\mu v < 0, \text{ and} \\
\frac{\partial f}{\partial \gamma_n} &= -\sigma\mu v(t_{hH} + \theta) < 0.
\end{aligned}
$$

Meanwhile, the execution effort incentive constraint (8) does not depend on α_n or γ_n.

2. Observe from (7) and (10) that α_t, β_t, and γ_t enter these expressions only via g. They are therefore interchangeable as far as the optimal full-information contract is concerned. It follows that one optimal contract is given by $\alpha_t = 0$ and $\beta_t = \beta^x$ (to satisfy (8)).

3. The effect of β_n on training incentives is given by

$$
\frac{\partial f}{\partial \beta_n} = \sigma\kappa_J\phi\Delta q - \sigma\mu v(t_{HH} + \theta).
$$

If $\partial f/\partial \beta_n < 0$, then setting β_n to the smallest value consistent with (8), $\beta_n = \beta^x$, is optimal. If $\partial f/\partial \beta_n > 0$ and $\beta_n > \beta^x$, consider a change in β_n and a simultaneous change in γ_t by

$$
\left.\frac{d\gamma_t}{d\beta_n}\right|_{f=const} = -\frac{\partial f/\partial \beta_n}{\partial f/\partial \gamma_t} = -\frac{\sigma\kappa_J\phi\Delta q - \sigma\mu v(t_{HH} + \theta)}{\sigma\mu v(t_{HH} + \theta)}
$$

that leaves f unchanged. Then the resulting change in the wage cost is

$$
\begin{aligned}
\frac{dw^L}{d\beta_n} &= \frac{\partial w^L}{\partial \beta_n} + \frac{\partial w^L}{\partial \gamma_t}\frac{d\gamma_t}{d\beta_n} \\
&= y^0 - \sigma\mu v(t_{HH} + \theta) + \sigma\mu v y_G - \sigma\mu v(t_{HH} + \theta)\frac{\sigma\kappa_J\phi\Delta q - \sigma\mu v(t_{HH} + \theta)}{\sigma\mu v(t_{HH} + \theta)} \\
&= y^0 + \sigma\mu v y_G - \sigma\kappa_J\phi\Delta q \\
&> \sigma t_{HH} - \sigma\kappa_J\phi\Delta q = \sigma t_{HL} > 0.
\end{aligned}
$$

This means that *decreasing* β_n and increasing γ_t to leave the training incentives f unchanged reduces the wage cost. It follows that irrespective of the sign of $\partial f/\partial\beta_n$, it is optimal to set $\beta_n = \beta^x$.

4. With $\alpha_n = \alpha_t = \gamma_n = 0$ and $\beta_n = \beta_t = \beta^x$, g reduces to $(t_{HH} + \theta)\gamma_t - y_L\beta^x$, and plugging this expression into f and solving $f = 0$ then leads to the expression for γ_t stated in the proposition. ■

Proof of Proposition 3: 1. We show first that it is weakly optimal to set α_t and β_t to their minimal feasible values, namely $\alpha_t = 0$ to satisfy limited liability, and $\beta_t = \beta^x$ to satisfy (8). To see this, observe that f,g, and w all have the same slope in γ_t and α_t; that is,

$$-\frac{\partial f/\partial\gamma_t}{\partial f/\partial\alpha_t} = -\frac{\partial g/\partial\gamma_t}{\partial g/\partial\alpha_t} = -\frac{\partial w^L/\partial\gamma_t}{\partial w^L/\partial\alpha_t} = -(t_{HH} + \theta).$$

By contrast, we have

$$-\frac{\partial h/\partial\gamma_t}{\partial h/\partial\alpha_t} = -(t_{LH} + \theta).$$

It follows that for any $\alpha_t > 0$, an increase in γ_t accompanied by a decrease in α_t that leaves f,g, and w^L unchanged, relaxes the constraint $h \geq 0$:

$$\frac{dh}{d\gamma_t} = \frac{\partial h}{\partial\gamma_t} + \frac{\partial h}{\partial\alpha_t}\left.\frac{d\alpha_t}{d\gamma_t}\right|_{f,g,w^L \text{ const.}} = -(t_{LH} + \theta) + (t_{HH} + \theta) > 0.$$

It is therefore weakly optimal to set $\alpha_t = 0$. Likewise,

$$-\frac{\partial f/\partial\gamma_t}{\partial f/\partial\beta_t} = -\frac{\partial g/\partial\gamma_t}{\partial g/\partial\beta_t} = \frac{\partial w^L/\partial\gamma_t}{\partial w^L/\partial\beta_t} = -\frac{t_{HH} + \theta}{t_{Hl} + \theta},$$

whereas

$$-\frac{\partial h/\partial\gamma_t}{\partial h/\partial\beta_t} = -\frac{t_{LH} + \theta}{t_{Hl} + \theta}.$$

Therefore,

$$\frac{dh}{d\gamma_t} = \frac{\partial h}{\partial\gamma_t} + \frac{\partial h}{\partial\beta_t}\left.\frac{d\beta_t}{d\gamma_t}\right|_{f,g,w^L \text{ const.}} = -(t_{LH} + \theta) + (t_{Hl} + \theta)\frac{t_{HH} + \theta}{t_{Hl} + \theta} = \kappa_S\phi\Delta q > 0,$$

which similarly implies that an increase in γ_t and an appropriate decrease in β_t would relax $h \geq 0$ while leaving all other constraints and the wage cost unchanged. It is therefore weakly optimal to set $\beta_t = \beta^x$.

