USC FBE APPLIED ECONOMICS WORKSHOP
presented by: Michael Raith
Friday, Oct. 18, 2013
1:30 pm - 3:00 pm; Room: **HOH-706**

Managers, Mentoring, and Internal Labor Markets

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

October 15, 2013

[1]Contact: gfriebel@wiwi.uni-frankfurt.de; michael.raith@simon.rochester.edu. We would like to thank Ricardo Alonso, Guillermo Caruana, Marco Celentani, Kong-Pin Chen, Mikhail Drugov, Bob Gibbons, Richard Holden, Mike Powell, Patrick Rey, Thierry Verdier, Julie Wulf, and seminar audiences at MIT, New South Wales, Carlos III, Queen's, Rochester, and the Spring 2012 NBER Organizational Economics Working Group meeting for their helpful comments and suggestions.

Abstract

We propose a theory of production and allocation of human capital in firms that focuses on the role of line managers, who both invest time to mentor junior employees, and acquire information about the juniors' abilities that is valuable for job assignments. We argue that this dual role of managers implies that internal labor markets and incentive contracts for managers must be determined jointly. Specifically, an internal labor market may be structured in two distinct ways: in a "silo," junior workers are eligible only for a transfer or promotion in the division they currently work in; in a "lattice," they can also be assigned to another division. The prospect of losing a good worker to another division undermines a managers's mentoring incentives, and may encourage her to misrepresent strategically information she provides about her workers. A lattice therefore leads to better job assignments but entails greater agency costs for the firm. Both silos or a lattice can be optimal. Our analysis suggests several reasons that may help explain a recent trend for firms to switch from silos to lattices with cross-divisional mobility.

Keywords: internal labor markets, organizational structure, intra-firm mobility, careers, mentoring, war for talent, career lattice

JEL-codes: D2, D8, L2, M5

1 Introduction

This paper makes two main arguments. The first is that line managers, or "bosses," play an important role for both the production and the allocation of human capital within firms. This dual role of managers introduces a strategic dimension into the study of internal labor markets that has received little attention in economics. Our second argument is that the role of managers helps explain the challenges that many large firms are recently experiencing in attempting to create greater cross-divisional mobility of their employees. Our paper is the first to theoretically examine internal labor markets in multidivisional firms, and thus expands the reach of economics beyond a large literature focused on traditional vertical job ladders.

Following Becker (1964), economists have distinguished between investments in human capital that are made by employees themselves, and those made by the firm. What has received little attention is that a significant share of training is not carried out by "the firm" but by managers. While formalized training is administered by Human Resources departments, many less tangible skills are imparted to employees through *mentoring* by bosses (Whittaker and Marchington 2003, Perry and Kulik 2008). Mentoring can take many forms, such as explaining technical matters, imparting leadership skills, providing advice on career decisions, or delegating a problem to an employee as an opportunity to learn, when in the short run it would be more efficient for the manager to handle the problem (Conaty and Charan, 2010, list many such examples). Our arguments apply without change to managers' investments in the recruiting of junior employees. These activities all have in common that they are costly to managers.

Both mentoring and its fruits are difficult to measure and therefore to incentivize, however. The main payoff from mentoring an employee, typically, is that a good employee enhances the productivity of the manager's unit, and thus indirectly contributes to the manager's performance—a truism to any manager. Central to this argument is the team nature of production in firms; see Alchian and Demsetz (1972).

Managers are in addition involved in the *allocation* of human capital because in the course of working with their employees, they acquire private information about their abilities. While firms collect information about employees that can be made available to anyone, bosses still have residual private information, especially about their employees' "soft" skills. Although learning about employees' abilities plays an important role in some economic theories (e.g., Gibbons and Waldman 1999), in reality it is not "the firm" that learns, but individual managers, whose input into personnel decisions is essential.

The dispersion of information potentially creates an additional agency problem because managers may be inclined to use their information strategically, as Peter Drucker observed 60 years ago: "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).

There are thus two agency problems between managers and firms: moral hazard with respect to mentoring, and adverse selection with respect to information provided about employees. Both agency problems are minimal when junior employees remain within their manager's unit. That is the case in traditional internal labor markets that are organized as vertical job ladders—the basis of all leading theories of the internal labor market since Doeringer and Piore (1971) and Williamson (1971).

Large multidivisional firms, however, are increasingly concerned with filling top positions with the best people from *anywhere* in the organization. Since the days of Doeringer and Piore, many firms have become bigger, more multi-divisional, and more global. At the same time, the requirements of managerial positions have become less industry-specific (Murphy and Zabojnik 2004, 2007). Moving employees through traditional job ladders is increasingly perceived as inefficient, and job ladders are increasingly referred to as "silos" in which employees get stuck.[1] Economists have documented highly convex returns to talent (Kaplan and Rauh, 2013), which can be attributed to "superstar" effects (Rosen 1981) that arise from the interaction of skill-biased technological change and scale economies. This evidence, too, points to increased returns to filling top positions in firms with top talent.

[1] "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).

Subsequent to an influential McKinsey publication from 1998 (Chambers et al.), "talent management" has become both a buzzword and a self-declared priority of almost any global firm, and is believed to be a key strategic issue for modern management (Cappelli 2008, Ready and Conger 2007, Collings and Mellahi 2009). It includes the internal development and monitoring of talented people, and their allocation to jobs in which they are most productive, including across divisional boundaries (Bryan et al. 2006, Conati and Charan 2010).[2] Following both management scholars and practitioners, we will refer to an internal labor market with cross-divisional mobility as a "lattice."[3]

We argue that because of the role of managers discussed above, establishing a lattice leads to higher agency costs for firms compared with silos, which helps explain both the historical longevity of silos and the difficulties that firms appear to encounter in transitioning to lattices. First, if a manager's main reward from mentoring is to have a productive employee working for her, then reassigning talented employees to another part of the firm directly undermines mentoring incentives.[4] Second, in line with Drucker's observations, "given that subsidiaries are generally rewarded (or punished) for their own performance, it may be in the subsidiary's self-serving interest to keep their best talent, rather than bring them to the attention of headquarters" (Mellahi and Collings, 2010, page 146; see also Capelli, 2008).

To understand the effects at work and possible solutions to the problem, we develop a model of a multidivisional firm and show that organizing its internal labor market as silos (job ladders) leads to inefficient job assignments, but minimizes agency problems at the managerial level. Organizing the internal labor market as lattice instead, and without

[2]Conaty and Charan's 2010 business bestseller "Talent Masters" is devoted to informal and formal ways in which firms can develop and identify talented employees, based on the experience of companies that have taken the lead.

[3]See e.g. Cleaver (2012). The term "corporate lattice" (which has the same meaning), is a registered trademark of Deloitte LLP, see e.g. Benko and Anderson (2010). We thank Ricardo Alonso for bringing the term "lattice" to our attention.

[4]Practitioners have recognized this problem too: "What reward does a line manager get for developing one of their people to the point where they are transferred to another department because they are demonstrating so much talent?" (Chick 2008).

undermining mentoring incentives, requires a redesign of managerial incentive contracts. We show that under quite general conditions, the managerial wage costs associated with a lattice exceed those of silos, creating a tradeoff in choosing between silos and a lattice whose resolution depends on the value of having the best people in top positions. We offer several reasons that may explain firms' recent talent management efforts.

We model a firm, headed by a CEO, that has two divisions, each of which consists of a manager and a worker.[5] Each division's output depends on the abilities of the manager and the worker, and the manager's execution (or production) effort. Prior to production, the manager invests mentoring effort to increase (stochastically) her worker's ability. The worker's ability is unverifiable, and is observable only to the division manager in our main case of interest. The division managers are risk-neutral and protected by limited liability. We consider different sets of feasible incentive contracts that are linear functions of both divisions' outputs but may depend, for instance, on whether or not a worker is transfered to the other division.[6]

For exogenous reasons, a managerial position may become vacant. The CEO would prefer to fill it with a qualified worker ideally from the same division, or possibly from the other division, or otherwise from outside. Promoting a good worker is efficient because a higher position in the firm's hierarchy is associated with a greater productivity of a manager's ability, consistent with other theories and with evidence (see our literature discussion below).[7]

With silos (Section 3.1), only the worker in the same division is eligible for promotion. Both the moral-hazard problem and the adverse-selection problem described above are minimal: optimal contracts are based on own divisional performance only; each manager reaps the maximal returns from her mentoring effort, knowing that her worker will stay

[5] We use the terms "division", "CEO" and "manager" purely for convenience; our analysis applies just as well to any adjacent tiers in in a larger hierarchy that generally satisfy the assumptions of our model.

[6] To exaggerate our emphasis on middle managers, the workers are reduced in our model to non-strategic pawns in a game between their bosses and top management.

[7] We focus on promotions to another division in order to juxtapose the lattice 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.

in her unit; and there is no reason to misrepresent her worker's ability to anyone.

A lattice offers the additional option of filling a managerial vacancy with a good worker from the other division, which we assume is preferable to hiring a new manager from outside (otherwise silos would always be optimal). With a lattice, the structure of the optimal contract and the associated wage cost for the firm depend on details about feasible contracts, information about workers, and the importance of execution effort.

A first benchmark result (Section 3.2) highlights the conditions under which a lattice may lead to *lower* agency costs for the firm compared with silos, contrary to our main message. When the CEO can observe the workers' abilities, and contracts can be contingent on the transfer of a worker to the other division, then the optimal incentive contract incentivizes execution effort based on the own division's performance, and incentivizes mentoring through a lump-sum referral bonus that is paid if the manager's (good) worker is promoted to the other division. If in addition mentoring effort is very costly relative to execution effort, then mentoring is less costly to induce with a lattice than with silos, because a referral bonus directly rewards the event of developing a good worker through mentoring, whereas with silos mentoring can be rewarded only indirectly through output-based incentive pay. A lattice unambiguously dominates silos in this case, because it leads to more efficient personnel assignments *and* lower managerial wage costs for the firm.

Our subsequent results, however, show that more costly execution effort, more restrictive contracting options, and especially private information about workers held by managers, *each* tend to make it more costly to incentivize both mentoring and execution effort with a lattice compared to silos. Whether a lattice is preferred to silos then depends on the value of having more efficient personnel assignments relative to the higher wage costs associated with a lattice.

Specifically, the costlier execution effort, the more the firm must reward own-division performance, which limits its ability to focus mentoring incentives on the event of the transfer of a worker. Because each manager stands to lose her good worker to the other division with some probability, incentives that with silos are sufficient to induce mentoring, are insufficient with a lattice, requiring the firm to pay an additional reward in the event of a transfer. Likewise, if transfers cannot be rewarded directly, for instance with message-contingent contracts (Section 3.4) or wages that depend only on the divisions' outputs (Section 3.5), then mentoring incentives are always weaker with a lattice and lead to higher wage costs for the firm.

Most importantly, when managers are privately informed about their workers, then under weak conditions, inducing mentoring is always more costly with a lattice than with silos, even if bonuses can depend on worker transfers, and even if execution effort is irrelevant (Section 3.3). Simple referral bonuses are no longer effective in this case because they would provide an incentive for managers to "kick upstairs" bad workers—a common concern with employee referral bonuses in many companies. Ensuring that a manager is willing to relinquish a good worker to the other division without being tempted to get rid of a bad worker as well, requires a combination of bonuses based on the own and the other division's performance that overall raises the wage cost to the firm above the level of silos.

We show how different parameters of the model influence the choice between silos and a lattice (Section 4). In particular, the greater the value of divisional output, the more important it is to get the best people to the top of each division, and for a sufficiently high value, a lattice dominates silos. One can interpret the value of output as the value of reducing costs or raising quality, which in turn has been linked to the degree of competition in product markets in both theoretical and empirical work. Further, the more difficult it is to hire managers from outside, the more firms will prefer to fill vacancies from inside, and thus the more likely a lattice is optimal. Both results help to explain why firms' efforts facilitate cross-divisional mobility are a relatively recent phenomenon, and coincide with firms' engagement in a "War for Talent" (Chambers et al. 1998).

Overall, then, our results show that contrary to the popular impression that silos are a symptom of organizational dysfunction, they provide the best incentives for managers to recruit and mentor people. Breaking up silos to establish cross-divisional mobility is not simply a matter of asking managers who their good people are; it requires changing the firm's incentive system (or other organizational practices) in order to induce managers to engage in mentoring and reveal information about talented workers. The associated higher agency costs for the firm are worth incurring only if the value of finding the best employee for each high-level position is large enough. The recent trend towards greater mobility of employees, in turn, may explain the apparent decline in mentoring by bosses observed by Capelli (2011).

Our paper relates to several strands of literature. First, two key features of our theory, asymmetric information about workers and the possible "poaching" of good workers, are reminiscent of Waldman (1984). Waldman argues that firms know more about their workers than do other firms, and shows how firms will optimally design promotions and wages to minimize poaching of good workers by other firms. In our model, the information asymmetry is *within* the firm, and the organizational problem is, in a sense, the opposite of Waldman's, namely how to design internal labor markets to *facilitate* rather than prevent the mobility of good workers. More generally, many theories of internal labor markets emphasize learning about workers (for instance, Gibbons and Waldman, 1999b) but assume that "the firm" that does the learning. Our theory, in contrast, posits that bosses usually know more about their workers than do others in the organization.[8]

Second, our theory focuses on managers as agents whose interests may diverge from the firm's. Fairburn and Macolmson (1994) and Prendergast and Topel (1996) have looked at biased evaluations by managers, and Carmichael (1985) and Friebel and Raith (2004) study bosses' potential incentives to hire unthreatening but inferior subordinates. The present paper highlights what we believe to be a more fundamental incentive problem that concerns both the production of human capital through mentoring and the allocation of human capital, and leads to the joint determination of both incentive contracts and the organization of internal labor markets.

Third, our assumption that managers are more important than workers builds on theories such as Rosen (1982), Qian (1994) or Garicano (2000). The importance of "bosses" is also emphasized in recent empirical work by Lazear et al. (2011), who estimate that the average boss is 1.75 times as productive as the average worker.

Fourth, we focus on the demand side of internal labor markets, that is, decisions on how to fill vacancies. Slot constraints and "job vacancy chains" (created when filling

[8] Kim (2011), too, assumes dispersed information about workers. His model studies optimal contracting to achieve efficient evaluation of workers *by peers*.

one vacancy chain creates new vacancy elsewhere) have been studied in the industrial-relations literature; see Chase (1991) and Pinfield (1995). In contrast, most economic theories focus on the supply side, that is, decisions on how to allocate given employees to different possible positions. An exception is Demougin and Siow (1994), which also focuses on the demand side but is otherwise very different from our paper.

Beyond the literature on internal labor markets, our paper overlaps with two more strands of organizational economics. One is 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 (2009, 2011), Dessein et al. (2010), Friebel and Raith (2010), and Dessein (2012).

Finally, our argument that the choice between silos and lattice is partly driven by product and labor market conditions links our paper to a diverse literature that studies the interaction between the internal organization of firms and their market environment. Theoretical papers include, for instance, Grossman and Helpman (2002), Raith (2003), Alonso et al. (2012), Gibbons et al. (2012), and Legros and Newman (2013). A rapidly growing empirical literature includes the work of Cuñat and Guadalupe (2005, 2009), Bloom and Van Reenen (2007), Rajan and Wulf (2006), Guadalupe and Wulf (2010), and Giroud and Müller (2010, 2011).

2 Model

Our model consists of a firm with two divisions headed by a CEO (male); see the organizational chart. Division $i = A, B$ consists of a manager M_i (female) and a worker W_i (male). The managers are the main strategic players in our model; the workers are not players in a game-theoretic sense. The job titles (CEO, manager) are chosen only for convenience; the model can more generally be interpreted as representing any two adjacent tiers of a multi-tier organization. Wage contracts are designed by the firm. The CEO acts in the firm's interest in filling vacant manager positions.