2. Next, we show that it is optimal to set $\alpha_n = \gamma_n = 0$. This is not as obvious as in the silo case because both variables help to satisfy the truthtelling constraint (12), but can be shown using arguments similar to step 1 of this proof. Specifically, observe that g and h have the same slopes between β_n,α_n and between β_n,γ_n; that is,

$$\begin{aligned}
\left.\frac{d\alpha_n}{d\beta_n}\right|_{g,h \text{ const.}} &= -\frac{\partial g/\partial\beta_n}{\partial g/\partial\alpha_n} = -\frac{\partial h/\partial\beta_n}{\partial h/\partial\alpha_n} = -(t_{HH}+\theta) \text{ and} \\
\left.\frac{d\gamma_n}{d\beta_n}\right|_{g,h \text{ const.}} &= -\frac{\partial g/\partial\beta_n}{\partial g/\partial\gamma_n} = -\frac{\partial h/\partial\beta_n}{\partial h/\partial\gamma_n} = -\frac{t_{HH}+\theta}{t_{hH}+\theta}.
\end{aligned}$$

Then, for $\alpha_n > 0$, an increase in β_n accompanied by a decrease in α_n that leaves g and h unchanged relaxes the training incentive constraint (10) because α_n enters (10) only via g, whereas β_n also increases the first term in (10). The same holds (if $\gamma_n > 0$) for an increase in β_n accompanied by a decrease in γ_n that leaves g and h unchanged. These same changes in the contract also reduce the firm's wage cost, as we show next. For an increase in β_n and a decrease in α_n, we have

$$\begin{aligned}
\frac{dw^L}{d\beta_n} &= \frac{\partial w^L}{\partial\beta_n} + \left.\frac{\partial w^L}{\partial\alpha_n}\frac{d\alpha_n}{d\beta_n}\right|_{f,g \text{ const.}} \\
&= y^0 + \sigma\mu v[-(t_{HH}+\theta) + y_G] - (1-\sigma\mu v)(t_{HH}+\theta) \\
&= y^0 + \sigma\mu v y_G - (t_{HH}+\theta) \\
&= \sigma t_{HH} + (1-\sigma)\mu t_{Hl} + v t_{hH} + \sigma\mu v(t_{HH}-t_{hH}) - (t_{HH}+\theta) \\
&< \sigma t_{HH} + (1-\sigma)\mu t_{Hl} + v(t_{hH}+t_{HH}-t_{hH}) - t_{HH} \\
&= -(1-\sigma)t_{HH} + (1-\sigma)[\mu t_{Hl} + (1-\mu)t_{HH}] < 0.
\end{aligned}$$

For an increase in β_n and a decrease in γ_n, we have

$$\begin{aligned}
\frac{dw}{d\beta_n} &= \left.\frac{\partial w}{\partial \beta_n} + \frac{\partial w}{\partial \gamma_n}\frac{d\gamma_n}{d\beta_n}\right|_{f,g\text{ const.}} \\
&= y^0 + \sigma\mu v[-(t_{HH}+\theta) + y_G] - \left\{y^0 + \sigma\mu v[-(t_{hH}+\theta) - y_L]\right\}\frac{t_{HH}+\theta}{t_{hH}+\theta} \\
&= \frac{1}{t_{hH}+\theta}\left[y^0(t_{hH}+\theta) + \sigma\mu v(t_{hH}+\theta)y_G - y^0(t_{HH}+\theta) + \sigma\mu v(t_{HH}+\theta)y_L\right] \\
&< \frac{y_G}{t_{hH}+\theta}\left[-y^0 + \sigma\mu v(t_{hH}+\theta) + \sigma\mu v(t_{HH}+\theta)\right] \\
&\quad \text{(using } t_{HH} - t_{hH} = y_G \text{ and } y_L < y_G\text{)} \\
&= \frac{y_G}{t_{hH}+\theta}\{-[\sigma t_{HH} + (1-\sigma)\mu t_{Hl} + v t_{hH} + \theta] + \sigma\mu v t_{hH} + 2\sigma\mu v\theta + \sigma\mu v t_{HH}\} \\
&= \frac{y_G}{t_{hH}+\theta}[-\sigma(1-\mu v)t_{HH} - (1-2\sigma\mu v)\theta - v(1-\sigma\mu)t_{hH} - (1-\sigma)\mu t_{Hl}] < 0.
\end{aligned}$$

It follows that any contract with $\alpha_n > 0$ or $\gamma_n > 0$ can be improved upon, and therefore that $\alpha_n = \gamma_n = 0$ is optimal.

3. It remains to determine the optimal values of β_n and γ_t. Given the linearity of all constraints, they lie in one of the corners of the feasible set defined by $f \geq 0$, $g \geq 0$, $h \geq 0$, and the execution effort constraint (8), which is a subset of the feasible set in the full-information case of Proposition 2. The shape of this set was already partly determined in Proposition 2, see Figure 2. The truthtelling constraints (11) and (12) have slopes -1 and $-\frac{t_{HH}+\theta}{t_{LH}+\theta} < -1$, respectively, and intersect at $\beta_n = \frac{t_{Hl}+\theta}{t_{HH}+\theta}$ and $\gamma_t = 0$, as shown in Figure 3.

Then three possible cases arise, depending on how β^x relates to the intersection of f and h (at $\tilde{\beta}$) and to the intersection of f and g (at β^S). If $\beta^x \in (\tilde{\beta}, \beta^S]$, the full-information solution of Proposition 2 is optimal because both truthtelling constraints are satisfied. That is Case 2.

If $\beta \leq \beta^x$ (Case 1), then the logic of Proposition 2 still applies and it is optimal to choose the leftmost feasible point along the $f = 0$ line. However, now the feasible set is constrained by the no-kicking-upstairs constraint (12), and the optimal solution is at $\beta_n = \tilde{\beta}$. The optimal contract variables for this case are

$$\tilde{\beta} = \frac{(t_{LH}+\theta)\beta^S + \mu v\frac{\kappa_S}{\kappa_J}(t_{HL}+\theta)\beta^x}{t_{LH}+\theta+\mu v\frac{\kappa_S}{\kappa_J}(t_{HH}+\theta)} \text{ and } \tilde{\gamma} = \frac{(t_{HH}+\theta)\beta^S - (t_{HL}+\theta)\beta^x}{t_{LH}+\theta+\mu v\frac{\kappa_S}{\kappa_J}(t_{HH}+\theta)}. \tag{18}$$

Finally, if $\beta^x > \beta^S$ (Case 3), then $\beta_n = \beta^x$ is optimal, but γ_t still needs to be chosen to satisfy $g \geq 0$; that is, to prevent talent hoarding. With $g = 0$ and $\beta_n > \beta^S$, the training incentive constraint $f \geq 0$ is slack in this case. ■

Proof of Proposition 4: With simple output-based contracts, the truthtelling constraints (11) and (12) reduce to

$$\begin{aligned} g &= \gamma y_G - \beta y_L \geq 0 \text{ and} \qquad (19) \\ h &= \beta y_L + \gamma \kappa_S (q_h - \phi q_l) \geq 0, \qquad (20) \end{aligned}$$

and using (7), we obtain

$$w = \alpha + (\beta + \gamma) y^0 + \sigma \mu v (\gamma y_G - \beta y_L + \beta y_G - \gamma y_L) = \alpha + (\beta + \gamma)[y^0 + \sigma \mu v (y_G - y_L)]. \quad (21)$$

The constant payment α increases the wage cost but plays no role in any constraint. Therefore, $\alpha = 0$ is optimal. Next, (10) translates to

$$f = \sigma \kappa_J \phi \Delta q \beta + \sigma \mu v (\gamma y_G - \beta y_L) - \tau \geq 0. \quad (22)$$

Given the linearity of the problem, there are two candidates for an optimal solution, one in which γ is as small as possible, and one where it is as large as possible, with β chosen to satisfy (19), (20), and (22).

The first solution entails $\gamma = (y_L/y_G)\beta$ according to (19), in which case h in (20) simplifies to $\beta \kappa_S \phi \Delta q \frac{y_L}{y_G} > 0$. In (22), the second term vanishes and the condition reduces to (17), and solving $f = 0$ leads to $\beta = \beta^S$. Therefore (taking into account the execution incentive constraint as well), $\beta = \max\{\beta^S, \beta^x\}$ is optimal in this case. The resulting wage cost is

$$\left(1 + \frac{y_L}{y_G}\right) \beta [y^0 + \sigma \mu v (y_G - y_L)] > \beta y^0 = w^0. \quad (23)$$

The second solution (γ as large as possible) must satisfy (22) with equality. It is therefore preferred to the first solution if and only if it is preferred when we ignore all other constraints, because the other constraints only constrain the feasible solutions along the $f = 0$ line, but not how the iso-profit lines and $f = 0$ intersect.

Consider therefore $\beta = 0$, in which case solving (22) leads to

$$\gamma = \frac{\tau}{\sigma \mu v y_G} = \frac{\kappa_J \phi \Delta q}{\mu v y_G} \beta^S$$

and a wage cost of

$$\frac{\kappa_J \phi \Delta q}{\mu v y_G} \beta^S [y^0 + \sigma \mu v (y_G - y_L)]. \tag{24}$$

Equation (23) is smaller than (24), and thus the first solution is optimal, if and only if $\frac{y_G + y_L}{y_G} < \frac{\kappa_J \phi \Delta q}{\mu v y_G}$, which is the condition stated in the proposition.

Otherwise, it is optimal to choose γ as large as possible along $f = 0$, but the constraints (19), (20), and $\beta \geq \beta^x$ also need to be satisfied. First, since the first solution (with $\gamma = (y_L/y_G)\beta$) satisfies $g = 0$, for any solution to $f = 0$ with a larger γ, (19) will be slack. Second, if $q_h \geq \phi q_l$, then (20) is always satisfied. In this case, the optimal contract is $\beta = \beta^x$, and γ solves (22) for $\beta = \beta^x$. If $q_h < \phi q_l$, however, then (20) will be binding for small ξ and hence small β^x. Thus, as stated in the proposition, the optimal contract has β at its smallest feasible value, namely $\beta = \max\{\beta^x, \hat{\beta}\}$, where $\hat{\beta}$ is part of the solution to $f = 0$ and $h = 0$ in β, γ. The corresponding γ solves $f = 0$ for $\beta = \max\{\beta^x, \hat{\beta}\}$.

It remains to show that with either solution, the wage cost is strictly greater than with a transfer-contingent contract. It suffices to show that for each of the two cases, there exists a strictly-dominant transfer-contingent one, which in turn is still dominated by the contract of Proposition 3. Specifically, consider a transfer-contingent deviation from a simple contract that reduces γ_n by $d\gamma > 0$, and increases β by $\frac{t_{hH}+\theta}{t_{HH}+\theta} d\gamma$. Then both g and h according to (6) and (12) are unchanged, whereas (10) becomes slack because g is unchanged and β_n increases. Thus, this deviation satisfies all constraints. The wage cost according to (7) changes by

$$\begin{aligned}
dw &= \left(\frac{t_{hH} + \theta}{t_{HH} + \theta} - 1 \right) y^0 d\gamma + \sigma \mu v \left(\frac{t_{hH} + \theta}{t_{HH} + \theta} y_G - y_L \right) d\gamma \\
&= \frac{1}{t_{HH} + \theta} \left\{ -y_G y^0 + \sigma \mu v [(t_{hH} + \theta) y_G - (t_{HH} + \theta) y_L] \right\} d\gamma \\
&< \frac{y_G}{t_{HH} + \theta} \left[-y^0 + \sigma \mu v (t_{hH} + \theta) \right] d\gamma < 0,
\end{aligned}$$

where the last inequality follows from $y^0 = t^0 + \theta > v t_{hH} + \theta > v(t_{hH} + \theta)$. It follows that there exists a deviation from a simple contract that decreases γ_n, increases β_n, satisfies all constraints, and lowers the wage cost. ■

Proof of Proposition 5: Given Proposition 1, the firm's (per-division) profit with silos is $\pi^0 = (1 - \beta^0)(t^0 + \theta)$. For a lattice, the three cases listed in Proposition 3 all have in common that $\alpha_n = \gamma_n = \alpha_t = 0$ and $\beta_t = \beta^x$. Maximizing the lattice profit π^L with respect to the remaining two contract variables, β_n and γ_t (see step 3 of the proof of Proposition 3), is then equivalent to maximizing the difference $\pi^L - \pi^0$. Moreover, since the problem is linear, we can use the envelope theorem for constrained optimization (see e.g. Theorem 19.5 in Simon and Blume 1994) to examine the comparative statics with respect to the model's parameters. To determine the corresponding Lagrangian, consider first the difference in equilibrium profits. Using (5), the difference between the lattice output y^L and the silo output y^0 is $\sigma\mu v(y_G - y_L)$. The silo wage cost is $w^0 = \beta^0 y^0$, whereas the lattice wage cost according to (7) is $\beta_n y^0 + \sigma\mu v(g + \beta_n y_G)$. Given constraints (8) through (10), the Lagrangian for maximizing $\Delta\pi$ with respect to β_n and γ_t then is

$$\begin{aligned} \mathcal{L} &= \sigma\mu v(y_G - y_L) + (\beta^0 - \beta_n)(t^0 + \theta) - \sigma\mu v(g + \beta_n y_G) \\ &\quad + \lambda_f f + \lambda_g g + \lambda_h h + \lambda_\beta(\beta_n - \beta^x) \end{aligned} \tag{25}$$

along with the nonnegativity constraints $\lambda_f, \lambda_g, \lambda_h, \lambda_\beta \geq 0$ and the complementary slackness conditions $\lambda_f f = \lambda_g g = \lambda_h h = \lambda_\beta(\beta_n - \beta^x) = 0$.