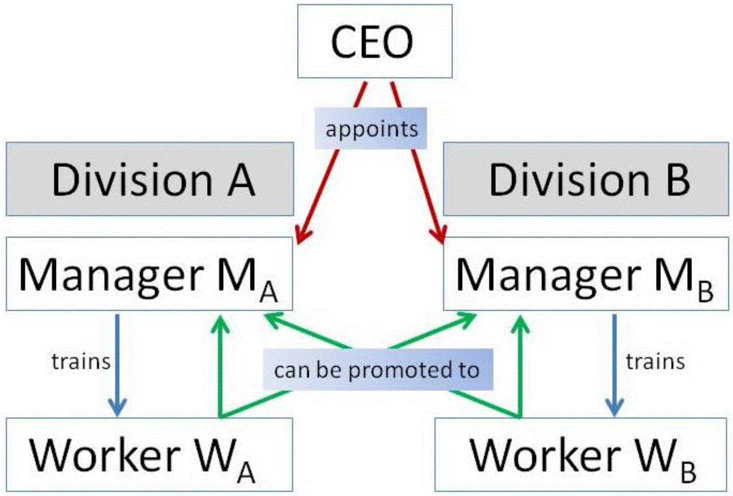

Figure 1: Organizational chart of the firm

2.1 Team production

Output in each division i depends on the productivity of the manager (q_i^m) and the worker (q_i^w), and manager M_i's execution effort x_i. Both the manager and the worker can be either "good" or "bad": $q_i^m, q_i^w \in \{q_g, q_b\}$, where $1 \geq q_g > q_b > 0$. Let $\Delta q = q_g - q_b$. These productivities will be affected by the managers' mentoring efforts and the CEO's assignment decisions. The "team productivity" of division i is defined as

$$t_i = \kappa q_i^m + (1 - \kappa) q_i^w. \qquad (1)$$

We assume that $\kappa \geq 1/2$ to reflect the greater importance of the manager for division output. Division i's output $y_i \in [0, 1]$ is a random variable belonging to a family of distributions with c.d.f. $F(y, t)$ such that $E[y|t] = t$. The shift parameter t is the division's productivity t_i, times M_i's execution effort $x_i \in \{0, 1\}$; thus $E[y_i|t_i] = x_i t_i$. An example of such a family of distributions is the Beta distribution with parameters α and β, with $\beta = 1$ and $\alpha = t/(1 - t)$.

The specification (1) has two implications. First, division i's expected output is increasing in the productivity of both M and W, but it is impossible to infer the productivity of M_i or W_i from realized output. It is therefore in each manager's interest to have a productive worker. Second, $\kappa \geq 1/2$ means that it may be in the firm's interest to move a good worker into vacant manager position where he can be more productive.

2.2 Timing

1. The firm hires M_A and M_B, offering each manager a contract whose resulting expected utility weakly exceeds her reservation utility. Each manager is good with probability p_m and bad with probability $1 - p_m$; denote the expected productivity of a manager by $q_m = p_m q_g + (1 - p_m) q_b$. Managers do not know their type before being hired, but learn their type upon joining the firm; a manager's type is therefore best interpreted as reflecting the quality of the *match* between the firm and the manager.

2. Each manager M_i $(i = A, B)$ hires a worker W_i from a pool of *ex ante* identical agents and invests effort $e_i \in \{0, 1\}$ in mentoring the worker, at cost $\psi_m e_i$. As a result, the worker is either good or bad; the probability of ending up with a good worker is a function of effort: $p_i(1) = p_h$ and $p_i(0) = p_l < p_h$. Let $\Delta p = p_h - p_l$.

3. M_i learns the type of W_i, which is private information.

4. With probability $1 - \sigma$, each manager leaves for exogenous reasons (thus σ is the probability of staying). The CEO then fills the vacant position with a manager hired from outside or by promoting a worker internally; see the next subsection.

5. Each manager M_i (which is either the original manager, or if she left, a newly appointed one) invests execution effort $x_i \in \{0, 1\}$ at cost $\psi_x x_i$.

6. Division i's output $y_i \in [0, 1]$ is realized, which depends on the division's team productivity and M_i's execution effort as described above.

Although mentoring and execution effort occur at different stages of the game, there are no periods and no discounting; our model is technically a static multi-stage game.[9]

[9] We separate the timing of mentoring and execution based on the idea that the benefits of mentoring tend to accrue with greater delay than those of many other managerial actions.

2.3 Vacancies and communication

We compare two types of internal labor markets that differ in the options available for filling a vacancy left by a departing manager:

1. With "silos," either the worker from the division with the vacancy is promoted, or a manager is hired from outside. A promoted good worker becomes a good manager with probability $\phi \leq 1$; a bad worker for sure becomes a bad manager[10]. A manager hired from outside is good with probability p_o; denote the expected productivity of such a manager by $q_o = p_o q_g + (1 - p_o) q_b$. Since the initial manager, who is good with probability p_m, may have been hired from outside or promoted from within the company, it is natural to assume $p_m \geq p_o$. A promoted worker is replaced with a new worker from outside, who is good with probability p_w.[11] Denote the expected productivity of such a worker by $q_w = p_w q_g + (1 - p_w) q_b$.

2. With a "lattice," there is a third option, to promote the worker from the *other* division. In this case, a promoted good worker is a good manager with probability $\delta\phi$. The additional "cross-divisional" discount factor $\delta \leq 1$ allows for human capital to be partly division-specific. This constitutes an intermediate case between firm-specific and task-specific human capital; see Gibbons and Waldman (2004).

Below, we state parameter constraints that ensure that the CEO will want to fill a vacancy (1) ideally with a good worker from the own division, (2) alternatively, with a good worker from the other division, and otherwise (3) by hiring from outside.

The decision how to fill a vacancy also depends on what the the CEO *knows* about the workers' abilities. We assume that a departing manager truthfully reveals her own worker's type because there is no reason to misrepresent it.[12] In contrast, a manager who

[10] If $\phi < 1$, a worker might be promoted "beyond his level of competence" — as claimed by the "Peter Principle," see Peter and Hull (1969), Fairburn and Malcomson (2001), and Lazear (2004).

[11] We can leave open whether a replacement worker is hired from outside or is transferred or promoted internally from an unmodeled lower tier. Cf. our remarks above about more general interpretations of our two-tier model.

[12] In our model, there is nothing at stake for the departing manager, so truthtelling is weakly optimal. In reality, managers departing by choice likely care about their legacy and therefore have a positive

stays with the firm will report her worker's type truthfully only if it is in her best interest, which it may not be if she has an incentive to "hoard" a good worker or "kick upstairs" a bad one. Consistent with casual observation of business practice, we assume that workers are unable to ascertain (or prove) their own suitability for a managerial position, see also Friebel and Raith (2004).

Note that worker turnover in our model is generated only by manager departures. A bad worker is not fired even after his manager or the CEO has learned his type. This assumption can be motivated by turnover costs: any *vacancy* at the manager or worker level must be filled either by promoting a worker or by hiring a manager or worker from outside, but *replacing* people because they are below average is prohibitively costly.[13]

2.4 Payoffs and contracts

Managers are risk-neutral and are protected by limited liability; specifically, assume that M_i's compensation w_i must be non-negative. If manager M_i stays with the firm, her utility is given by her wage, minus the cost of mentoring and execution effort:

$$U_i = w_i - \psi_m e_i - \psi_x x_i.$$

A manager accepts to work for the firm if her expected equilibrium utility is at least $\underline{U}$. The firm's profit from division i is given by

$$\pi_i = R y_i - w_i$$

for some $R > 0$.

We assume that division outputs y_1 and y_2 are verifiable, and consider wage contracts that are linear in both. The linearity of the wage function implies that in computing expected profits and wages, only the expected value of y_i matters; we therefore do not need any assumptions about the distribution of y_i other than $E(y_i|t_i) = x_i t_i$.

incentive to be honest about suitable successors.

[13] Allowing for the ability to replace a bad worker would have opposite effects on two incentive constraints. It would reduce a manager's incentive to invest in mentoring, thus tighening an effort incentive constraint. On the other hand, it would also reduce a manager's incentives to misrepresent a bad worker as a good one, i.e., would relax a truthtelling incentive constraint.

Aside from the restriction to linear contracts, our assumptions about feasible contracts are quite general. Specifically, we study three types of contracts that differ in what other variables or events are assumed to be contractible:

1. Simple, linear output-based contracts that cannot be conditioned on anything else: $w_i = \alpha + \beta y_i + \gamma y_j$ for $i \in \{A, B\}, j \neq i$. In this case, neither the workers' types nor messages about them are verifiable; that is, the managers' reports about their workers are cheap talk. The CEO makes promotion decisions that are ex-post optimal for the firm, based on the information available to him.[14]

2. Output-based contracts that can in addition be based on a manager's message about her worker's type: $w_i = \alpha_{\widehat{\theta}} + \beta_{\widehat{\theta}} y_i + \gamma_{\widehat{\theta}} y_j$ for $i \in \{A, B\}, j \neq i$; and $\widehat{\theta} \in \{g, b\}$ corresponding to whether M_i reports having a good or bad worker. We do not have a literal interpretation of such contacts in mind. Instead, the idea is that managers can be held accountable not only for outcomes but also for what they say about their workers, through the effects on their reputations (see also Friebel and Raith, 2010). In keeping with this interpretation, we consider only wage contracts based on the report made by the respective manager, not contracts based on joint reports.

3. Output-based contracts that can in addition be based on the event of a transfer of the own worker to the other division:

$$w_i = \begin{cases} \alpha_n + \beta_n y_i + \gamma_n y_{j \neq i} & \text{if } W_i \text{ stays in division } i \text{ (is not transferred)} \\ \alpha_t + \beta_t y_i + \gamma_t y_{j \neq i} & \text{if } W_i \text{ is transferred (promoted) to division } j;\ j \neq i \end{cases}$$

   with symmetric coefficients $(\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t)$ for each manager. The motivation for this case is that even if neither types nor messages are verifable, job assignments based on them may well be verifiable.

Because $y_i \in [0, 1]$, limited liability for every realization of y_1, y_2 requires the salary α

[14] If Ry represents revenue or gross profit, it might seem more natural for incentive contracts to be based on Ry than on just y. However, for our results this makes no difference. We chose to condition contracts on y because that way, changes in R (which play an important part in Section 4) have no effect on the contract, whereas in the alternative version the bonus parameters would be decreasing in R.

for each case to be nonnegative. To simplify the analysis, we assume that $\underline{U}$ is low enough such that the limited-liability constraint is binding and the participation constraint is not.

Any manager who leaves for exogenous reasons receives her reservation utility $\underline{U}$. The firm must pay the manager who replaces her. We assume that a replacement manager works under the same incentive contract as an incoming manager at stage 1 of the game, and will fill in details as we proceed in the analysis. In general, however, our results do not depend on the details of how replacement managers are paid.[15]

With a lattice and potentially even silos, the two divisions are linked because division A's output depends on M_B's effort, and M_A's compensation depends on y_B, and vice versa. Nevertheless, 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, as the revenue from output in division A (Ry_A), minus the total compensation paid to M_A *or her replacement*. Thus, π_A includes payments γy_B to M_A (the original manager or a replacement) based on division B's output, but does not include γy_A paid to M_B based on A's output.

2.5 Parameter conditions

Two parameter conditions ensure that replacing a manager from within the own division, the other division, or from outside, are all relevant options. The most basic condition for the existence of an internal labor market with upward mobility is that the firm prefers to *promote a worker within the same division* who is known to be good, than to hire a manager from outside:

$$\kappa[\phi q_g + (1-\phi)q_b] + (1-\kappa)q_w > \kappa q_o + (1-\kappa)q_g. \tag{2}$$

The left-hand side of (2) is the expected team productivity that results from promoting a good worker to manager, and hiring a replacement worker. The right-hand side is the

[15]Strictly speaking, since our model is static, replacement managers need not be given incentives for anything. In an ongoing organization, though, new managers need to be given incentives to invest in mentoring new workers just like the departing managers, which is why it is reasonable to assume that they are given the same wage contracts.

productivity of a team where a good worker remains in place and a new manager is hired from outside. Condition (2) is equivalent to

$$\kappa(\phi - p_o) > (1-\kappa)(1-p_w). \tag{3}$$

For *promoting a good worker across divisions* to be relevant requires a stricter version of (2), with ϕ replaced with $\delta\phi$. An equivalent stricter version of (3) is

$$\text{(C1)} \quad \kappa(\delta\phi - p_o) > (1-\kappa)(1-p_w).$$

Because of $\delta \leq 1$, promoting a good worker in the same division to fill a vacancy weakly dominates promoting a good worker from the other division.

Hiring from outside, in turn, requires at the least that an outside hire is preferred to promoting a worker known to be bad:

$$\kappa q_o + (1-\kappa)q_b > \kappa q_b + (1-\kappa)q_w,$$

which is equivalent to

$$\kappa p_o > (1-\kappa)p_w. \tag{4}$$

We will use a stricter version, however, namely

$$\text{(C2)} \quad p_o \geq \delta\phi p_w.$$ [16]

Condition (C2) implies that the firm would rather hire a manager from outside, than promote a worker from another division whom it knows nothing about. This assumption reflects the very idea that an internal labor market is an assignment mechanism through which the company learns about people's type. Unless the firm has good news about a worker, it prefers hiring from the outside pool of managers. This assumption is in line with the empirical evidence of Baker et al. (1994), who found that their focus firm hired from the outside at all levels. Below, we assume throughout that (C1) and (C2) hold.

Let $v = (1-\sigma)(1-p_h)$; it is the probability, say from M_A's perspective, that a vacancy arises for the position of M_B that will not filled by W_B (who is bad with probability $1-p_h$ if M_B chooses high mentoring effort). A weak restriction that we will use on occasion is that v is sufficiently small:

$$\text{(C3)} \quad v \leq 1-\kappa,$$

while Propositions 4 and 6 make use of a stronger version:

$$\text{(C3')} \quad \frac{q_b}{q_g} \geq \frac{\kappa v}{(1-\kappa)(1-v)},$$

which implies (C3) because $q_b/q_g \leq 1$. Both (C3) and (C3') ensure that the firm cares sufficiently about the quality of workers in their position as workers, and *not mainly* about how valuable a worker would be if promoted to manager in a different division.

Finally, we will assume throughout that with both silos and a lattice, it is optimal for the firm to induce both mentoring and execution effort; that is, $e_i = x_i = 1$. We refrain from stating the relevant parameter conditions formally; they amount to upper bounds on the effort costs ψ_m and ψ_x. The analysis then consists of determining the cost-minimizing contract for each form of internal labor market, and to comparing the costs and benefits of inducing high effort between silos and a lattice. Little is lost by making this assumption. Without it, for high enough mentoring effort cost (or low marginal benefit of effort), for instance, the firm would sometimes prefer a lattice with low mentoring effort to silos with high effort. The tradeoffs at work, however, are the same as in our analysis below.

3 Analysis

We begin our analysis with the silo case, and proceed to study optimal incentives with a lattice under different contracting assumptions. The three contract spaces defined in Section 2 are listed in increasing order of the fineness of conditioning events, which will be reflected in the complexity of the optimal contracts that we derive below. It will be convenient to cover the contract spaces in reverse order, beginning with transfer-contingent contracts.

3.1 Silos

To determine optimal contracts for silos, we begin by constructing the firm's and the managers' payoffs. In accordance with (1), denote the productivity of a team consisting of a manager of productivity q_x and worker q_y by $t_{xy} = \kappa q_x + (1-\kappa)q_y$, where $x, y \in \{g, b\}$. More generally, since q_m, q_o, and q_w represent the expected productivities of a manager hired at the beginning, a manager hired from outside, and a replacement worker, respectively, we can extend the above definition of t_{xy} to apply to $x \in \{m, g, b, o\}$ and $y \in \{g, b, w\}$.

The expected productivity of the original manager-worker team, as of stage 1 of the game, depends on the manager's mentoring effort e, which results in having either a good or a bad worker. Denote this productivity by $t_0(p(e))$, where $p(1) = p_h$ and $p(0) = p_l$. It is given by

$$t_0(p) = pt_{mg} + (1-p)t_{mb} = \kappa q_m + (1-\kappa)[pq_g + (1-p)q_b]. \tag{5}$$

Suppose, instead, the manager leaves and must be replaced (hence the subscript 'r' in the equation below). If the manager has a good worker (with probability p), the worker is promoted to manager but becomes a good manager only with probability ϕ. A new worker is hired from outside. If the worker is bad (probability $1-p$), then the manager's position is filled from outside while the worker stays in his position. The resulting expected output of the division is given by

$$t_r(p) = p[\phi t_{gw} + (1-\phi)t_{bw}] + (1-p)t_{ob}. \tag{6}$$

With silos, by definition, workers remain in the same division. Contracts conditioned on transfers are irrelevant; so let us consider simple output-based contracts (α, β, γ) first, and comment on message-contingent contracts later. It is also clear that without interaction between the divisions, there is no benefit from paying manager M_i based on division j's output; therefore $\gamma = 0$.