The first-order conditions for maximization with respect to β_n and γ_t are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\beta_n} &= -(t^0 + \theta) + \sigma\mu v(t_{HH} + \theta) - \sigma\mu v y_G + \lambda_f[\sigma\kappa_J\phi\Delta q - \sigma\mu v(t_{HH} + \theta)] \\ &\quad -\lambda_g(t_{HH} + \theta) + \lambda_h(t_{HH} + \theta) + \lambda_\beta = 0 \text{ and} \end{aligned} \tag{26}$$

$$\frac{\partial \mathcal{L}}{\gamma_t} = -\sigma\mu v(t_{HH} + \theta) + \lambda_f\sigma\mu v(t_{HH} + \theta) + \lambda_g(t_{HH} + \theta) - \lambda_h(t_{LH} + \theta) = 0. \tag{27}$$

For case 1 of Proposition 3, we have $\beta^0 = \beta^S$ (for the silo case), $\lambda_f, \lambda_h > 0$, and $\lambda_g = \lambda_\beta = 0$. The values of λ_f and λ_h that solve (26) and (27) are

$$\lambda_f = \frac{t^0 + \theta + \sigma\mu v(y_G + \kappa_S\phi\Delta q\frac{t_{HH}+\theta}{t_{LH}+\theta})}{(\mu v\kappa_S\frac{t_{HH}+\theta}{t_{LH}+\theta} + \kappa_J)\sigma\phi\Delta q} \text{ and } \lambda_h = \mu v\frac{t^0 + \theta + \sigma\mu v y_G - \sigma\kappa_J\phi\Delta q}{(\mu v\kappa_S\frac{t_{HH}+\theta}{t_{LH}+\theta} + \kappa_J)\phi\Delta q} \ \frac{t_{HH} + \theta}{t_{LH} + \theta}. \tag{28}$$

For case 2, we have $\beta^0 = \beta^S$, $\lambda_f, \lambda_\beta > 0$ and $\lambda_g = \lambda_h = 0$. Specifically, (27) implies that $\lambda_f = 1$. For case 3, we have $\beta^0 = \beta^x$ and $\lambda_g, \lambda_\beta > 0$, and $\lambda_f = \lambda_h = 0$. Specifically, (27) implies that $\lambda_g = \sigma\mu v$.

Part (a): τ does not appear in (25) directly, but both β^S and f depend on τ. Thus, for cases 1 and 2, for which $\beta = \beta^S$,

$$\frac{d\mathcal{L}}{d\tau} = \frac{\partial\beta^S}{\partial\tau}(t^0 + \theta) - \lambda_f. \tag{29}$$

For case 1, substitute λ_f from (28):

$$\frac{d\mathcal{L}}{d\tau} = \frac{t^0+\theta}{\sigma\kappa_J\phi\Delta q} - \frac{t^0+\theta+\sigma\mu v\left(y_G+\kappa_S\phi\Delta q\frac{t_{HH}+\theta}{t_{LH}+\theta}\right)}{\left(\mu v\kappa_S\frac{t_{HH}+\theta}{t_{LH}+\theta}+\kappa_J\right)\sigma\phi\Delta q},$$

which has the same sign as

$$\begin{aligned}
& (t^0+\theta)\left(\mu v\kappa_S\frac{t_{HH}+\theta}{t_{LH}+\theta}+\kappa_J\right) - \kappa_j\left[t^0+\theta+\sigma\mu v\left(y_G+\kappa_S\phi\Delta q\frac{t_{HH}+\theta}{t_{LH}+\theta}\right)\right] \\
= \quad & \mu v\left[(t^0+\theta)\kappa_S\frac{t_{HH}+\theta}{t_{LH}+\theta} - \sigma y_G\kappa_J - \sigma\kappa_S\kappa_J\phi\Delta q\frac{t_{HH}+\theta}{t_{LH}+\theta}\right]. \qquad (30)
\end{aligned}$$

Since $t^0+\theta > t^0 > \sigma t_{HH} = \sigma(\kappa_S+\kappa_J)\phi q_h$ according to (4), (30) is greater than

$$\sigma\mu v\kappa_S\frac{t_{HH}+\theta}{t_{LH}+\theta}\left[(\kappa_S+\kappa_J)\phi q_h - \frac{t_{LH}+\theta}{t_{HH}+\theta}\kappa_J(\phi-1)q_h - \kappa_J\phi\Delta q\right],$$

which is positive. We therefore have $\frac{dL}{d\tau} > 0$.

For case 2, we have $\frac{\partial\beta^S}{\partial\tau}(t^0+\theta) = \frac{t^0+\theta}{\sigma\kappa_J\phi\Delta q} > 1$ and $\lambda_f = 1$, therefore $\frac{d\mathcal{L}}{d\tau} > 0$. For case 3, $\frac{dL}{d\tau} = 0$ because $\beta^0 = \beta^x$ and the training incentive constraint $f \geq 0$ is slack. Overall, we have $\frac{d\mathcal{L}}{d\tau} \geq 0$.

Part (b): The execution effort cost ξ scales the bonus β^x. Thus,

$$\frac{d\mathcal{L}}{d\xi} = \frac{\partial\mathcal{L}}{\partial\beta^x}\frac{\partial\beta^x}{\partial\xi} = \frac{1}{\theta}\frac{\partial\mathcal{L}}{\partial\beta^x},$$

where

$$\frac{\partial\mathcal{L}}{\partial\beta^x} = (-\sigma\mu v + \lambda_f\sigma\mu v + \lambda_g)\frac{\partial g}{\partial\beta^x} + \lambda_h\frac{\partial h}{\partial\beta^x} - \lambda_\beta,$$

$\frac{\partial g}{\partial\beta^x} = t_{Hl}+\theta$ and $\frac{\partial h}{\partial\beta^x} = -(t_{Hl}+\theta) = -\frac{\partial g}{\partial\beta^x}$. Therefore,

$$\frac{\partial\mathcal{L}}{\partial\beta^x} = (-\sigma\mu v + \lambda_f\sigma\mu v + \lambda_g - \lambda_h)(t_{Hl}+\theta) - \lambda_\beta. \qquad (31)$$

To evaluate the first term in (31), rewrite the first-order condition (27) as

$$(-\sigma\mu v + \lambda_f\sigma\mu v + \lambda_g - \lambda_h)(t_{HH}+\theta) + \lambda_h\kappa_S\phi\Delta q = 0,$$

which is equivalent to

$$-\sigma\mu v + \lambda_f\sigma\mu v + \lambda_g - \lambda_h = -\frac{\kappa_S\phi\Delta q}{t_{HH}+\theta}\lambda_h. \qquad (32)$$

Therefore, from (31),

$$\frac{\partial \mathcal{L}}{\partial \beta^x} = -\frac{\kappa_S \phi \Delta q}{t_{HH}+\theta}(t_{Hl}+\theta)\lambda_h - \lambda_\beta < 0$$

because $\lambda_h, \lambda_\beta \geq 0$ and either $\lambda_h > 0$ (in case 1) or $\lambda_\beta > 0$ (cases 2 and 3).

Part (c): κ_S appears in both y_G and in all t_{xy} expressions. Thus,

$$\begin{aligned}\frac{d\mathcal{L}}{d\kappa_S} &= \sigma\mu v(\phi-1)q_h + (\beta^0-\beta_n)\frac{\partial t^0}{\partial \kappa_S} - \sigma\mu v\left[\frac{\partial g}{\partial \kappa_S} + (\phi-1)q_h\beta_n\right] \\ &\quad + \lambda_f \sigma\mu v \frac{\partial g}{\partial \kappa_S} + \lambda_g \frac{\partial g}{\partial \kappa_S} + \lambda_h \frac{\partial h}{\partial \kappa_S}. \end{aligned} \tag{33}$$

From (27) we have

$$-\sigma\mu v + \lambda_f\sigma\mu v + \lambda_g = \frac{t_{LH}+\theta}{t_{HH}+\theta}\lambda_h.$$

Equation (33) therefore simplifies to

$$\frac{d\mathcal{L}}{d\kappa_S} = \sigma\mu v(1-\beta_n)(\phi-1)q_h + (\beta^0-\beta_n)\frac{\partial t^0}{\partial \kappa_S} + \lambda_h\frac{t_{LH}+\theta}{t_{HH}+\theta}\ \frac{\partial g}{\partial \kappa_S} + \lambda_h\frac{\partial h}{\partial \kappa_S}, \tag{34}$$

where

$$\frac{\partial g}{\partial \kappa_S} = (\beta^x+\gamma_t-\beta_n)\phi q_h \text{ and } \frac{\partial h}{\partial \kappa_S} = \phi(\beta_n q_h - \beta^x q_h - \gamma_t q_l) = -\frac{\partial g}{\partial \kappa_S} + \gamma_t\phi\Delta q.$$

The first two terms in (34) are strictly positive, and the last two equal

$$\lambda_h\left[\left(\frac{t_{LH}+\theta}{t_{HH}+\theta}-1\right)(\beta^x+\gamma_t-\beta_n)\phi q_h + \gamma_t\phi\Delta q\right]. \tag{35}$$

In cases 2 and 3, $\lambda_h = 0$, whereas in case 1, $h = 0$ is equivalent to

$$\beta_t(t_{Hl}+\theta) + \gamma_t(t_{LH}+\theta) = \beta_n(t_{HH}+\theta)$$

and therefore

$$\beta_t(t_{HH}+\theta) + \gamma_t(t_{HH}+\theta) > \beta_n(t_{HH}+\theta)$$

or $\beta^x + \gamma_t - \beta_n > 0$. It follows that the λ_h-terms in (34) are nonnegative, and that overall $\frac{\partial \mathcal{L}}{\partial \kappa_s} > 0$.

Part (d): κ_J appears in β^S, y_L and in all t_{xy} expressions. Thus,

$$\frac{d\mathcal{L}}{d\kappa_J} = -\sigma\mu v(\phi q_h - q_l) + (t^0+\theta)\frac{\partial \beta^0}{\partial \kappa_J} - \sigma\mu v\frac{\partial g}{\partial \kappa_J} + \lambda_f\sigma\mu v\frac{\partial g}{\partial \kappa_J} + \lambda_g\frac{\partial g}{\partial \kappa_J} + \lambda_h\frac{\partial h}{\partial \kappa_J}. \tag{36}$$

For case 3, we have $\beta^0 = \beta_n = \beta^x$, $\lambda_f = \lambda_h = 0$ and $\lambda_g = \sigma\mu v$ per (27). Thus, all terms in (36) except the first vanish, and we have $\frac{\partial \mathcal{L}}{\partial \kappa_J} < 0$. For case 2, we have $\beta^0 = \beta^S$ with $\frac{\partial \beta^S}{\partial \kappa_J} < 0$, $\lambda_g = \lambda_h = 0$, and $\lambda_f = 1$ per (27). Thus, all terms in (36) except the first two vanish, and we have again $\frac{\partial \mathcal{L}}{\partial \kappa_J} < 0$.

For case 1, the first two terms in (36) are negative just as in case 2. To evaluate the remaining terms, note that $\frac{\partial g}{\partial \kappa_J} = \beta^x q_l + \gamma_t \phi q_h - \beta_n \phi q_h$ and $\frac{\partial h}{\partial \kappa_J} = -\frac{\partial g}{\partial \kappa_J}$. With $\lambda_g = 0$, the $\frac{\partial g}{\partial \kappa_J}$ and $\frac{\partial g}{\partial \kappa_J}$ terms in (36) then simplify to

$$(-\sigma\mu v + \lambda_f \sigma\mu v - \lambda_h)(\beta^x q_l + \gamma_t \phi q_h - \beta_n \phi q_h). \tag{37}$$

Per (32), the first term in (37) equals $-\frac{\kappa_S \phi \Delta q}{t_{HH}+\theta}\lambda_h$. Moreover, given (18), it is straightforward that $\beta^x q_l + \gamma_t \phi q_h - \beta_n \phi q_h$ is linear in β^x because both β_n and γ_t are. To show that (36) is negative for case 1, it therefore suffices to show that (36) is negative for the minimal and maximal values of β^x, namely 0 and $\tilde{\beta}$. For $\beta^x = 0$, $h = 0$ is equivalent to $\beta_n(t_{HH} + \theta) = \gamma_t(t_{LH} + \theta)$, which implies $\gamma_t > \beta_n$ and therefore $\beta^x q_l + \gamma_t \phi q_h - \beta_n \phi q_h > 0$. This makes (37) negative and it follows overall that $\frac{d\mathcal{L}}{d\kappa_J} < 0$. For $\beta^x = \tilde{\beta} = \beta_n$, we have for the second term in (37):

$$\frac{\partial g}{\partial \kappa_J} = \beta^x q_l + \gamma_t \phi q_h - \beta_n \phi q_h = \gamma_t \phi q_h - \tilde{\beta}(\phi q_h - q_l). \tag{38}$$