We focus on equilibria that induce both managers to invest in mentoring, such that $p_A = p_B = p_h$. The firm's net profit from division A is then simply

$$\pi_A^S = [\sigma t_0(p_h) + (1-\sigma)t_r(p_h)](R-\beta) \equiv t_S(R-\beta), \tag{7}$$

which is the product of the expected output $t_S = \sigma t_0(p_h) + (1-\sigma)t_r(p_h)$ of a siloed division (based on the probabilities of M_A staying or leaving), and the firm's per-unit profit net of the manager's wage, $R - \beta$.

M_A's expected payoff upon hiring, but before investing mentoring or execution effort, is a weighted average of the payoffs obtained if M_A herself stays or leaves, minus the costs of mentoring and execution effort. The two effort costs enter M_A's payoff differently because the mentoring effort e is invested before, but the execution effort x after, M_A knows whether she will stay with the firm. In addition, each manager knows her own type $\mu \in \{g, b\}$ upon joining the firm. The expected division output as of stage 2 of the game is therefore given by $t_\mu(p) = pt_{\mu g} + (1-p)t_{\mu b}$. Her net payoff then is

$$V_A^S(\mu, e, x) = \sigma\{t_\mu(p(e))\beta - \psi_x x\} + (1-\sigma)\underline{U} - \psi_m e. \tag{8}$$

Managers exert execution effort in stage 5 of the game. At that point, all uncertainty about manager departures, division team composition, and agent types has been resolved: M_i knows whether she and her worker are good or bad. We assume that regardless of team productivity, it is optimal for the firm to induce managerial execution effort. From $E[y_i|t_i] = x_i t_i$ we then obtain the incentive constraint

$$\alpha + \beta t_i - \psi_x \geq \alpha \Longleftrightarrow \beta t_i \geq \psi_x.$$

This constraint is most restrictive for the worst possible team with $q_i^m = q_i^w = q_b$ and thus $t_i = \kappa q_b + (1-\kappa)q_b = q_b$. For this team, the manager's execution incentive constraint simplifies to $\beta q_b \geq \psi_x$, or

$$\beta \geq \frac{\psi_x}{q_b} =: \beta^x. \tag{9}$$

In contrast, M_A invests mentoring effort under uncertainty about both her own and the other manager's tenure in the firm, and hence chooses e to maximize $V_A^S(e, x)$. The incentive constraint $V_A^S(1,1) \geq V_A^S(0,1)$, which due to the additive production function does not depend on the manager's own type, leads to the minimal bonus β that the firm must pay to induce high mentoring effort:

Proposition 1 *With silos, the minimal own-division bonus β that induces $e_i = 1$ is given by*

$$\beta^S = \frac{\psi_m}{\sigma(1-\kappa)\Delta p \Delta q}. \tag{10}$$

If $\beta^S \geq \beta^x$, then the firm's optimal silo contract is given by $(\alpha, \beta, \gamma) = (0, \beta^S, 0)$.

For all proofs, see Appendix B. The expression for β^S is intuitive. As in any model with binary effort, effort is easier to induce—and thus the optimal β is smaller—the smaller the cost of (mentoring) effort ψ_m and the larger the value of effort $\Delta p \Delta q$. In addition, in our setting, effort is easier to induce the higher the probability that a manager will stay with the firm, for the manager benefits from having trained a good worker only as long as she remains with the firm. Finally, effort is easier to induce the greater the importance of the worker, $1 - \kappa$, in the division's production.

Let $\overline{\psi}_x = q_b \beta^S$; it is the value of ψ_x where $\beta^x = \beta^S$. As indicated in Proposition 1, we will focus on the case $\beta^S \geq \beta^x$ where incentivizing mentoring is the binding constraint, and therefore consider the range of execution effort costs $\psi_x \in [0, \overline{\psi}_x]$. For larger values of ψ_x, (9) would become binding, and incentivizing execution effort would automatically incentivize mentoring as well. Even in this case, however, the choice between silos and lattice involves a tradeoff, unless the costs of execution effort are so high that providing the necessary incentives makes mentoring incentives irrelevant.

We show in Appendix A that if message-contingent contracts of the form $w_i = \alpha_{\widehat{\theta}} + \beta_{\widehat{\theta}} y_i + \gamma_{\widehat{\theta}} y_j$ for $\widehat{\theta} \in \{g, b\}$ (reports about the worker's type) are feasible, a pooling contract that coincides with the one of Proposition 1 is optimal. Although it is possible to separate a manager with a good worker from a manager with a bad worker, it is not possible to do so in a way that improves upon the simple output-based contract derived above. The reason is that separation requires paying a manager with a good worker $\beta_g \geq \beta^S$ in order to induce mentoring, but since a manager with a bad worker can select this contract too, there is no way for the firm to save on wage costs.

3.2 Lattice with transfer-contingent contracts and observable worker types

Consider now a lattice and assume that transfer-contingent contracts are feasible. In this subsection we also assume that the CEO can perfectly observe the workers' abilities. Aside from providing a useful benchmark, this case is interesting in its own right because some firms are good at getting to know people with leadership potential (see Section 5) to the point that there is little remaining asymmetric information. The problem of providing mentoring incentives remains, however.

The advantage of a lattice is that if (say) M_B leaves the firm, she can be replaced by W_A if he happens to be good. Having a lattice instead of silos makes a difference only if exactly one manager leaves (if both leave, each manager is replaced from within her own division or from outside, like in the silo case). Even then, a lattice matters only if the departing manager's worker is bad and the other division's worker is good, for otherwise the departing manager is replaced in the same fashion as with silos.

Thus, focusing on division A, outcomes differ from the silo case in two different events: with probability $\sigma(1-\sigma)p_A(1-p_B)$, manager M_B leaves and is replaced by W_A. Manager M_A, who stays, loses a good worker and must hire a replacement worker from outside. The resulting productivity loss for division A is $y_L = t_{mg} - t_{mw} = (1-\kappa)(1-p_w)\Delta q$. Conversely, with probability $\sigma(1-\sigma)(1-p_A)p_B$, manager M_A leaves and is replaced by W_B. W_B, in turn, is a good M_A with probability $\delta\phi$ and a bad one with probability $1-\delta\phi$. Compared to hiring a new M_A from outside, which is the next best alternative, the net productivity gain for division A is therefore $y_G = \delta\phi t_{gb} + (1-\delta\phi)q_b - t_{ob} = \kappa(\delta\phi - p_o)\Delta q$. It follows that in a symmetric equilibrium in which $p_A = p_B = p_h$, a lattice leads to a higher expected output than silos if and only if the net gain in output is positive:

$$y_G > y_L \Leftrightarrow \kappa(\delta\phi - p_o) > (1-\kappa)(1-p_w) > 0, \tag{11}$$

which is what condition (C1) in Section 2 states.

Stage 5 of the game, where the managers choose their execution effort, looks much the same as with silos: all uncertainty about manager and worker types is resolved, and the worst possible team productivity is $t_{bb} = q_b$. Because x_i affects only y_i, neither the salary component nor any bonus γ_n or γ_t for the other division's output affects execution incentives. The own-division bonus is either β_n or β_t depending on whether the own worker was transferred to the other division. Inducing managers to choose $x_i = 1$ in all situations leads to the same lower bound as with silos:

$$\beta_n, \beta_t \geq \frac{\psi_x}{q_b} = \beta^x. \tag{12}$$

We can now construct the firm's profit from division A, and M_A's payoff. For M_A's payoff, we need to leave p_A unspecified in order to determine M_A's mentoring incentive constraint. For the firm's profit, however, we can focus on the (equilibrium) case where both managers invest mentoring effort and hence $p_A = p_B = p_h$. For both payoffs, we take as given $x_A = x_B = 1$ as equilibrium choice.

Manager M_A's expected payoff as of stage 2 of the game can be obtained by starting with the payoff for a "silo case" without cross-divisional transfers, and adjusting it for the case in which if M_B leaves, M_A stays, and W_A is promoted to M_B, as follows:

$$\begin{aligned} V_A^L(p_A, p_B) &= \sigma\{\alpha_n + t_\mu(p(e))\beta_n + t_S\gamma_n - \psi_x x\} + (1-\sigma)\underline{U} - \psi_m e & (13)\\ &\quad +\sigma v p(e)\{\alpha_t + t_{\mu w}\beta_t + [\delta\phi t_{gb} + (1-\delta\phi)q_b]\gamma_t - \alpha_n - t_{\mu g}\beta_n - t_{ob}\gamma_n\} & (14)\\ \text{for } \mu &\in \{g, b\}. & (15)\end{aligned}$$

The first line of (13) is the same as (8), but with an added payment γ_n times the expected output of division B, under the equilibrium assumption that M_B chooses high effort. The second line states the probability of a promotion of W_A to M_B (assuming that M_A's effort is $p(e) \in \{p_h, p_l\}$, while M_B's is p_h), times (in $\{\}$) the difference in M_A's compensation relative to the silo case. The first three terms in the $\{\}$-brackets are payments to M_A, where after the transfer of W_A the productivity of division A is $t_{\mu w}$, and that of division B is $[\delta\phi t_{gb} + (1-\delta\phi)q_b]$. The last two terms in $\{\}$ subtract what M_A would be paid if her worker wasn't transfered, to correct for terms already included in the first line of (13). A contract that induces mentoring effort regardless of manager type must satisfy the following mentoring incentive constraint, which is straightforward to derive from (13):

$$\text{(MIC)} \quad \begin{aligned} &\xi \equiv \sigma\Delta p\{(1-\kappa)\Delta q\beta_n + v[\alpha_t + t_{\mu w}\beta_t - t_{\mu g}\beta_n \\ &+(\delta\phi t_{gb} + (1-\delta\phi)q_b)\gamma_t - t_{ob}\gamma_n]\} \geq \psi. \end{aligned} \quad \text{for } \mu \in \{g, b\} \qquad (16)$$

Condition (MIC), which by symmetry is the same for manager M_B, is derived under the assumption that the other division's manager has chosen high mentoring effort. Thus, it describes the condition for $e_A = e_B = 1$ to be a Nash equilibrium of stage 2 of the game.

Because of the terms $t_{\mu w}\beta_t - t_{\mu g}\beta_n = \kappa q_\mu(\beta_t - \beta_n) + (1-\kappa)(q_w\beta_t - q_g\beta_n)$ in ξ, (MIC) depends on the manager's own type, in contrast to the silo case. Specifically, if $\beta_n > \beta_t$, ξ is smaller, and thus (MIC) more restrictive, for $\mu = g$. Intuitively, because division A's output depends on the division *team*, losing a good worker to the other division matters more to a good manager, who is more likely to attain high output in her own division. If $\beta_t > \beta_n$, (MIC) is more restrictive for $\mu = b$.

Like M_A's payoff, the firm's expected profit from division A π_A^L can be determined by starting from the "silo case" without transfers, and adding and subtracting terms pertaining to a transfer of W_A or W_B, either of which occurs with probability $\sigma v p_h$:

$$\begin{aligned} \pi_A^L &= t_S(R - \beta_n - \gamma_n) \\ &\quad + \sigma v p_h \{-\alpha_t + (R - \beta_t)t_{mw} - [\delta\phi t_{gb} + (1 - \delta\phi)q_b]\gamma_t - (R - \beta_n)t_{mg} + t_{ob}\gamma_n\} \\ &\quad + \sigma v p_h[(R - \beta_n)y_G + y_L\gamma_n]. \end{aligned} \tag{17}$$

The first line of (17) corresponds to (7), with an added payment γ_n. The second line is the probability that a good W_A is promoted to a vacant M_B position, times (in $\{\}$) the associated profit, net of what the firm's profit would be without transfer. The third line of (17) covers the case in which M_A leaves and is replaced by W_B, and is easy to express in terms of the productivity gain y_G in division A and the loss y_B in division B because of the assumption that the new M_A is paid according to the contract terms $(\alpha_n, \beta_n, \gamma_n)$.

Denoting $\zeta = (\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t)$, the firm's contracting problem can thus be stated as

$$\max_{\zeta} \pi_A^L(\zeta) \text{ s.t. MIC}(\zeta) \text{ for } m \in \{g, b\}, \text{ (12), and } \alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t \geq 0. \tag{18}$$

The following result states a simple optimal contract:

Proposition 2 *With a lattice and observable worker types, an optimal contract that induces $e_A = e_B = x_A = x_B = 1$ is given by $\alpha_n = 0$,*

$$\alpha_t = \max\left\{0, \frac{\psi_m}{\sigma v \Delta p} - (1 - \kappa)\Delta q \frac{1 - v(1 - p_w)}{v q_b}\psi_x\right\}, \tag{19}$$

$\beta_n = \beta_t = \beta^x$, and $\gamma_n = \gamma_t = 0$. For $\psi_x = \overline{\psi}_x$, (19) reduces to $\alpha_t = (1 - p_w)/(\sigma\Delta p) > 0$.

Proposition 2 states that when the workers' types are known, an optimal contract consists of own-division bonuses β_n, β_t just large enough to induce execution effort, and a reward $\alpha_t > 0$ that is paid, say to M_A, conditional on the promotion of W_A to M_B but not conditional on either division's output. In plain words, it is optimal to pay an own-division output bonus to induce execution effort, and a simple "referral bonus" to induce mentoring.[17]

The intuition for Proposition 2 is that a good worker is most valuable if he is promoted to manager, irrespective how rare the event. Because both firm and managers are risk neutral, it is therefore optimal to put all mentoring incentives on this event (with a reward that is inversely proportional to the probability of the worker's promotion), which leads to the first term in the expression for α_t. Since the event of a transfer already rewards having a good worker, a simple referral bonus suffices, i.e. it is unnecessary to provide further incentives through β_t or γ_t. However, the own-division bonuses β_n and β_t must be at least β^x to induce execution effort. Since both bonuses incentivize mentoring as well (thought not as efficently as α_t), higher values of ψ_x and hence β_n and β_t lead to a lower optimal value of α_t in order to satisfy (MIC).

Observation 1 *If the firm's expected wage bill with a lattice is smaller than with silos, a lattice always dominates silos; if it is larger, then a lattice may or may not dominate silos.*

This observation (which does not depend on the assumed contract space) is straightforward: the firm's profit equals R times expected output, minus the expected wage bill (including wage payments to replacement managers). Since the expected output is always higher with a lattice than with silos, a lattice strictly dominates if the wage bill is lower as well. If the wage bill is higher, then it depends on the parameters whether a lattice or silos are optimal. In particular, since the wage bill does not depend on R, silos dominate a lattice if R is sufficiently small.

Proposition 3 *Consider a lattice with observable worker types and a wage contract as specified in Proposition 2. If* $\psi_x = 0$*, then the firm's expected wage bill with a lattice is smaller than with silos; if* $\psi_x = \overline{\psi}_x$*, the wage bill is larger.*

Proposition 3 shows that the execution effort plays an important part for the cost of a lattice. Fundamentally, mentoring effort is more productive with a lattice because good

[17] This is not the uniquely optimal contract but the simplest one: a profit-equivalent contract is one that puts all incentive weight on γ_t instead of α_t, in addition to $\beta_n = \beta_t = \beta^x$.

workers are more valuable if they can be promoted to the other division. It follows that if incentives can be sufficiently well targeted at the event of training a good worker, the firm's wage bill will be lower as well. That is the case if execution effort is costless, for then mentoring can incentivized by rewarding an event (transfer of the own worker) that can occur only if the worker is good.[18]

The more costly the execution effort, the greater is the necessary reward for own-divison output (β_n, β_t). Both of these wages also incentivize mentoring effort, allowing the firm to scale back α_t. However, as β_n and β_t approach β^S, the firm's wage bill must eventually exceed the silo wage bill: if $\beta_n = \beta_t = \beta^S$, both rewards combined would suffice to induce mentoring in a silo, where a manager is assured of keeping a good worker in her division. In a lattice, however, the manager loses a good worker to the other division with some probability, and gets a new worker who is worse in expected terms. It follows that $\beta_n = \beta_t = \beta^S$ alone cannot incentivize mentoring effort in a lattice. An additional reward $\alpha_t > 0$ is necessary, leading to a wage bill strictly higher than with silos.