Substituting $\beta_n = \beta^x = \tilde{\beta}$ into $h = 0$, we obtain $y_t = \frac{y_L}{t_{LH}+\theta}\tilde{\beta}$, and (38) equals

$$\tilde{\beta}\frac{y_L}{t_{LH}+\theta}\phi q_h - \tilde{\beta}(\phi q_h - q_l) = \tilde{\beta}\frac{(\phi q_h - q_l)}{t_{LH}+\theta}[\phi q_h \kappa_J - (t_{LH}+\theta)] = -\tilde{\beta}\frac{(\phi q_h - q_l)(\kappa_S \phi q_l + \theta)}{t_{LH}+\theta}.$$

Note that $\frac{\partial \beta^S}{\partial \kappa_J} = -\frac{1}{\kappa_J}\beta^S$, and plugging $\beta^x = \tilde{\beta}$ into (18) leads to $\tilde{\beta} = \beta^S/[1 + \mu v \frac{\kappa_S}{\kappa_J}\frac{y_L}{t_{LH}+\theta}]$. Collecting terms, we have

$$\begin{aligned}
\frac{d\mathcal{L}}{d\kappa_J} &= -\sigma\mu v(\phi q_h - q_l) - \beta^S\frac{t^0+\theta}{\kappa_J} + \beta^S\frac{\kappa_S\phi\Delta q}{t_{HH}+\theta}\,\lambda_h\,\frac{(\phi q_h - q_l)(\kappa_S\phi q_l + \theta)}{t_{LH}+\theta}\,\frac{1}{1+\mu v\frac{\kappa_S}{\kappa_J}\frac{y_L}{t_{LH}+\theta}} \\
&< -\beta^S\frac{t^0+\theta}{\kappa_J} + \beta^S\frac{\kappa_S\phi\Delta q}{t_{HH}+\theta}\,\lambda_h\,\frac{(\phi q_h - q_l)(\kappa_S\phi q_l + \theta)}{t_{LH}+\theta},
\end{aligned}$$

and a sufficient condition for $\frac{\partial \mathcal{L}}{\partial \kappa_J} < 0$ is therefore

$$\lambda_h < \frac{t^0+\theta}{\kappa_S\phi\Delta q}\,\frac{t_{HH}+\theta}{y_L}\,\frac{t_{LH}+\theta}{\kappa_S\phi q_l + \theta}. \tag{39}$$

Next, λ_h according to (28) has the structure $\lambda_h = \frac{a+b\mu v}{c+d\mu v}\mu v$ with positive coefficients a, b, c, d, and it is straightforward to show that λ_h is increasing in μv. Provided that condition (16) holds, an upper bound for μv is given by $\frac{\kappa_J \phi \Delta q}{y_G+y_L}$, and substituting this upper bound into (28) gives us

$$\begin{aligned}
\lambda_h &\le \frac{\kappa_J \phi \Delta q}{y_G + y_L} \; \frac{t^0 + \theta + \sigma y_G \frac{\kappa_J \phi \Delta q}{y_G+y_L} - \sigma \kappa_J \phi \Delta q}{\left(\frac{\kappa_J \phi \Delta q}{y_G+y_L} \kappa_S \frac{t_{HH}+\theta}{t_{LH}+\theta} + \kappa_J\right) \phi \Delta q} \; \frac{t_{HH} + \theta}{t_{LH} + \theta} \\
&= \frac{t^0 + \theta - \sigma \kappa_J \phi \Delta q (1 - \frac{y_G}{y_G+y_L})}{\kappa_S \phi \Delta q (t_{HH} + \theta) + (y_G + y_L)(t_{LH} + \theta)} \, (t_{HH} + \theta) \\
&< \frac{t^0 + \theta}{\kappa_S \phi \Delta q}.
\end{aligned}$$

This upper bound is clearly smaller than the sufficient upper bound in (39) because the first terms in each are the same, and both the second and the third terms in (39) exceed 1. This completes the proof that $\frac{d\mathcal{L}}{d\kappa_J} < 0$ for case 1, and thus for all three cases. ■

References

Acemoglu, Daron, and Jörn-Steffen Pischke (1998): "Why do firms train? Theory and evidence." *The Quarterly Journal of Economics* 113(1): 79-119.

Alchian, Armen A. and Harold Demsetz (1972):, "Production, Information Costs, and Economic Organization." *American Economic Review*, 62 (5): 777-795

Alok, Shashwat, and Radhakrishnan Gopalan (2018): "Managerial compensation in multi-division firms." *Management Science* 64 (6): 2856-2874.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek (2008): "When Does Coordination Require Centralization?" *American Economic Review*, 98 (1): 145-179

Athey, Susan and John Roberts (2001): "Organizational Design: Decision Rights and Incentive Contracts." *American Economic Review Papers and Proceedings* 91(2): 200-205

Baker, George P. (2002): "Distortion and Risk in Optimal Incentive Contracts." *Journal of Human Resources*, 37 (4): 728-751

Baker, George P., Michael Gibbs, and Bengt Holmström (1994): "The internal economics of the firm: Evidence from personnel data." *Quarterly Journal of Economics* 109 (4): 881-919.

Bar-Isaac, Heski, and Clare Leaver (2021): “Investment in human capital under endogenous asymmetric information.” CEPR Discussion paper 16305.

Becker, Gary S. (1964): *Human Capital.* Columbia University Press

Benko, Cathleen, and Molly Anderson (2010): *The Corporate Lattice: Achieving High Performance In the Changing World of Work.* Harvard Business Review Press.

Benson, Alan, Danielle Li, and Kelly Shue (2024), “Potential” and the Gender Promotion Gap.” Working paper.

Bianchi, Nicola, Giulia Bovini, Jin Li, Matteo Paradisi, and Michael L. Powell (2021): “Career spillovers in internal labor markets.” No. w28605. National Bureau of Economic Research

Bidwell, Matthew (2011): “Paying more to get less: The effects of external hiring versus internal mobility.” *Administrative Science Quarterly* 56 (3): 369-407.

Bidwell, Matthew, and James Keller (2014): “Within or without? How firms combine internal and external labor markets to fill jobs.” *Academy of Management Journal* 57 (4): 1035-1055.