3.3 Lattice with export-contingent contracts and privately informed managers

Now suppose only M_i knows the productivity of W_i. All payoff functions and constraints derived above remain unchanged, but new constraints come into play. With privately informed managers, the CEO can fill a vacant M_B position with a good W_A only if M_A (assuming she stays) reports the type of her worker truthfully. Drucker's observation that "the promotion system must ... make difficult alike kicking upstairs and hoarding good people" enters our analysis in the form of truthtelling incentive constraints that must hold for a manager with a good and a bad worker, respectively. Like in the case with full information, a lattice matters to M_A only if she stays and M_B leaves, and if W_B is bad but W_A good. We can therefore focus on this case in determining the conditions for

[18] In our model, the savings on wages with a lattice may well exceed the productivity gain $\sigma v p_h(y_G - y_L)$ because the ability to condition wages on a worker's transfer enables the firm to incentivize precisely what mentoring aims at, namely to produce a good worker. With silos, in contrast, this incentive can only be provided through noisy output incentives, which drive up the rents that need to be paid to managers.

truthtelling: even though M_A does not know W_B, irrelevant events simply cancel out on both sides of M_A's truthtelling constraints.

If M_A has a *good* worker, she will report the worker's type truthfully if

$$\text{(TTg)} \quad \alpha_t + \beta_t t_{mw} + \gamma_t[\delta\phi t_{gb} + (1-\delta\phi)q_b] \geq \alpha_n + \beta_n t_{mg} + \gamma_n t_{ob}. \tag{20}$$

The left-hand side of (TTg) is M_A's expected payoff if a good W_A is promoted to M_B and a new W_A is hired. The right-hand side of (TTg) is M_A's expected wage if she reports her good worker to be bad and thus "hoards" her worker, in which case the productivity of division A is t_{mg}, whereas in division B a new manager is hired and the bad worker remains. Like the mentoring incentive constraint (MIC), (TTg) depends on the manager's type whenever $\beta_n \neq \beta_t$: If $\beta_n > \beta_t$, (TTg) is more restrictive for a good manager. Intuitively, the difference $\beta_n - \beta_t$ matters more to a good manager, who is more likely to produce high output in her own division.

If M_A has a *bad* worker, she will report the worker's type truthfully if

$$\text{(TTb)} \quad \alpha_n + \beta_n t_{mb} + \gamma_n t_{ob} \geq \alpha_t + \beta_t t_{mw} + \gamma_t q_b. \tag{21}$$

The left-hand side of (TTb) is M_A's expected wage from keeping her bad worker, while in division B a new M_B is hired. If, on the other hand, M_A reports her worker W_A to be good, then W_A is promoted to M_B, in which case a new worker is hired in division A (resulting in team productivity t_{mw} there), while in division B both manager and worker are bad ($t_{bb} = q_b$). Condition (TTb), too, depends on the manager's type: If $\beta_n > \beta_t$, (TTb) is more restrictive for a *bad* manager, who has less to lose than a good manager from "kicking upstairs" a bad worker.

When managers have private information, the full-information contract of Proposition 2 is no longer feasible:

Lemma 1 *Any contract that is optimal with a lattice and full information violates (TTb).*

Intuitively, when the workers' types are known, the optimal contract pays an own-division output bonus and a simple referral bonus. With private information, however, managers who have a bad worker will take advantage of the referral bonus. This, precisely, is a main concern with referral bonuses in practice. Referral bonuses are common

(WorldatWork 2011, HRWorld 2008), but companies are aware of the perverse incentives they potentially create, and put safeguards in place. One is a careful screening process that approximates the observable-types case considered above. Another is to make the payment of a bonus contingent on the referred person's successful performance in the new position for a specified amount of time, such as six months or a year (Bilski 2011). This practice resembles the type of contract derived next:

Proposition 4 *Consider a lattice with privately informed managers, and assume (C3') holds. If $\psi_x = 0$, then an optimal contract is characterized by $\alpha_n = \gamma_n = \alpha_t = 0$, $\beta_n = \beta^S$, and*

$$\begin{aligned} \text{either (1) } \alpha_n &= \frac{1-\kappa}{\kappa} q_w \beta^S, \ \beta_t = \beta^S, \text{ and } \gamma_t = \frac{1-\kappa}{\delta\varphi\kappa}\beta^S; \\ \text{or (2) } \alpha_n &= \frac{1-\delta\phi\kappa}{\delta\phi\kappa} q_b \beta^S, \beta_t = 0, \text{ and } \gamma_t = \frac{1}{\delta\varphi\kappa}\beta^S. \end{aligned}$$

Contract (1) is optimal if $\delta\phi < 1$ and q_b sufficiently large; contract (2) is optimal if both $\delta\phi$ and σ are sufficiently close to 1. If $\psi_x > 0$, the optimal contract approximates one of the stated contracts subject to $\beta_n, \beta_t \geq \beta^x$; the full expressions are stated in the proof.

Thus, the optimal contract is characterized by $\alpha_n > 0$, $\beta_n = \beta^S$, $\gamma_t > 0$, and either $\beta_t = \beta^S$ or $\beta_t = 0$. To understand the result, recall first from Proposition 2 that mentoring is best incentivized by rewarding the transfer of a good worker. When the manager is privately informed about the worker, however, it is less costly to incentivize truthtelling through the use incentives (either $\beta_t > 0$ or $\gamma_t > 0$) than through an unconditional payment (α_t).

Because of Lemma 1, with privately informed managers a version of (TTb) must be binding for any optimal contract, which requires α_n, β_n or γ_n to be positive. Of these, β_n has a positive effect on mentoring incentives whereas both α_n and γ_n have a negative effect. Intuitively, therefore, it is least costly from an incentive perspective to satisfy (TTb) by choosing $\beta_n > 0$. The necessary level of β_n would lead to a violation of (TTg), however. Consequently, in both solutions stated in the proposition, (TTg) is binding as well, and given the resulting constraint on β_n, both solutions have $\alpha_n > 0$ to help satisfy

(TTb).[19]

The value of β_n at which (TTg) becomes binding happens to be the silo wage β^S. More generally, any contract that satisfies (MIC) and (TTg) with equality must have $\beta_n = \beta^S$. Intuitively, if (TTg) holds with equality, a manager with a good worker is indifferent between reporting her worker as good and having him promoted to manager, and reporting him as bad, in which case the no-transfer wages $(\alpha_n, \beta_n, \gamma_n)$ apply. In turn, incentivizing mentoring effort when the worker always stays in the division is exactly the silo situation and leads to $\beta_n = \beta^S$ (recall that α_n and γ_n have no effect on effort incentives in the silo case).

We showed above that with full information about workers, it depends on the cost of execution effort whether the firm's wage bill exceeds the wage bill with silos. With privately informed managers, the prediction is much clearer:

Proposition 5 *The wage bill with a lattice, using either of the contracts stated in Proposition 4, exceeds the wage bill with silos for any $\psi_x \in [0, \overline{\psi}_x]$.*

The intuition relates to the intuition for why $\beta_n = \beta^S$ for this contract: When (TTg) is binding, the manager receives the silo wage β^S if her worker is not transferred, and receives the same amount (in expected terms) if her worker *is* transferred. It follows that the lattice wage bill must be at least as large as in the silo case. In addition, however, (TTb) needs to be satisfied, which requires $\alpha_n > 0$ and thus leads to a wage bill strictly higher than with silos.[20]

[19] The argument relies on condition (C3'). If (C3') does not hold (which occurs if q_b/q_g is very small), the optimal contract may take a different form, with $\alpha_n > 0$, $\gamma_t > 0$, and all other contract variables set to zero. This may happen because while α_n negatively affects mentoring incentives, it can be effective in helping to satisfy (TTb), especially for a bad manager for whom the coefficient t_{mb} of β_n in (TTb) reduces to q_b, which makes it costly to satisfy (TTb) using β_n when q_b is very small.

[20] In contrast, if (C3') does not hold and hence the optimal contract takes a different form as described above, then as $q_b \to 0$, the contract converges to the full-information case according to Proposition 2. In that case, the wage bill may or may not exceed the silo wages bill, cf. Proposition 3.

3.4 Lattice with message-contingent contracts

Suppose now that a manager's wage cannot depend on the *transfer* of a worker but can depend on the manager's *report* about the worker's type. This scenario is relevant only when managers are privately informed about their workers. Like in Section 3.1, at stage 3 of the game, each manager reports her worker's type after observing it, and her wage can depend on her report: $w_A = \alpha_g + \beta_g y_A + \gamma_g y_B$ or $w_A = \alpha_b + \beta_b y_A + \gamma_b y_B$. What is different compared to the constraints (TTg) and (TTb) from Section 3.3 is that a manager's report matters not only when there is a vacancy in the other division, but unconditionally.

At stage 2, when the managers invest mentoring effort, M_A's expected payoff (assuming both managers exert execution effort later) can be constructed from the silo payoff (29) in Appendix A by adding a term that accounts for the event of a transfer of W_A:

$$V_A^{LM} = V_A^{SM} + \sigma(1-\sigma)p_A(1-p_B)(-y_L\beta_g + y_G\gamma_g). \tag{22}$$

At Stage 3, if M_A has a good worker, then her payoff from reporting truthfully (assuming high execution effort later, and assuming high mentoring effort on part of M_B) is

$$\sigma\{\alpha_g + t_{mg}\beta_g + t_S\gamma_g - \psi_x\} + \sigma v(-y_L\beta_g + y_G\gamma_g) + (1-\sigma)\underline{U}, \tag{23}$$

for $m \in \{g, b\}$ representing M_A's type. Her payoff from reporting having a bad worker instead is

$$\sigma\{\alpha_b + t_{mg}\beta_b + t_S\gamma_b - \psi_x\} + (1-\sigma)\underline{U}, \tag{24}$$

which reflects the fact that a worker reported to bad will not be transferred. Similar payoffs can be constructed for a manager with a bad worker who reports the worker's type truthfully or untruthfaully. Manager M_A's mentoring effort constraint can be constructed from (22), while her truthtelling constraints can be constructed from (23) and (24) and their counterparts for a manager with a bad worker. Let us move straight to the main result and leave the formal details for the proof:

Proposition 6 *Consider a lattice with privately informed managers, and assume (C3') holds. If message-contingent contracts are feasible, an optimal contract is given by* $\alpha_g =$

$\alpha_b = 0; \beta_g = \beta_b = \beta^S$;

$$\gamma_g = \frac{1-\kappa}{\delta\phi\kappa}\beta^S, \ \ and \ \gamma_b = \gamma_g - \frac{v(p_o - \delta\phi p_w)}{\sigma\Delta p\delta\phi t_S}\psi_m.$$

The logic of the result $\beta_b = \beta^S$ in Proposition 6 is similar to that of the result $\beta_n = \beta^S$ in Proposition 4: if M_A has a good W_A, she can always hold on to him by providing a false report to the CEO. If M_A does hold on to a good W_A, the bonus necessary to induce high effort is the silo bonus $\beta_b = \beta^S$, because the bonus γ_b has no incentive effect in this case.

In contrast to the silo case, the ability to transfer a manager's worker implies that it is possible to separate managers with a good and a bad worker, respectively. This can be achieved by choosing $\gamma_g > \gamma_b$ (which holds for the expressions stated becuase of Condition (C1)): a manager who truthfully reveals her good worker is rewarded by receiving a larger stake in the other division (which a manager with a bad worker would weakly forgo because kicking upstairs a bad worker would decrease the other division's productivity). Incentive compatibility then requires $\beta_g \leq \beta_b$, and because mentoring incentives are best provided through the own-division bonus, $\beta_g = \beta_b$ is optimal.

3.5 Lattice with simple output-based contracts

The contracts derived above may not be easy to implement, and hence may be rare to observe in reality. For instance, in the case of observable worker types, the referral bonus α_t according to Proposition 2 may be very large because it is inversely proportional to the probability of a relevant vacancy in another division (measured by v in the denominator of (19)). In reality, though, very large referral bonuses are rare; typical amounts (ranging from several hundred to several thousand dollars) offer adequate incentives for getting rank-and-file employees to refer friends, but are likely insufficient for getting a boss to refer a good employee of hers. Why do firms not just pay more, then?

One obstacle is that in reality, the impact of a 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 W_A is promoted to M_B at time period τ, the referring manager M_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 horizon may be 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.

A wage contract that is likely be feasible, however, is one conditional on the own and the other division's output *only*. In this section, therefore, we consider contracts of the form $w_i = \alpha + \beta y_i + \gamma y_j$. Under the assumptions of Section 2, it is optimal to set $\alpha = 0$. The expression for $\pi_A^L(p_A, p_B)$ in (17) simplifies to

$$\pi_A^L(p_h, p_h) = [t_S + \sigma v p_h(y_G - y_L)](R - \beta - \gamma), \tag{25}$$

and Manager M_A's expected payoff is (cf. (13))

$$V_A^L(p_A, p_h) = \sigma[t_\mu(p_A)\beta + t_S\gamma - \psi_x] + (1-\sigma)\underline{U} - \psi e_A + \sigma v p_A(-y_L\beta + y_G\gamma). \tag{26}$$

With privately informed managers, the truthtelling constraint (TTg) from Section 3.3 becomes

$$\beta t_{mw} + \gamma[\delta\phi t_{gb} + (1-\delta\phi)q_b] \geq \beta t_{mg} + \gamma t_{ob},$$

which can be expressed more simply as

$$\gamma y_G \geq \beta y_L$$

with y_G and y_L as defined in Section (3.2). This leads to a lower bound for γ as fraction of β:

$$\gamma \geq \gamma_g = \frac{y_L}{y_G}\beta = \frac{(1-\kappa)(1-p_w)}{\kappa(\delta\phi - p_o)}\beta. \tag{27}$$

Notice that $\gamma_g/\beta < 1$ per condition (C1). For a manager with a bad worker, the constraint (TTb) from Section 3.3 becomes

$$\beta t_{mb} + \gamma t_{ob} > \beta t_{mw} + \gamma q_b,$$

which, too, leads to requires γ to be a minimal fraction of β:

$$\gamma \geq \gamma_b = \frac{t_{mw} - t_{mb}}{t_{ob} - q_b}\beta = \frac{(1-\kappa)p_w}{\kappa p_o}\beta, \tag{28}$$

with $\gamma_b/\beta < 1$ per condition (4). Thus, with simple linear contracts, both truthtelling constraints can be satisfied by giving a manager a large enough stake in the other division. We can then show

Proposition 7 *Consider a lattice with privately informed managers, assume that only output-based contracts are feasible, and assume that (C3) holds. Then the optimal contract that induces $e_A = e_B = x_A = x_B = 1$ is given by $\beta = \beta^S$ and $\gamma = \gamma_g \beta^S$. The firm's expected wage bill is higher than with silos.*

Both β and γ provide incentives for mentoring, but as intuition would suggest, it is less costly to incentivize mentoring through β provided that (C3) holds, i.e., if the probability of a relevant vacancy is no too large. The only reason to choose $\gamma > 0$, then, is to induce truthtelling. Therefore, it is optimal to set $\gamma = \max\{\gamma_g, \gamma_b\}$, and it is straightforward to establish that (C1) implies $\gamma_g \geq \gamma_b$, which means that with an optimal contract, (TTg) is binding but (TTb) is not. The result $\beta = \beta^S$ then follows for the same reason we already encountered in Sections 3.3 and 6: if M_A has as a good W_A, she can always hold on to him by providing a false report to the CEO. If M_A does hold on to a good W_A, the bonus necessary to induce high effort is the silo bonus β^S. Paying $\gamma < \gamma_g \beta$ would have no effect on incentives because M_A would still hoard a good W_A. Paying $\gamma = \gamma_g \beta$ weakly induces truthtelling and thus makes M_A indifferent between hoarding and exporting a good worker. Compared to the case in which M_A hoards a good W_A, therefore, paying γ increases the firm's wage bill without providing additional effort incentives for the manager.