Boon, Corine, Rory Eckardt, David P. Lepak, and Paul Boselie (2018): “Integrating strategic human capital and strategic human resource management.” *The International Journal of Human Resource Management* 29 (1): 34-67.

Bryan, Lowell L, Claudia I. Joyce and Leigh M. Weiss (2006): “Making a Market in Talent.” *McKinsey Quarterly*, May 2006

Bushman, Robert M., Raffi J. Indjejikian, and Abbie Smith (1995): “Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies.” *Journal of Accounting research* 33: 101-128.

Campbell, Dennis, Yuan Zou, and Shu Lin (20223): “Haidilao: changing your future with your own hands.” HBS case 9-122-031

Cappelli, Peter (2008): "Talent management for the twenty-first century." *Harvard Business Review* 86 (3) p74

Cappelli, Peter (2013): "HR for neophytes." *Harvard Business Review*, October

Carmichael, Lorne H. (1985): "Incentives in Academics: Why Is There Tenure?" *Journal of Political Economy*, 96 (3): 453-472

Chartered Institute of Personnel and Development (2020): "Talent Management Survey." retrieved on Janury 10th 2022 from `http://go.hibob.com/rs/466-KLY-124/images/bob\_2020\_Talent\%20ManagementSurvey.pdf`

Chase, Ivan (1991): "Vacancy Chains." *Annual Review of Sociology*, 17, 133-154

Collings, David G., and Kamel Mellahi (2009): "Strategic talent management: A review and research agenda." *Human Resource Management Review* 19.4: 304-313.

Conaty, Bill and Ram Charan (2010): *Talent Masters*, Crown Business

Cowgill, Bo, Jonathan M.V. Davis, B. Pablo Montagnes, Patryk Perkowski, and Bettina Hammer (2023): "How to Design an Internal Talent Marketplace." *Harvard Business Review* May-June

Cowgill, Bo, Jonathan MV Davis, B. Pablo Montagnes, and Patryk Perkowski (2024): "Stable Matching on the Job? Theory and Evidence on Internal Talent Markets." *Management Science*, June

Cullen, Zoë B., and Ricardo Perez-Truglia (2019): "The old boys' club: Schmoozing and the gender gap." No. w26530. National Bureau of Economic Research.

Cyert, Richard M., and James March (1963): *A behavioral theory of the firm.* University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.

Demougin, Dominique and Aloysius Siow (1994): "Careers in Ongoing Hierarchies." *The American Economic Review,* 84 (5): 1261-1277

Dessein, Wouter (2014): "Incomplete Contracts and Firm Boundaries: New Directions." *Journal of Law, Economics, and Organization* 30 (Supplement 1): i13-i36.

Dessein, Wouter, Luis Garicano and Robert Gertner (2010): "Organizing for Synergies." *American Economic Journal: Microeconomics*, 2 (4): 77-114

De Varo, Jed and Michael Waldman (2012): "The Signaling Role of Promotions: Further Theory and Empirical Evidence." *Journal of Labor Economics*, 30 (1), 91-147

Diaz, Brayan, Andrea Neyra-Nazarrett, Julian Ramirez, Raffaella Sadun, and Jorge Tamayo (2024): "Training within Firms." Working paper

Doeringer, Peter and Michael Piore (1971): *Internal Labor Markets and Manpower Analysis*, Heath, Lexington, Mass.

Drechsel-Grau, Moritz, and Felix Holub (2024): "Are Male Bosses Bad for Women's Careers? Evidence from a Multinational Corporation." Working paper, LMU Munich and Goether University Frankfurt

Drucker, Peter F. (1954): *The Practice of Management.* Harper & Row, New York

Fairburn, James A. and James M. Malcomson (1994): "Rewarding Performance by Promotion to a Different Job." *European Economic Review* 38 (3/4): 683-690

Feltham, Gerald A., and Jim Xie (1994): "Performance measure congruity and diversity in multi-task principal/agent relations." *Accounting review* 429-453.

Frederiksen, Anders, Lisa B. Kahn, and Fabian Lange (2020): "Supervisors and performance management systems." *Journal of Political Economy* 128 (6): 2123-2187.

Friebel, Guido, Matthias Heinz, and Nikolay Zubanov (2022): "Middle Managers, Personnel Turnover, and Performance: A Long-Term Field Experiment in a Retail Chain." *Management Science* 68(1): 211-229.

Friebel, Guido and Michael Raith (2004): "Abuse of Authority and Hierarchical Communication." *RAND Journal of Economics*, 35 (2): 224-244

Friebel, Guido and Michael Raith (2010): "Resource Allocation and Organizational Form." *American Economic Journal: Microeconomics*, 2 (2): 1-33

Fudenberg, Drew, and Luis Rayo (2019): "Training and effort dynamics in apprenticeship." *American Economic Review* 109 (11): 3780-3812.

Gallup (2015): "State of the American Manager: Analytics and Advice for Leaders." Retrieved on Jan. 12, 2022 from www.gallup.com/services/182216/state-american-manager-report.aspx

Garicano, Luis (2000): "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy* Vol 108 (5)

Garicano, Luis, and Luis Rayo (2017): "Relational knowledge transfers." *American Economic Review* 107 (9): 2695-2730.

Gibbons, Robert (2005): "Four formal (izable) theories of the firm?" *Journal of Economic Behavior & Organization* 58 (2): 200-245.

Gibbons, Robert (2013): "Cyert and March (1963) at fifty: A perspective from organizational economics." MIT and NBER April 7.

Gibbons, Robert and Michael Waldman (1999): "A Theory of Wage and Promotion Dynamics Inside Firms." *Quarterly Journal of Economics,* 114 (4): 1321-1358.

Green, Jeff, and Cristin McClave (2021): "How Midsize Firms Can Attract — and Retain — Talent Right Now." *Harvard Business Review*, December

Groysberg, Boris, and Sarah Abbott (2013): "A.P. Moller-Maersk Group: Evaluating Strategic Talent Management Inititives." HBS case 9-412-147

Haegele, Ingrid (2024): "Talent Hoarding in Organizations." Working paper. `https://arxiv.org/pdf/2206.15098`

Hall, Brian, and Jonathan Lim (2002): "Massachusetts Financial Services." HBS Case 9-902-132

He, Qiwei, annd Michael Waldman (2024): “Intra-firm employer learning, talent hoarding, and managerial practices.” Working paper, Cornell University

Hoffman, Mitchell, and Christopher T. Stanton (2024): “People, Practices, and Productivity: A Review of New Advances in Personnel Economics.” NBER Working Paper 32849

Hoffman, Mitchell, and Steven Tadelis (2021): “People management skills, employee attrition, and manager rewards: An empirical analysis.” *Journal of Political Economy* 129 (1): 243-285.

i4cp (2016): “Talent Mobility Matters.” Report by the Institute for Corporate Productivity, retrieved from `https://www.i4cp.com/file/1519/download` on January 17, 2022

Kahn, Lisa (2013): “Asymmetric Information between Employers.” *American Economic Journal: Applied Economics*, 5(4): 165-205.

Kahn, Lisa, and Fabian Lange (2014): “Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures.” *Review of Economic Studies* 81 (4): 1575-1613.

Ke, Rongzhu, Jin Li, and Michael Powell (2018): “Managing careers in organizations.” *Journal of Labor Economics* 36 (1): 197-252.

Keller, James R., and Kathryn Dlugos: “Advance’Em to Attract’Em: How Promotions Influence Applications in Internal Talent Markets.” *Academy of Management Journal* 66 (6): 1831-1859.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton (2015): “The value of bosses.” *Journal of Labor Economics* 33 (4): 823-861.

Levitt, Steven and Christopher Snyder (1997): “Is No News Bad News? Information Transmission and the Role of ”Early Warning” in the Principal-Agent Model.” *RAND Journal of Economics* 28 (4): 641-61

Linebaugh (2012): "The New GE Way: Go Deep, Not Wide." *Wall Street Journal*, March 7, 2012

Marino, Anthony M., and Jan Zabojnik (2004): "Internal competition for corporate resources and incentives in teams." *RAND Journal of Economics*, 710-727.

Martin, Jean, and Conrad Schmidt (2010): "How to keep your top talent." *Harvard Business Review* 88.5 (2010): 54-61.

Mellahi, Kamel, and David G. Collings (2010): "The barriers to effective global talent management: The example of corporate élites in MNEs." *Journal of World Business* 45 (2): 143–149

Mercer (2021): "Win With Empathy. Global Talent Trends 2020-2021." Retrieved on January 28, 2022 from `https://www.mercer.com/content/dam/mercer/attachments/private/global-talent-trends/2021/gl-2021-gtt-global-eng-mercer.pdf`

Meyer, Margaret A. (1994): "The dynamics of learning with team production: Implications for task assignment." *The Quarterly Journal of Economics* 109 (4): 1157-1184.

Michaels, Ed, Helen Handfield-Jones, and Beth Axelrod (2001): *The war for talent*. Harvard Business Press

Milgrom, Paul, and Sharon Oster (1987): "Job Discrimination, Market Forces, and the Invisibility Hypothesis." *The Quarterly Journal of Economics* 102 (3), pp. 453–476.

Milgrom, Paul, and John Roberts (1992): *Economics, Organization and Management*.

Minni, Virginia (2023): "Making the invisible hand visible: Managers and the allocation of workers to jobs." CEP Working paper 1948.

Morrison, Alan D., and William J. Wilhelm Jr. (2004): "Partnership firms, reputation, and human capital." *American Economic Review* 94 (5): 1682-1692.

Pearson, Andrall and Johanna Hurstak (1992): "Johnson & Johnson in the 1990s." HBS case 9-393-001

Perry, Elissa and Carol Kulik (2008): "The Devolution of HR to the line: Implications for Perceptions of People Management Effectiveness." *International Journal of Human Resource Management*, 19 (2): 262-273

Pinfield, Lawrence T. (1995): *The Operation of Internal Labor Markets: Staffing Practices and Vacancy Chains*, Plenum Press New York

Prendergast, Canice, and Robert H. Topel (1996): "Favoritism in Organizations." *Journal of Political Economy*, 104 (): 958-78.

Qian, Yingyi (1994): "Incentives and Loss of Control in an Optimal Hierarchy." *Review of Economic Studies*, 61: 527-544

Qing, Qu, Cao Shanshan, and Li Chaoping (2021): "The internal transfer program at Tencent: keeping the water fresh." Tsinghua University case TU0133

Rantakari, Heikki (2008): "Governing Adaptation." *The Review of Economic Studies*, 75 (4): 1257-85

Rantakari, Heikki (2011): "Organizational Design and Environmental Volatility." *Journal of Law, Economics and Organization*

Ready, DA and JA Conger (2007): "Make Your Company a Talent Factory." *Harvard Business Review* 85(6):68-77

Rosen, Evan (2010): "Smashing Silos." *Business Week* February 10, 2010

Rosen, Sherwin (1982): "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics,* 13 (2): 311-323

Rouen, Ethan, and Suraj Srinivasan (2019): "Joe & The Juice crosses the Atlantic." HBS case 9-118-039

Sandvik, Jason, Richard Saouma, Nathan Seegert, and Christopher T. Stanton (2021): "Treatment and Selection Effects of Formal Workplace Mentorship Programs." National Bureau of Economic Research No. w29148

Simon, Carl P., and Lawrence Blume (1994): *Mathematics for Economists.* W.W. Norton

Tadelis, Steven, and Oliver Williamson (2012): "Transaction-Cost Economics." In: Robert Gibbons and John Roberts (eds.), *Handbook of Organizational Economics*, Princeton University Press.

Waldman, Michael (1984): "Job Assignments, Signalling, and Efficiency." *RAND Journal of Economics* 15(2): 255-67

Waldman, Michael (2013): "Theory and Evidence in Internal Labor Markets." In: Robert Gibbons and John Roberts (eds.), *Handbook of Organizational Economics*, Princeton University Press, pp. 520-571.

Williamson, Oliver (1985): *The Economic Institutions of Capitalism*

Whittaker, Susan and Mick Marchington (2003): "Devolving HR Responsibility to the Line: Threat, Opportunity or Partnership?" *Employee Relations*, 25 (3): 245-261

Work Institute (2020): "2020 Retention Report." Retrieved from `https://info.workinstitute.com/hubfs/2020\%20Retention\%20Report/Work\%20Institutes\%202020\%20Retention\%20Report.pdf` on January 31, 2022

Zheng, Xiaoming, and Ziqian Zhao (2018): "Haidilao 2018: Demystifying Restaurant Employee Motivation." Tsinghua University (China Business Case Center) Case TU0102 (HBSP Product TU0102)