In conclusion, if wages can depend only on output and not on referrals, it is most efficient to incentivize mentoring just through the own-division bonus β. With observable types, the necessary bonus is larger than with silos because managers stand to lose a good worker with some probability *without* any direct compensation. With privately informed managers, it is still most efficient to incentivize mentoring through β, but the truthtelling constraint for both a manager with a good worker and a manager with a bad worker require a minimal other-division bonus γ. Compared to the case of observable types, wage costs are higher still, and thus a lattice costlier to implement, because of the additional adverse-selection problem.

3.6 Discussion of results

Let us interpret the results of Sections 3.1-3.5 in the context of Williamson's (1985) selective-intervention puzzle. Even without any other synergies between divisions A and B, selectively transfering employees between divisions is a form of value-increasing "selective intervention", much like other forms of resource reallocation.[21] Williamson and others have argued that intervention from the top tends to undermine division managers' incentives. This argument, however, relies on assumptions about contracting constraints. If division managers can be given a stake in the value-increasing interventions, then incentives need not be weakened, in which case integration is unambiguously optimal. We obtained this result in Proposition 3 for the full-information case and $\psi_x = 0$.

However, both costly execution effort and private information held by managers impose additional restrictions that create a tradeoff between the benefits of cross-divisional transfers (selective intervention) and the associated wage costs. The new restrictions work in different ways: execution effort creates a moral hazard problem with regard to mentoring, whereas private information creates adverse-selection costs, as follows.

Incentivizing costly execution effort requires rewarding own-division output, which prevents the firm from targeting incentives fully at the transfer of a good worker. As $\psi_x \to \overline{\psi}_x$, the wage bill approaches that of the silo case, but having $\beta_n = \beta_t = \beta^x$ does not suffice to incentivize mentoring even as $\beta^x \to \beta^S$, because in a lattice managers stand to lose a good worker to the other division. Thus, selective intervention directly "taxes"

[21] General Electric is a good example of a conglomerate held together significantly by synergies resulting form an actively managed corporate internal labor market; see Linebaugh (2012) and Footnote 23.

the managers and undermines mentoring incentives. The lattice is eventually more costly than are silos because inducing mentoring requires an additional reward in the form of $\alpha_t > 0$ or $\gamma_t > 0$.

In contrast, private information creates adverse-selection costs which alone can raise the wage bill above the silo case, even when execution effort is costless (Proposition 4). This is always the case if the truthtelling constraint (TTg) for a manager with a good worker is binding: Satisfying (TTg) requires paying the manager as much when the worker is transferred as when he is not transferred. That would suffice to induce mentoring, but would fail to prevent a manager from "kicking upstairs" a bad worker. Satisfying (TTb) as well, then, pushes the firm's wage bill above the silo case.

Wage costs are higher still, and definitely higher than with silos, when contracts cannot be conditioned on the transfer of a worker, whether they can be conditioned on messages about workers (Proposition 6) or only on outputs (Proposition 7).

A practically important implication of these arguments is that a lattice cannot be implemented simply by *allowing* cross-divisional transfers while retaining the same incentive system: a silo-like incentive contract with $(\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t) = (0, \beta^S, 0, 0, \beta^S, 0)$ would both undermine managers' mentoring effort and create incentives for them to misrepresent information about their people. Establishing a lattice, therefore, requires a change in the firm's incentive system that is likely to impose additional costs on the firm.

4 Silos or Lattice?

What forces drive a firm's choice between silos and a lattice? What motivates firms to change their internal labor markets from vertical job ladders to ones that facilitate cross-divisional mobility of employees? As it turns out, several parameters of our model have an unambiguous effect on this choice. Our first result concerns the value of output R:

Proposition 8 *For all three contracting scenarios (export-contingent, message-contingent, output-based), there exists $\tilde{R}$ such that a lattice with private information is preferred over silos if and only if $R \geq \tilde{R}$.*

The proof of the result is simple. Expected output is greater with a lattice than with silos (per (C1)), and the difference in output is multiplied by R. The expected wage cost, meanwhile, is higher with a lattice according to Propositions 5, 6, and 7, but the difference does not depend on R. It follows that there must be a critical value for R above which a lattice is more profitable, and below which silos are more profitable.

Observe that not just output but also the *difference* in the value of having a good employee in the position of a manager or worker is scaled by R. Thus, when employees in higher positions are more important, then a higher value of output implies not just a high value of having good employees in general, but a high value of moving the best people to the top.

We have so far left open how to interpret divisional output y_i and its value R. For instance, one can think of y_i not as actual output but as the outcome of efforts to innovate, improve quality, or reduce costs. A growing body of theoretical and empirical work suggests that the marginal value of such efforts is greater in more competitive product markets. The intuition is that in more competitive markets, firms' demand functions are more elastic, which magnifies the effects of cost or quality advantages or disadvantages, and therefore raises the value of investing to "stay ahead of the competition." This effect can be counterbalanced by a negative effect of competition on profits and hence the value of investments. However, when market structure is endogenously determined by free entry, the first effect tends to dominate (Raith 2003). Markets are more competitive the more substitutable the products and the greater the market. Results that establish a positive relation between product substitutability or market size and the value of investments can be found in Symeonidis (2000, Property 2), Raith (2003, Proposition 5), Vives (2008), and Martin (2009, Theorem 3). Evidence in line with this argument includes Syverson (2004), Cuñat and Guadalupe (2005, 2009), and Nocke and Asplund (2006).

Proposition 8 thus suggests a link between product market competition and the organization of internal labor markets. Intuitively, a general effect of product market competition is to raise a firm's marginal benefit of increasing quality or lowering costs. When doing so is a (hierarchical) team production process in which higher-level employees are more important, then greater competition also increases the value of moving talented employees into higher-level positions, and may make it worthwhile for the firm to incur the incentive costs of facilitating cross-divisional mobility. The link between competition and the value of searching for talent internally is recognized by practitioners too:

> "As global markets become more dynamic and competitive, companies will need to deploy talent even more flexible across broader swaths of the organization. Since management must develop and execute value-creating initiatives so quickly, talent is becoming more critical to corporate performance" (Bryan et al. 2006).

Our next result shows how choice between silos and lattice is driven by the external market for managers, and by the division specificity of human capital:

Proposition 9 *With export-contingent contracts, and with simple output-based contracts, the difference in profit between a lattice with private information and silos, $\pi^{LP} - \pi^{S}$, is decreasing in p_o.*

In words, a tighter external labor market for managers (smaller p_o) favors the adoption of a lattice organization internally. The result is driven by a direct effect on profit, and in the case of output-based wages, in indirect wage effect. The direct effect is that although both the silo output and the lattice output are increasing in p_o, the effect is smaller in a lattice because internal promotion of a worker is more likely than with silos, and hence hiring from outside (where p_o matters) less likely. With a lattice, a higher p_o also increases the incentive to hoard a good worker (tightens (TTg)), and lowers mentoring incentives (tightens (MIC)) but raises the costs of kicking upstairs a bad worker (relaxes (TTb)). These effects cancel out with export-contingent contracts, while the first effect reinforces the direct effect in the case of output-based contracts, by leading to an increase in γ.

Proposition 9 thus establishes a link between the "war for talent"—businesses' perception that recruiting good people externally has become harder—and the trend towards more elaborate internal talent management practices. Tighter labor markets, too, have been linked to greater product market competition; for a theoretical analysis see Gersbach and Schmutzler (2012). Propositions 8 and 9 thus jointly support the argument that increased product market competition, greater competition for talented people, and the ongoing reorganization of internal labor markets are all causally linked.[22]

Proposition 10 *For all three contracting scenarios, the difference in profit between a lattice with private information and silos,* $\pi^L - \pi^S$*, is increasing in* δ *and* ϕ*.*

Proposition 10 first states that a lattice is more likely to be preferred the larger the cross-divisional discount factor δ, which is intuitive: the more similar the managerial skill requirements across different divisions in a company, the greater the option value of moving employees across divisions, and hence the greater the profit with a lattice, whereas with silos δ is irrelevant for profit. In addition, a higher δ also relaxes both (MIC) and (TTg), which results in lower wages, cf. Propositions 4, 6, and 7.

A striking example of the link between division similarity and cross-divisional mobility is the case of erstwhile Newell Co. (today Newell Rubbermaid Co.). Newell's divisions produce picture frames, paintbrushes, curtain rods and countless other mundane products, and sell them to mass merchandisers such as Wal-Mart. Although the products are technologically unrelated, the business model of each division is the same, namely to be a "no-problem" supplier to its much larger customers (Montgomery 1999). As Montgomery argues, the synergies holding the corporation together revolve around the exchange of expertise among division managers and headquarters. To support this strategy, division managers frequently move from one division to another. The ease with which managers can move laterally is closely linked to the similarity of the divisions' operations (high δ in terms of our model).

The example, however, also points to the partial nature of our analysis: in reality, δ is endogenous as it results from the corporation's strategic decisions on what lines of business to pursue. We thus obtain an interesting link between corporate strategy and internal

[22] The causality need not be unidirectional from product market competition to external and internal labor markets. Greater mobility of managers, whatever the reason (see e.g. Murphy and Zabojnik 2004, 2007), could also lead to greater product market competition, thus reinforcing the causality we have emphasized. This could be the case, for instance, if the returns to superior managerial talent are not fully appropriated by the managers but in part also by (heterogeneous) firms hiring them.

labor markets, and a potentially testable prediction: other things equal, cross-divisional mobility is more likely in firms whose divisions are more similar.[23]

The second result of Proposition 10 is that a lattice is more likely to be preferred the larger the probability that a promoted worker is good as manager (ϕ). The logic is similar as for δ (as reflected in the appearance of $\delta\phi$ in many expressions of the analysis), except that a larger ϕ also increases the silo profit because by increasing the success of within-division promotion of a workers. Nevertheless, in a lattice the impact of ϕ on output is greater because of the higher probability of promoting a worker. Like δ, a larger ϕ also relaxes the constraints (MIC) and (TTg) and thus decreases γ_n and γ_t.

Proposition 10 shows that greater vertical mobility of employees should also tend to favor cross-divisional mobility, because a lattice relies on internal promotions more than does a silo organization. In practice, the same firm may offer different mobility opportunities for different types of jobs. Proposition 10 then suggests that measured across jobs, cross-divisional mobility of employees ought to be positively related to upward mobility. Like other parameters, ϕ is in practice endogenous and depends on managers' efforts not only to develop employees for their current jobs (as already captured by our model) but for higher-level jobs too. This suggests that across firms or jobs, cross-divisional mobility ought to be positively related to efforts to groom managers for higher-level positions.

5 Alternative Solutions

We confined our formal analysis to the design of incentive contracts in order to highlight two key obstacles to establishing cross-divisional mobility, managers' mentoring incentives and their private information about employees. In reality, of course, firms' efforts to reorganize their internal labor markets are more complex than that. In this section, we discuss potential and actual alternative solutions.

[23]Exceptions to this relation would be firms whose corporate strategy *is*, to a great extent, the internal labor market. The best known example is General Electric, whose divisions are very dissimilar but which as a corporation is famous for its talent management. Interestingly, GE itself appears to see limits to the viability of this strategy, see Linebaugh (2012).

Direct monitoring of mentoring effort: All solutions discussed so far provide incentives for mentoring that are based on outcomes. Some firms choose to measure mentoring efforts more directly. Doing so provides more targeted incentives, but is also costly to administer. Conaty and Charan (2010, page 160) cite the case of instrument maker Agilent, where (according to CEO Bill Sullivan) two thirds of the compensation of top managers are based on performance, while the other third is related to the HR development, or "building organizational capability by building the leaders of the future."

Aside from looking at mentoring outcomes, one way in which to measure mentoring effort is through 360-degree reviews, in which subordinates are asked about their boss's mentoring efforts. A successful example is mutual fund company MFS Investment Management, where evaluations of portfolio managers and analysts are each compiled from up to 60 evaluation forms submitted by peers, subordinates and superiors (Hall and Lim 2002)[24]. Most strikingly, these forms are not aggregated by HR staff but by C-level managers. The very elaborate process makes it feasible to ascertain not only performance "by numbers" but many other softer dimensions such as teamwork and mentoring.

The example of MFS is quite exceptional and illustrates the high costs. In general, rewarding inputs instead of outcomes is less reliable, may lead to distortions in the allocation of effort, and fails to take advantage of agents' (here, bosses) private information about how best to allocate their time (Baker 2002, Raith 2008).

Job rotation: General Electric, Novartis, SAS Institute and many others have programs in which junior and mid-career managers go through different assignments across functions and divisions (recall also our example of Newell Co. in Section 4).[25] The primary objective of job rotation is often to facilitate employees' human capital development and to prepare them for higher-level positions (e.g. Conaty and Charan 2010, 228-229). An additional benefit is that through rotation, talented employees become more visible to senior managers in different units, which reduces the asymmetry of information about

[24] We cite MFS as an example of an elaborate subjective evaluation processes. Due its lack of a divisional structure it is *not* an example of a firm with a lattice-like internal labor market.

[25] For a by now classical description about the effects of cross-firm rotation of the French engineering elite from the Ecole Polytechnique, see Crozier (1964).

employees held by senior managers. Job rotation thus alleviates the adverse-selection problem emphasized in this paper.

But job rotation also leads to worse incentives for the mentors, who have even less reason to invest in juniors who will leave their unit for sure, than when they stand to lose good people only if a better opportunity arises. A second downside of job rotation is that employees may end up learning too little about too many things. Through rotation, employees acquire many general managerial skills, but may have too little time and incentive to acquire division- (or industry-)specific expertise. After many years of emphasizing general managerial skills, General Electric recently decided to decided to keep senior managers in their divisions longer to help them acquire the expertise needed to compete in their industries (Linebaugh 2012). Finally, senior managers who rotate through leadership positions may be too focused on short-term results and may fail to make optimal long-run decisions for their divisions.

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, runs leadership courses and "work out" sessions in which high-potential employees are encouraged to publicly challenge their bosses' views. Former CEO Jack Welch used to travel to Crotonville every other week to meet the high-potential employees (Martin et al, 2010). Procter & Gamble puts much emphasis on social networks to make sure that people within the same cohort get to know each other and can draw on each other when staffing jobs.

Many other firms have adopted similar initiatives, cf. Conaty and Charan (2010). One could argue that as a result, the extent of remaining 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 2 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 too. 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 holders of private information.

6 Conclusion

Traditional vertical job ladders in firms—the principal subject of economic research on internal labor markets—have recently been giving way to active "talent management" aimed at optimally matching people with positions, which includes efforts to promote cross-divisional mobility of employees. We have argued in this paper that such efforts generally entail agency costs for the firm that arise because of the role that line managers (i.e., bosses) play in internal labor markets. Managers invest time to develop their employees' human capital through mentoring, and they acquire information about their employees' abilities that provides important input for personnel decisions.

Agency problems arise because managers prefer to have good employees working for them. It follows that efforts to increase employees' cross-divisional mobility undermine managers' incentives to invest in mentoring, and create incentives to use private information about employees strategically, either by "hoarding" good employees or by "kicking upstairs" bad ones. Our formal model captures the contractual origins of these agency problems: team production in firms, and an inability to measure mentoring effort or its outcomes directly. Our analysis explored how optimal contractual solutions that implement a "lattice" (under different contracting assumptions) differ from the simple incentive contracts that are optimal in traditional "silos."

Our results help explain the historical prevalance of silos (job ladders), and shed light on the challenges faced by companies transitioning to lattice structures. Silos lead to inefficient matches of people to positions, but create relatively good incentives for managers to mentor their employees. We have argued that firms' recent efforts to increase employees' mobility can be explained by a greater importance of getting the best people to the top, which in turn may have both internal or external causes such as skill-biased technological change, superstar effects, or product market competition.

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 in addition requires changes to the incentives provided to higher-level managers or other supporting practices such as job rotation, monitoring of mentoring effort, and other development initiatives. Consistent with our conclusion, the experience of many companies suggests that transitioning to lattice structures is harder than it looks; meanwhile, commentators lament a decline in managerial mentoring (Capelli 2012) that may be an unintended consequence of changes in internal labor markets.

To highlight the role of managers, we ignored the workers' (in real life, junior managers') point of view. Workers clearly stand to benefit from the greater mobility that a lattice offers. Pearson and Hurstak (1992) describe the consequences of Johnson & Johnson's silo structure in the early 1990s: "Many junior executive 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]." In response, top management took measures to facilitate cross-company mobility which, however, were unpopular with many senior managers. In any case, aside from the allocational advantage emphasized in our paper, a lattice can help firms to retain employees, and improves employees' own incentives to invest in firm-specific human capital. Cross-divisional mobility as a retention device is likely to be more valuable to firms with "flat" hierarchies in their divisions, in which the opportunities to "climb" are limited. This creates a link (and a potentially testable prediction) between the trends towards flatter organizational structures (Rajan and Wulf 2006, Guadalupe and Wulf 2010) and towards lattice structures.

Our paper departs from much of the literature in two main ways, each of which lends itself to further analysis. The first is our argument 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 the firm they work for. The second departure is to study the internal labor market of a *multi-divisional* firm, which extends the reach of economic analysis to questions of importance to today's large companies (Roberts 2004). Both departures, we hope, contribute to embracing within organizational economics Cyert and March's (1963) insight that managers spend substantial amounts of time on managing their coalitions by providing transfers or favors or access to information, or by influencing members through the means of communciation.

Appendix A: Silos with message-contingent contracts

Suppose that at stage 3 of the game, each manager reports her worker's type after observing it, and that her wage can depend on her report: $w_A = \alpha_g + \beta_g y_A + \gamma_g y_B$ or $w_A = \alpha_b + \beta_b y_A + \gamma_b y_B$. At stage 2, when deciding whether to invest in mentoring, M_A's expected payoff (assuming $p_B = p_h$, in equilibrium) then is

$$V_A^{SM} = \sigma\{p_A[\alpha_g + t_{mg}\beta_g + t_S\gamma_g] + (1-p_A)[\alpha_b + t_{mb}\beta_b + t_S\gamma_b] - \psi_x\} + (1-\sigma)\underline{U} - \psi_m e_A \quad (29)$$

As before, because there is no interaction between the divisions, it is optimal to set $\gamma_g = \gamma_b = 0$.

At stage 3, if M_A has a good worker, then her payoff from reporting truthfully (assuming high execution effort later, and assuming high mentoring effort on part of M_B) is $\sigma\{\alpha_g + t_{mg}\beta_g - \psi_x\} + (1-\sigma)\underline{U}$, while her payoff from reporting having a bad worker instead is $\sigma\{\alpha_b + t_{mg}\beta_b - \psi_x\} + (1-\sigma)\underline{U}$, for $m \in \{g, b\}$ representing M_A's type. Manager M_A thus has an incentive to report truthfully if

$$\alpha_g - \alpha_b + t_{mg}(\beta_g - \beta_b) \geq 0. \quad (30)$$

For manager with a bad worker, analogous payoffs can be derived, and lead to the truthtelling condition

$$\alpha_b - \alpha_g + t_{mb}(\beta_b - \beta_g) \geq 0. \quad (31)$$

Adding the left-hand sides of (30) and (31) leads to $\beta_g \geq \beta_b$ as necessary condition to satisfy both constraints. With an optimal contract, (31) must be binding. If it were not, then it would be optimal to set $\alpha_b = \beta_b = 0$ because of their negative effect on mentoring incentives (see (29)), which in turn would violate (31) and lead to a contradiction. Provided $\beta_g \geq \beta_b$, (30) then holds whenever (31) does.

Upon substitution of $\alpha_b = \alpha_g + t_{mb}(\beta_g - \beta_b)$ from (31) into (29)), and observing that $p_A t_{mg} + (1-p_A)t_{mb} = t_0(p_A)$, the manager's payoff (29) reduces to (8) with α_g, β_g in place of α, β. It follows from the analysis in Section 3.1 that $\beta_g = \beta^S$ is optimal, and that any contract that in addition satisfies (31) and $\beta_g \geq \beta_b$ is optimal. For instance, it is possible to separate managers with a good or bad worker, respectively, by setting $\beta_g = \beta^S, \beta_b = 0, \alpha_g = 0$, and $\alpha_b = t_{mb}\beta^S$: a manager with a good worker selects an incentive contract, while a manager with a bad worker chooses a fixed-wage contract. However, a pooling contract identical to a simple silo contract (with $\beta_g = \beta_b = \beta^S$) is optimal too. Thus it is not possible to get any extra mileage from message-contingent contracts. The reason is that separation requires paying a manager with a good worker $\beta_g \geq \beta^S$ in order to induce mentoring, but since a manager with a bad worker can select this contract too, there is no way for the firm to save on wage costs. We have thus shown:

Proposition 11 *With silos and when message-contingent contracts are feasible, a pooling contract with $\beta_g = \beta_b = \max\{\beta^S, \beta^x\}$ and $\alpha_g = \alpha_b = \gamma_g = \gamma_b = 0$ is optimal.*

Appendix B: Proofs

Proof of Proposition 1: Using (8), the incentive constraint $V_A^S(1,1) \geq V_A^S(0,1)$ takes the form

$$\sigma[t_\mu(p_h)\beta - \psi_x x] + (1-\sigma)\underline{U} - \psi_m \geq \sigma[t_\mu(p_l)\beta - \psi_x x] + (1-\sigma)\underline{U}, \tag{32}$$

which using (5) reduces to

$$\sigma(1-\kappa)\Delta p \Delta q \beta \geq \psi_m, \tag{33}$$

which leads to the expression for β^S stated in the proposition. If $\beta^S \geq \beta^x$, then β^S satisfies both (32) and (9), and thus the optimal contract is $(\alpha, \beta, \gamma) = (0, \beta^S, 0)$. QED

Proof of Proposition 2: We begin by separately stating the following lemma.

Lemma 2 *For the left-hand side of M_A's mentoring incentive constraint (MIC), μ, and the firm's profit π_A^L, we have*

$$\frac{\partial\mu/\partial\beta_n}{|\partial\pi_A^L/\partial\beta_n|} < \frac{\partial\mu/\partial\alpha_t}{|\partial\pi_A^L/\partial\alpha_t|} \quad \textit{and} \quad \frac{\partial\mu/\partial\beta}{|\partial\pi_A^L/\partial\beta|} < \frac{\partial\mu/\partial\alpha_t}{|\partial\pi_A^L/\partial\alpha_t|}, \tag{34}$$

where the derivatives in the second inequality are evaluted for $\beta_n = \beta_t \equiv \beta$.

Proof of the Lemma: From (16) and (13), the right-hand side of (34) is easily computed as $\Delta p/p_h$. For the left-hand side, we have

$$\begin{aligned} \frac{\partial \mu}{\partial \beta_n} &= \sigma \Delta p[(1-\kappa)\Delta q - v t_{\mu g}] \\ &< \sigma \Delta p\{(1-\kappa)\Delta q - v(1-\kappa)\Delta q] \\ &= \sigma \Delta p(1-v)(1-\kappa)\Delta q, \end{aligned} \tag{35}$$

where the inquality follows because $t_{\mu g} > (1-\kappa)\Delta q$ for $\mu \in \{g, b\}$; and

$$\begin{aligned} \left|\frac{\partial \pi_A^L}{\partial \beta_n}\right| &= t_S - \sigma v p_h t_{mg} + \sigma v p_h y_G > \sigma t_0(p_h) - p_t t_{mg} \\ &= \sigma(1-v)t_0(p_h) + \sigma v t_0(p_h) - \sigma v p_h t_{\overline{m}g} \\ &= \sigma(1-v)t_0(p_h) + \sigma v(1-p_h)t_{\overline{m}b} \\ &> \sigma(1-v)t_0(p_h). \end{aligned}$$

Overall, therefore, we have

$$\frac{\partial \mu/\partial \beta_n}{|\partial \pi_A^L/\partial \beta_n|} < \frac{\sigma \Delta p(1-v)(1-\kappa)\Delta q}{\sigma(1-v)t_0(p_h)} = \Delta p \frac{(1-\kappa)\Delta q}{t_0(p_h)} < \frac{\Delta p}{p_h},$$

where the last inequality follows from

$$t_0(p_h) = \kappa q_m + (1-\kappa)[p_h q_g + (1-p_h)q_b] < (1-\kappa)[q_b + p_h \Delta q] < (1-\kappa)\Delta q.$$

This proves the first inequality in (34). For the second inequality, observe that

$$\frac{\partial \mu}{\partial \beta_n} = \sigma \Delta p[(1-\kappa)\Delta q - v(t_{\mu g} - t_{\mu w})],$$

where $t_{\mu g} - t_{\mu w} = (1-\kappa)\Delta q p_w$ for both $\mu \in \{g, b\}$, which means that (35) applies to β as well, whereas $|\partial \pi_A^L/\partial \beta| > |\partial \pi_A^L/\partial \beta_n|$ simply because of the additional wage cost associated with β_t. The inequalities (34) thus apply to the case of $\beta_n = \beta_t \equiv \beta$ as well, which proves the second result. QED (Lemma).

To prove Proposition 2, we need to look for the cost-minimizing contract that satisfy both (12) and (16), which immediately restricts attention to contracts with $\beta_n, \beta_t \geq \beta^x$. Next, α_n incentivizes neither mentoring nor execution; set $\alpha_n = 0$. Next, γ_n has no effect on execution effort incentives and a negative effect on mentoring incentives, cf. μ in (16). Therefore, set $\gamma_n = 0$ as well. Next, (13) and (16) lead to

$$\frac{\partial \pi^L / \partial \gamma_t}{\partial \pi^L / \partial \alpha_t} = \frac{\partial V_A / \partial \gamma_t}{\partial V_A / \partial \alpha_t} = \delta \phi t_{gb} + (1 - \delta \phi) q_b.$$

It follows that mentoring incentives can be provided at the same expected cost through either α_t or γ_t. Let us focus on the simpler solution of using α_t rather than γ_t (if at all).

Consider first contracts with $\beta_t \geq \beta_n$, in which case (16) is (weakly) more restrictive for a bad manager. In this case we have

$$\frac{\partial \pi^L / \partial \beta_t}{\partial \pi^L / \partial \alpha_t} = t_{mw} > t_{bw} = \frac{\partial V_A / \partial \beta_t}{\partial V_A / \partial \alpha_t}. \tag{36}$$

Since $\partial \pi^L / \partial \beta_t$ and $\partial \pi^L / \partial \alpha_t$ are both negative, (36) implies that it is less costly to provide mentoring incentives through α_t than through β_t; hence an optimal contract must have β_t at the lowest possible level, hence $\beta_t = \beta_n$ in order to satisfy both $\beta_t \geq \beta_n$ and (12).

For contracts with $\beta_t < \beta_n$, (16) is (strictly) more restrictive for a good manager, and we obtain

$$\frac{\partial \pi^L / \partial \beta_t}{\partial \pi^L / \partial \alpha_t} = t_{mw} < t_{gw} = \frac{\partial V_A / \partial \beta_t}{\partial V_A / \partial \alpha_t}. \tag{37}$$

Because the sign is the opposite of that of (36), it follows that is optimal to incentivize mentoring through β_t rather than α_t. But providing incentives through α_t is still less costly than providing incentives through β_n per Lemma 2. It follows that it is less costly to incentivize mentoring through β_t than through β_n, in which case a contract with $\beta_t < \beta_n$ cannot be optimal.

We have shown so far that an optimal contract must satisfy $\beta_t = \beta_n$. According to Lemma 2, it is less costly to incentivize mentoring through α_t than through $\beta_n = \beta_t \equiv \beta$. It follows that a cost-minimizing contract is given by $\alpha_n = \gamma_n = \gamma_t = 0$, $\beta_n = \beta_t = \beta^x$, and α_t chosen to satisfy (16) given the values of the other contract variables. The expression is stated in the proposition in {}-brackets. The optimal contract is not unique: a different, cost-equivalent contract is one with $\alpha_t = 0$ and γ_t chosen to satisfy (16). Any convex combination of these two corner solutions is optimal too. For $\psi_x = \overline{\psi}_x$ and thus $\beta_n = \beta_t = \beta^S$, solving (16) given $\alpha_n = \gamma_n = \gamma_t = 0$ and $\beta_n = \beta_t = \beta^S$ leads to $\alpha_t = (1 - p_w)/(\sigma \Delta p)$, as stated. For larger values of ψ_x, the general expression for α_t

becomes negative. In this case, setting $\beta_n = \beta_t = \beta^x$ and all other bonuses to zero suffices to induce both mentoring and execution effort. QED

Proof of Proposition 3: For silos, the expected wage cost is $t_S\beta$, cf. (7), which using $\beta = \beta^S$ from Proposition 1 equals

$$\begin{aligned} \frac{t_S}{\sigma(1-\kappa)\Delta p\Delta q}\psi_m &> \frac{\sigma t_0(p_h)}{\sigma(1-\kappa)\Delta p\Delta q}\psi_m \\ &= \frac{\kappa q_m + (1-\kappa)[q_b + p_h\Delta q]}{(1-\kappa)\Delta p\Delta q}\psi_m > \frac{p_h}{p_h - p_l}\psi_m. \end{aligned} \quad (38)$$

For a lattice, the wage cost for the case $\psi_x = 0$ is obtained from (17) for $\alpha_n = \beta_n = \beta_t = \gamma_n = \gamma_t = 0$, and is simply $p_t\alpha_t = (p_h/\Delta p)\psi_m$, which in conjunction with (38) proves the first part of the proposition.

For the second part, recall that if $\psi_x = \overline{\psi}_x$, then $\beta_n = \beta_t = \beta^S$ with an optimal contract according to Proposition 2. Observe in (17) that if $\beta_n = \beta^S$, then the first term alone leads to a wage cost equal to the silo wage cost. In addition, the wage costs captured in the second and third terms in (17) are both strictly positive, because $\beta_t = \beta^S$ and $\alpha_t > 0$ according the last part of Proposition 2. It follows that the expected wage cost with a lattice is strictly higher than with silos. QED

Proof of Lemma 1: The proof is immediate from the contract of Proposition 2 and inspection of (TTb): any optimal contract has $\beta_n = \beta_t = \beta^x$, which puts more weight on the right-hand side of (TTb) because $t_{mw} > t_{mb}$. In addition, an optimal full-information contract has α_t and γ_t nonnegative and $\gamma_n = 0$. Condition (TTb) is therefore violated. QED

Proof of Proposition 4: The proof proceeds in several steps to derive the characteristics of an optimal contract $\zeta = (\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t)$.

1. If ζ is optimal, then $\alpha_t = 0$. Suppose instead that for $\zeta = (\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t)$, α_t is positive. Given that

$$\frac{\partial \pi}{\partial \alpha_t} = -p_t \text{ and } \frac{\partial \pi}{\partial \alpha_t} = -p_t[\delta\phi t_{gb} + (1-\delta\phi)q_b],$$

the contract $\zeta' = (\alpha_n, \beta_n, \gamma_n, \alpha_t - \Delta, \beta_t, \gamma_t + \frac{\Delta}{\delta\phi t_{gb} + (1-\delta\phi)q_b})$ leads to the same profit; similarly, it can be shown that both (MIC) and (TTg) are unaffected by replacing ζ with

ζ'. The same substitution, however, relaxes (TTb) by reducing the right-hand side by Δ times $1-\frac{q_b}{\delta\phi t_{gb}+(1-\delta\phi)q_b} > 0$. All constraints are satisfied, and it follows that if ζ is optimal, then ζ' must be optimal too. But because of Lemma 1, ζ' cannot be optimal because (TTb) is slack. It follows that ζ cannot be optimal either, proving the result.

2. Next we show that if ζ is optimal, then $\beta_n \geq \beta_t$. Suppose instead that $\beta_t > \beta_n$. In this case, the most restrictive version of both (MIC) and (TTg) is for $\mu = b$, whereas the most restrictive version of (TTb) is for $\mu = g$ (the two versions of each constraint coincide if $\beta_t = \beta_n$). In this case,

$$\frac{\partial\mu_A/\partial\gamma_t}{\partial\mu_A/\partial\beta_t} = \frac{\delta\phi t_{gb} + (1-\delta\phi)q_b}{t_{bw}}.$$

It follows that the contract $\zeta' = (\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t - \Delta, \gamma_t + \frac{t_{bw}}{\delta\phi t_{gb}+(1-\delta\phi)q_b}\Delta)$ satisfies (MIC) iff ζ does. The same holds for (TTg). In addition, ζ' relaxes (TTb) by Δ times

$$t_{gw} - \frac{q_b}{\delta\phi t_{gb} + (1-\delta\phi)q_b} t_{bw} > 0.$$

Finally, ζ' leads to an increase in profit by $p_t(t_{\overline{m}w} - t_{bw})$. Thus ζ' is feasible and dominates ζ. It follows that conditional on $\beta_t \geq \beta_n$, it is optimal to choose β_t as small as possible, which leads to $\beta_t = \beta_n$ and thus contradicts $\beta_t > \beta_n$. We thus know that an optimal contract satisfies $\beta_n \geq \beta_t$, which means that for the rest of the proof, the relevant version of both (MIC) and (TTg) is for $\mu = g$, whereas the relevant version of (TTb) is for $\mu = b$.

3. Next, if ζ is optimal, then $\gamma_n = 0$. Suppose instead that γ_n is positive. Then by inspection of (TTb) for $m = b$, the contract $\zeta' = (\alpha_n + \Delta t_{ob}, \beta_n, \gamma_n - \Delta, \alpha_t, \beta_t, \gamma_t)$ would leave (TTb) unaffected. The same holds for both (MIC) and (TTg) for $m = g$. Profit, in turn, changes by Δ times

$$\sigma(t_0(p_h) - t_{ob}) + (1-\sigma)[t_r(p_h) - t_{ob} - \sigma p_h(1-p_h)y_L]. \tag{39}$$

The first term in (39) is positive because $t_0(p_h) > t_{\overline{m}b} > t_{ob}$, where the first inequality follows from (5) and the second from the assumption $p_m > p_o$. The second term in (39) is positive because

$$\sigma p_h(1-p_h)y_L < p_h y_G < p_h\kappa(\phi - p_o)\Delta q$$

from (11), and

$$t_r(p_h) > p[\phi t_{gb} + (1-\phi)q_b] + (1-p)t_{ob}$$

from (6). It follows that ζ' is feasible and dominates ζ. Therefore ζ with $\gamma_n > 0$ cannot be optimal.

4. We have shown so far that any optimal contract satisfies: $\beta_n \geq \beta_t$, (MIC) binding for $m = g$, (TTb) binding for $m = b$, and $\alpha_t = \gamma_n = 0$. Solving (MIC) and (TTb) for α_n and γ_t (with $\alpha_t = \gamma_n = 0$) leads to

$$\begin{aligned}
\alpha_n &= \frac{q_b}{\sigma\upsilon\delta\phi\kappa\Delta p\Delta q}\psi_m - \left[\frac{1-\kappa-\upsilon}{\upsilon\delta\phi\kappa} + 1\right] q_b\beta_n \\
&\quad - \left[\frac{1-\delta\phi}{\delta\phi}q_b - (1-\kappa)p_w\Delta q\right]\beta_t \text{ and} \\
\gamma_t &= \frac{1}{\sigma\upsilon\delta\phi\kappa\Delta p\Delta q}\psi_m - \frac{1-\kappa-\upsilon}{\upsilon\delta\phi\kappa}\beta_n - \frac{1}{\delta\phi}\beta_t.
\end{aligned} \tag{40}$$

Any optimal contract must satisfy (40).

5. Restricting attention to corner solutions for now, it is optimal to choose β_n either as small or as large as possible, subject to constraints. Lower bounds to β_n are given by β^x (per (12)) and by β_t (per step 2 of this proof); an upper bound is given by (TTg). Consider thus a constract ζ with β_n within these bounds, and consider the effect of an increase in β_n of the firm's profit, when α_n and γ_t are adjusted according to (40) to satisfy (MIC) and (TTg):

$$\begin{aligned}
\frac{d\pi_A}{d\beta_n} &= \frac{\partial\pi_A}{\partial\beta_n} + \frac{\partial\pi_A}{\partial\alpha_n}\frac{\partial\alpha}{\partial\beta_n} + \frac{\partial\pi_A}{\partial\gamma_t}\frac{\partial\gamma_t}{\partial\beta_n} \\
&= -t_S + p_t(t_{mg} - y_G) - (1-p_t)\frac{\partial\alpha}{\partial\beta_n} - p_t[\delta\phi t_{gb} + (1-\delta\phi)q_b]\frac{\partial\gamma_t}{\partial\beta_n} \\
&= -t_S - \frac{\partial\alpha}{\partial\beta_n} + p_t\left\{t_{mg} - y_G + \frac{\partial\alpha}{\partial\beta_n} - [\delta\phi t_{gb} + (1-\delta\phi)q_b]\frac{\partial\gamma_t}{\partial\beta_n}\right\}.
\end{aligned} \tag{41}$$

To evaluate the expression in {} in (41), notice from (40) that $\frac{\partial\alpha}{\partial\beta_n} = \left(\frac{\partial\gamma_t}{\partial\beta_n} - 1\right)q_b$. The expression in {} thus equals

$$t_{mg} - y_G + \left(\frac{\partial\gamma_t}{\partial\beta_n} - 1\right)q_b - [\delta\phi t_{gb} + (1-\delta\phi)q_b]\frac{\partial\gamma_t}{\partial\beta_n} = t_{mg} - y_G - q_b - y_G\frac{\partial\gamma_t}{\partial\beta_n},$$

where

$$t_{mg} - y_G - q_b = [1 - \kappa(1 - p_m + \delta\phi - p_o)]\Delta q > 0,$$

$y_G > 0$, and $\frac{\partial\gamma_t}{\partial\beta_n} < 0$. It follows that the p_t-term in (41) is positive. For the remaining terms, we have

$$
\begin{aligned}
&-[\sigma t_0(p_h) + (1-\sigma)t_r(p_h)] - \frac{\partial \alpha}{\partial \beta_n} \\
&= \left(\frac{1-\kappa-\upsilon}{\upsilon\delta\phi\kappa} + 1\right) q_b - t_S.
\end{aligned}
\tag{42}
$$

Since $t_0(p_h)$ and $t_r(p_h)$ cannot exceed q_g, and $\delta\phi \leq 1$, a sufficient condition for (42) to be nonnegative is

$$
\left(\frac{1-\kappa-\upsilon}{\upsilon\kappa} + 1\right) q_b \geq q_g \Leftrightarrow \frac{q_b}{q_g} \geq \frac{\kappa\upsilon}{(1-\kappa)(1-\upsilon)},
$$

which is condition (C3'). It follows that $\frac{d\pi_A}{d\beta_n}$ in (41) is positive, which means that it is optimal to choose β_n as large as possible, subject to other constraints.

6. As mentioned, β_n is constrained by (TTg), which means that with an optimal contract, (MIC), (TTb), and (TTb) must all be binding. Any contract that satisfies (MIC) and (TTg) with equality must satisfy $\beta_n = \beta^S$: Substituting the right-hand side of (TTg) for $\alpha_t + \beta_t t_{mw} + \gamma_t[\delta\phi t_{gb} + (1-\delta\phi)q_b]$ in (16) cancels the entire []-term in (16) and leaves the same effort incentive constraint as for the silo case. Therefore, $\beta_n = \beta^S$.

7. It remains to determine whether the optimal value of β_t is its minimal value β^x or its maximal value $\beta_n = \beta^S$ (since $\beta_n \geq \beta_t$ from step 2). Proceeding as in step 5, varying α_n and γ_t to satisfy (MIC) and (TTb), we get

$$
\begin{aligned}
\frac{d\pi_A}{d\beta_t} &= \frac{\partial \pi_A}{\partial \beta_t} + \frac{\partial \pi_A}{\partial \alpha_n}\frac{\partial \alpha_n}{\partial \beta_t} + \frac{\partial \pi_A}{\partial \gamma_t}\frac{\partial \gamma_t}{\partial \beta_t} \\
&= -\sigma\upsilon p_h t_{mw} + (1-\sigma\upsilon p_h)\left[\frac{1-\delta\phi}{\delta\phi} q_b - (1-\kappa)p_w\Delta q\right] + \sigma\upsilon p_h[\delta\phi t_{gb} + (1-\delta\phi)q_b]\frac{1}{\delta\phi} \\
&= \frac{1-\delta\phi}{\delta\phi} q_b - (1-\kappa)p_w\Delta q + \sigma\upsilon p_h \left\{-t_{mw} - \frac{1-\delta\phi}{\delta\phi} q_b + (1-\kappa)p_w\Delta q + [\delta\phi t_{gb} + (1-\delta\phi)q_b]\frac{1}{\delta\phi}\right\},
\end{aligned}
$$

where the expression in {} reduces to $\kappa(1-p_m)\Delta q > 0$, and the remaining terms are positive if

$$
(1-\delta\phi)q_b > \delta\phi(1-\kappa)p_w\Delta q. \tag{43}
$$

Thus, if (43) holds, $\frac{d\pi_A}{d\beta_t} > 0$, and the optimal value of β is $\beta_n = \beta^S$. However, for $\delta\phi \to 1$, (43) does not hold. Moreover, $\sigma\upsilon p_h = \sigma(1-\sigma)p_h(1-p_h)$ is at most 1/16 and may be very small if σ is close to 1. The derivative $\frac{d\pi_A}{d\beta_t}$ may therefore just as well be negative, leading to $\beta_t = \beta^x$ for the optimal contract. QED

Proof of Proposition 5: From (17), and using $\gamma_n = \alpha_t = 0$ from Proposition 4, the expected wage cost with a lattice is

$$\alpha_n + t_S\beta_n + p_t\{t_{\overline{m}w}\beta_t + [\delta\phi t_{gb} + (1-\delta\phi)q_b]\gamma_t - \alpha_n - t_{\overline{m}g}\beta_n + y_G\beta_n\}.$$

Since $\beta_n = \beta^S$ according to Proposition 4, the term $t_S\beta_n$ equals the wage cost with silos. With a lattice, therefore, the wage cost differs by the expression

$$\Delta w = \alpha_n + p_t\{t_{\overline{m}w}\beta_t + [\delta\phi t_{gb} + (1-\delta\phi)q_b]\gamma_t - \alpha_n - t_{\overline{m}g}\beta_n + y_G\beta_n\}.$$

From (TTg) for $m = g$, we have

$$\alpha_n = t_{gw}\beta_t + [\delta\phi t_{gb} + (1-\delta\phi)q_b]\gamma_t - q_g\beta_n.$$

Substituting for α_n in the second term of Δw, we obtain

$$\Delta w = \alpha_n + p_t[(1-p_m)(\beta_n - \beta_t) + (\delta\phi - p_o)\beta_n)\kappa\Delta q,$$

which is positive because $\beta_n \geq \beta_t$ in both solutions of Proposition 4, and $\delta\phi > p_o$ from (C1). QED

Proof of Proposition 6: 1. We begin by stating the firm's profit function and the relevant constraints. The firm's profit from division A (assuming high mentoring and execution effort, and truthtelling by the managers) is

$$\begin{aligned}\pi_A^{LM} \;=\; & p_h\{-\alpha_g + \sigma t_{mg} + (1-\sigma)[\phi t_{gw} + (1-\phi)t_{bw}](R - \beta_g - \gamma_g)\} + \\ & (1-p_h)\{-\alpha_b + [\sigma t_{mb} + (1-\sigma)t_{ob}](R - \beta_b - \gamma_b)\} + \\ & \sigma v p_h[-(R-\beta_b)y_L - y_G\gamma_g + (R-\beta_b)y_G + y_L\gamma_b],\end{aligned} \tag{44}$$

where the first two lines correspond to the silo profit from a manager with a good worker and a bad worker, respectively. The third line captures the net effect of a lattice on profits, where the first two terms in [] are the net loss from losing a good worker to the other division, while the last two terms reflect the net gain, upon departure of M_A, from promoting a good W_B instead of hiring from outside. We assume here (although it is not critical for our results) that the promoted worker inherits the contract of the departing

M_A, which stipulates bonuses β_b and γ_b because M_A's worker is bad (otherwise W_A would have been promoted).

M_A's mentoring incentive constraint can be constructed from (29), adjusting for the possible transfer of W_A:

$$\begin{aligned}(\text{MIC}) \quad \mu &\equiv \sigma\Delta p\{\alpha_g - \alpha_b + t_{mg}\beta_g - t_{mb}\beta_b + t_S(\gamma_g - \gamma_b)\} \\ +\sigma v \Delta p(-y_L\beta_g + y_G\gamma_g) &\geq \psi_{m.}\end{aligned}$$

The truthtelling constraint for a manager with a good worker can be obtained from (23) and (24):

$$(\text{TTg}) \quad \sigma[\alpha_g - \alpha_b + t_{mg}(\beta_g - \beta_b) + t_S(\gamma_g - \gamma_b)] + \sigma v(-y_L\beta_g + y_G\gamma_g) \geq 0,$$

and the constraint for a manager with a bad worker can be derived similarly:

$$(\text{TTb}) \quad \sigma[\alpha_b - \alpha_g + t_{mb}(\beta_b - \beta_g) + t_S(\gamma_b - \gamma_g)] - \sigma v[(t_{mw} - t_{mb})\beta_g + q_b\gamma_g] \geq 0.$$

The subscripts m in all three constraints stand for $m = g, b$ depending on the manager's own type. Like in Section 3.3, therefore, there are two versions of each constraint, and which of them is more restrictive depends on the sign of $\beta_g - \beta_b$. A first observation is that with an optimal contract, one version of (TTb) must be binding: since α_b, β_b, γ_b all affect both (MIC) and (TTg) negatively, and profit anyway, it would be optimal to set all three variables to zero in the relaxed problem without (TTb). Doing so, however, would violate (TTb), contradiction.

2. Next, we show that if a contract ζ is optimal, it must be that $\beta_g = \beta_b$. Suppose first that $\beta_b > \beta_g$, in which case (MIC) and (TTg) are most restrictive for $m = g$, and (TTb) is most restrictive for $m = b$. Consider a variation of the contract that reduces β_b by $d\beta_b$ and increases α_b by $xd\beta_b$, for $x \equiv \frac{\partial\pi/\partial\beta_b}{\partial\pi/\partial\alpha_b} = \sigma t_{mb} + (1-\sigma)t_{ob} + p_h\sigma(1-\sigma)y_G$. By construction, the variation leaves the firm's profit unchanged. Moreover,

$$-\frac{\partial\mu}{\partial\beta_b} + x\frac{\partial\mu}{\partial\alpha_b} = \sigma\Delta p(t_{gb} - x),$$

and

$$t_{gb} - x = [1 - p_h\sigma(1-\sigma)]t_{gb} - \sigma t_{mb} - (1-\sigma)(1-p_h\sigma)t_{ob} + \sigma(1-\sigma)p_h(1-\delta\phi)(t_{gb} - q_g),$$

which is positive because $t_{mb}, t_{ob} < t_{gb}$ and the weights of t_{mb}, t_{ob} add up to that of t_{gb}; ie., $\sigma + (1-\sigma)(1-p_h\sigma) = (1-p_h\sigma(1-\sigma))$. Next,

$$\begin{aligned} -\frac{\partial TTb}{\partial \beta_b} + x\frac{\partial TTb}{\partial \alpha_b} &= \sigma t_{mb} + (1-\sigma)t_{ob} - q_b + \sigma(1-\sigma)p_h y \\ \text{and } -\frac{\partial TTg}{\partial \beta_b} + x\frac{\partial TTg}{\partial \alpha_b} &= [1-p_h\sigma(1-\sigma)]t_{gb} - \sigma t_{mb} \\ -(1-\sigma)(1-p_h\sigma)t_{ob} + \sigma(1-\sigma)p_h(1-\delta\phi)(t_{gb}-q_g) + t_{bg} - q_b &> 0 \end{aligned}$$

It follows that the described variation (which hinges on the assumption $\beta_b > \beta_g$) leaves profit unchanged but relaxes all constraints, which means that the variation cannot be optimal, and therefore ζ cannot be optimal either.

Suppose in contrast that $\beta_g \geq \beta_b$, in which case (MIC) and (TTg) are most restrictive for $m = b$, and (TTb) is most restrictive for $m = g$. If the inequality is strict, $\beta_g > \beta_b$, we can construct a variation of the contract that reduces β_g by $d\beta_b$ and increases α_g by $\frac{\partial TTg/\partial \beta_g}{\partial TTg/\partial \alpha_g} = t_{bg} - (1-\sigma)(1-p_h)(t_{bg} - t_{bw}) \equiv x$. By construction, the variation has no effect on (TTg) unchanged. Moreover,

$$\begin{aligned} -\frac{\partial TTb}{\partial \beta_g} + x\frac{\partial TTb}{\partial \alpha_g} &= t_{gb} - t_{bg} + v(t_{bg} - t_{bw} - t_{gb} + t_{gw}) > 0, \\ -\frac{\partial \mu}{\partial \beta_g} + x\frac{\partial \mu}{\partial \alpha_g} &= 0, \text{ and} \\ -\frac{\partial \pi}{\partial \beta_b} + x\frac{\partial \pi}{\partial \alpha_b} &= p_h\left\{\sigma t_{mg} + (1-\sigma)[\phi t_{gw} + (1-\phi)t_{bw}] - t_{bg} + (1-\sigma)v(t_{bg} - t_{bw})\right\} > 0, \end{aligned}$$

where the last inequality follows from $\sigma t_{mg} + (1-\sigma)[\phi t_{gw} + (1-\phi)t_{bw}] - t_{bg} = [\kappa\phi - (1-\kappa)p_w]\Delta q > 0$ because of (4). It follows that the described variation increases profit while relaxing or leaving unchanged all constraints, which means that ζ cannot have been optimal. Both steps together leave $\beta_g = \beta_b$ as only option for an optimal contract. For the rest of the proof, we will continue to use (MIC) and (TTg) for $m = b$, and (TTb) for $m = g$.

3. Next, we show that it is optimal to set $\alpha_b = 0$. Suppose in contrast that $\alpha_b > 0$, and consider a variation that reduces α_b by $d\alpha_b$ and increases γ_b by $xd\alpha_b$, for $x \equiv \frac{\partial TTg/\partial \gamma_b}{\partial TTg/\partial \alpha_b} = t_S$. By construction, the variation leaves (TTg) unaffected. Moreover,

$$-\frac{\partial TTb}{\partial \alpha_b} + x\frac{\partial TTb}{\partial \beta_b} = -\frac{\partial \mu}{\partial \alpha_b} + x\frac{\partial \mu}{\partial \beta_b} = 0, \text{ and}$$

$$-\frac{\partial\pi}{\partial\alpha_b}+x\frac{\partial\pi}{\partial\beta_b}=(1-p_h)\left\{\sigma[t_0(p_h)-t_{mb}]+(1-\sigma)[t_r(p_h)-t_{ob}]\right\}+p_t y_L>0,$$

which means the variation is feasible and leads to a higher profit.

4. Substituting $\beta_g=\beta_b\equiv\beta$ into (TTg) and (TTb) shows that both truthtelling constraints are negatively affected by β. It follows that the only possible reason to have $\beta>\beta^x$ is to incentivize mentoring through β. Since $\alpha_b=0$, it also means that an optimal contract will have $\gamma_b>0$ in order to satisfy (TTb). Moreover, adding (TTg) and (TTb) leads to

$$v[\delta\phi(t_{gb}-q_b)\gamma_g-(t_{bg}-t_{bw}-t_{gb}+t_{gw})\beta]\geq 0, \tag{45}$$

which, because the coefficient of β is positive, means that γ_g must be positive. The optimal contract (or set of optimal contracts) must therefore be a subset of the (non-negative) solutions of (MIC) and (TTb) for γ_g and γ_b. These are

$$\begin{aligned}\gamma_g &= \frac{\psi_m}{\delta\phi\kappa\sigma v\Delta p\Delta q}-\frac{(1-\kappa)(1-v)}{\delta\phi\kappa v}\beta, \\ \gamma_b &= \frac{t_S-v\kappa p_o\Delta q}{\delta\phi\kappa\sigma v t_S\Delta p\Delta q}\psi_m+\frac{\alpha_g}{t_S}-\frac{(1-\kappa)}{\delta\phi\kappa v t_S}[(1-v)t_S-\kappa v((1-v)p_o+vp_w\delta\phi)\Delta q]\beta.\end{aligned} \tag{46}$$

Focusing on corner solutions, there are two cases to consider: one where β is set to its minimal possible value consistent with (12), β^x, and one where β is chosen to be as large as possible consistent with both γ_g, $\gamma_b\geq 0$ and with (TTg) (as it turns out, the relevant constraint is (TTg)). To determine which is optimal, consider a contract ζ satisfying (46) and consider the effect on profit of an increase in β accompanied by changes in γ_g and γ_b according to (46). Thus, we are interested in the sign of

$$\frac{\partial\pi}{\partial\beta}+\frac{\partial\pi}{\partial\gamma_g}\left[-\frac{(1-\kappa)(1-v)}{\delta\phi\kappa v}\right]+\frac{\partial\pi}{\partial\gamma_b}\left[-\frac{(1-\kappa)}{\delta\phi\kappa v t_S}[(1-v)t_S-\kappa v((1-v)p_o+vp_w\delta\phi)\Delta q]\right] \tag{47}$$

From inspection of (44), we have $\frac{\partial\pi}{\partial\gamma_g}+\frac{\partial\pi}{\partial\gamma_b}=\frac{\partial\pi}{\partial\beta}$, which means that (47) equals

$$\frac{\partial\pi}{\partial\gamma_g}\left[1-\frac{(1-\kappa)(1-v)}{\delta\phi\kappa v}\right]+\frac{\partial\pi}{\partial\gamma_b}\left[1-\frac{(1-\kappa)}{\delta\phi\kappa v t_S}[(1-v)t_S-\kappa v((1-v)p_o+vp_w\delta\phi)\Delta q]\right]. \tag{48}$$

Since both $\frac{\partial\pi}{\partial\gamma_g}$ and $\frac{\partial\pi}{\partial\gamma_b}$ are negative, a sufficient condition for (48) to be positive is that both terms in [] be negative, which is the case if the absolute values of the coefficients of β in (46) are both greater than 1. For the first term, that follows from (C3), because

$1-\kappa > v$ implies $(1-\kappa)(1-v) > \kappa v > \delta\phi\kappa v$ and hence $\frac{(1-\kappa)(1-v)}{\delta\phi\kappa v} > 1$. For the second term, we need to show that

$$(1-\kappa)[(1-v)t_S - \kappa v((1-v)p_o + vp_w\delta\phi)\Delta q] - \delta\phi\kappa v t_S > 0.$$

To do so, we use $\delta\phi \leq 1$, $t_S \geq t_{ob} = q_b + \kappa p_o \Delta q$, $\kappa > 1-\kappa$, $\kappa p_o > (1-\kappa)p_w$ (per (C1)), and from (C3),

$$(1-\kappa)(1-v)q_b \geq \kappa v q_g \Leftrightarrow (1-\kappa-v)q_b + \kappa v q_b \geq \kappa v q_g \Leftrightarrow (1-\kappa-v)q_b \geq \kappa v \Delta q.$$

Then we have:

$$\begin{aligned}
& (1-\kappa)[(1-v)t_S - \kappa v((1-v)p_o + vp_w\delta\phi)\Delta q] - \delta\phi\kappa v t_S \\
\geq\ & (1-\kappa)[(1-v)t_S - \kappa v((1-v)p_o + vp_w)\Delta q] - \kappa v t_S \\
\geq\ & (1-\kappa-v)(q_b + \kappa p_o\Delta q) - \kappa v[(1-\kappa)(1-v)p_o + v(1-\kappa)p_w]\Delta q \\
\geq\ & \kappa v\Delta q + (1-\kappa-v)\kappa p_o\Delta q - \kappa v[\kappa(1-v)p_o + v\kappa p_o]\Delta q \\
=\ & \kappa v\Delta q + (1-\kappa-v)\kappa p_o\Delta q - \kappa^2 v p_o\Delta q \\
=\ & \kappa v(1-\kappa p_o)\Delta q + (1-\kappa-v)\kappa p_o\Delta q > 0.
\end{aligned}$$

Thus, if (C3') holds, then it is optimal to set β in (46) to its highest possible value, which is the value at which (TTg) binds.

5. Like in Proposition 4, solving (MIC) and (TTg) as equalities leads to $\beta_b = \beta^S$; therefore, $\beta_g = \beta_b = \beta^S$. Also, using (45), solving both (TTg) and (TTb) as equalities leads to $\gamma_g = \frac{1-\kappa}{\kappa\delta\phi}\beta^S$ as stated in the proposition. Next, solve (TTb) for γ_b, which leads to an expression that is increasing in α_g. With all constraints taken care of, α_g serves no purpose but drives up the wage cost; it is therefore optimal to set $\alpha_g = 0$, in which case the solution for γ_b is the one stated in the proposition. QED

Proof of Proposition 7: Consider first the relaxed problem without truthtelling constraints. Based on (26), and taking as given $p_B = p(e_B)$, M_A's incentive constraint for high effort is given by

$$\sigma\{[(1-\kappa)\Delta q - (1-\sigma)(1-p_B)y_L]\beta + (1-\sigma)(1-p_B)y_G\gamma\}\Delta p \geq \psi. \tag{49}$$

Since both π_A^L and V_A are linear in β and γ, a corner solution is generically optimal with either $\beta > 0$ and $\gamma = 0$, or $\gamma > 0$ and $\beta = 0$. The first case obtains if and only if

$$\frac{\frac{\partial^2 V_A}{\partial e_A \partial \beta}}{-\frac{\partial \pi_A}{\partial \beta}} > \frac{\frac{\partial^2 V_A}{\partial e_A \partial \gamma}}{-\frac{\partial \pi_A}{\partial \gamma}}. \tag{50}$$

With symmetric contracts and equilibrium effort levels, we have $\partial \pi_A / \partial \beta = \partial \pi_A / \partial \gamma$; cf. (25). Therefore (50) holds if and only if $\frac{\partial^2 V_A}{\partial e_A \partial \beta} > \frac{\partial^2 V_A}{\partial e_A \partial \gamma}$, for which the relevant expressions can already be gleaned from (49) because the term in $\{\}$ is $\partial V_A / \partial e_A$. From that term, and using $q_w \geq q_b$, $q_o \geq q_b$, $\delta\phi \leq 1$, $p_B = p_h$ in equilibrium, and Condition (C3), we obtain

$$\begin{aligned} \frac{\partial^2 V_A}{\partial e_A \partial \beta} &= (1-\kappa)\Delta q - v(1-\kappa)(1-p_w)\Delta q \\ &\geq (1-\kappa)\Delta q(1-v) \geq v\kappa\Delta q \geq v y_G = \frac{\partial^2 V_A}{\partial e_A \partial \gamma}. \end{aligned}$$

It follows that mentoring is least costly to incentivize with $\beta > 0$ and $\gamma = 0$.

Because $\gamma = 0$ with observable workers (according to part (a)), the only purpose of choosing $\gamma > 0$ is to ensure truthtelling. This means that the smallest (i.e. cost-minimizing) value of γ that satisfies both (TTg) and (TTb) is $\gamma = \min\{\gamma_g, \gamma_b\}$. It is straightforward to show that $\gamma_g \geq \gamma_b$ if and only if (C1) holds. In this case, $\gamma = \gamma_g \beta$ as stated. Substituting this expression for γ in (49) and solving for β leads to $\beta = \beta^S$. QED

Proof of Proposition 9: According to Proposition 4, $\beta_n = \beta^S$ and $\gamma_n = 0$, like in the silo case. The profit difference therefore equals the last two lines of (17), which reduce to p_t times

$$(R - \beta_t)t_{mw} - [\delta\phi t_{gb} + (1-\delta\phi)q_b]\gamma_t - (R - \beta_n)t_{mg} + (R - \beta_n)y_G \tag{51}$$

This expression is decreasing in p_o via y_G. Moreover, none of the wages stated in Proposition 4 depends on p_o, which completes the proof for export-contingent contracts.

For output-based contracts, using (7) and (25), and $\beta = \beta^S$ according to Proposition 7, we have

$$\begin{aligned} \Delta\pi &= [t_S + p_t(y_G - y_L)](R - \beta - \gamma) - t_S(R - \beta) \\ &= p_t(y_G - y_L)(R - \beta) - [t_S + p_t(y_G - y_L)]\gamma, \end{aligned}$$

where the first term is decreasing in p_o via y_G, and the second term is decreasing as well because $\gamma = \gamma_g \beta^S$ is increasing in p_o, and

$$t_S + p_t y_G = \sigma t_0(p_h) + (1-\sigma)t_r(p_h) + p_t y_G$$

is increasing in p_o as well. QED

Proof of Proposition 10: From (17) and Proposition 4, there are three effects of δ and ϕ on the lattice profit: the first is a positive effect proportional to $p_t(R-\beta_n)(t_{gb}-q_b)$ (corresponding to the event of a promotion of W_B to M_A). The second is a profit increase due to a decrease in α_n in version (2) of the optimal contract, or an unchanged profit in version (1). The third is through the term

$$-[\delta\phi t_{gb} + (1-\delta\phi)q_b]\gamma_t = -[q_b + \delta\phi\kappa\Delta q]\frac{x}{\delta\phi\kappa}\beta^S = -\left[\frac{q_b x}{\delta\phi\kappa} + \Delta q\right]\ \beta^S$$

for $x \in \{1, 1-\kappa\}$ depending on the version of the optimal contract. Either way, this term is increasing in δ and ϕ because γ_t is decreasing in δ and ϕ.

For message-contingent contracts, the direct effect of an increase in δ and ϕ on the profit (44) affects the terms $-y_G\gamma_g + (R-\beta_b)y_G$, and arguments similar to the case of export-contingent contracts establish that it is positive. The indirect effect through γ_g is positive too because γ_g is clearly decreasing in $\delta\phi$. For the effect through γ_b, we have

$$\begin{aligned}\gamma_b &= \gamma_g - \frac{v(p_o - \delta\phi p_w)}{\sigma\delta\phi t_S \Delta p}\psi_m = \frac{1-\kappa}{\kappa\delta\phi}\beta^S - \frac{v(p_o-\delta\phi p_w)}{\sigma\delta\phi t_S\Delta p}\psi_m \\ &= \frac{\psi_m}{\sigma\kappa\Delta p\Delta q t_S}\frac{t_S - v(p_o-\delta\phi p_w)\Delta q}{\delta\phi},\end{aligned}$$

which is decreasing in δ and ϕ because $t_S - vp_o\Delta q \geq t_{ob} - vp_o\Delta q = q_b + p_o\Delta q(\kappa - v) > 0$ per (C3). Thus, the indirect effect on profit, running through γ_b, of an increase in δ or ϕ is positive too.

For simple output-based contracts, finally, the direct effect in (25) runs through $p_t y_G(R-\beta-\gamma)$ and is positive provided a lattice is profitable at all ($R-\beta-\gamma > 0$). Moreover, both versions of γ stated in Proposition 7 are decreasing in δ and ϕ, reinforcing the positive effect on profit. QED

7 References

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

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

Alonso, Ricardo, Wouter Dessein and Niko Matouschek (2012): "When Does Adaptation Require Decentralization?", working paper.

Anders, George (2011), "The Rare Find: Reinventing Recruiting", *Business Week,* Oct. 13

Athey, Susan and John Roberts (2001), "Organizational Design: Decision Rights and Incentive Contracts", *American Economic Review Papers and Proceedings* 91(2): 200-5

Baker, George P. (2002): "Distortion and Risk in Optimal Incentive Contracts", *Journal of Human Resources*, Vol. 37, No. 4, pp. 728-751

Baker, George P., Michael C. Jensen and Kevin J. Murphy (1988), "Compensation and Incentives: Practice vs. Theory", *Journal of Finance,* Papers and Proceedings, 43 (3): 593-616

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.

Bilski, Jared (2011): "Rethinking the employee referral bonus", retrieved from http://www.cfodailynews.com/rethinking-the-employee-referral-bonus/ on Nov. 5, 2012

Bloom, Nick and John Van Reenen (2007): "Measuring and Explaining Management Practices Across Firms and Countries, *Quarterly Journal of Economics* 122, 1351-1408

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

Capelli 2008

Capelli, Peter (2011), "Why Companies Aren't Getting the Employees They Need", *Wall Street Journal*, October 26

Carmichael, Lorne H. (1988), "Incentives in Academics: Why Is There Tenure?", *Journal of Political Economy*, 96: 453-472

Chambers, Elizabeth G., Mark Foulon, Helen Handfield-Jones, Steven M. Hankin, Edward G. Michaels (1998), "The War for Talent", *McKinsey Quarterly*, August 1998

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

Chick, Gareth (2008): Line managers – the key to driving your talent forward", published by Spring Partnerships; retrieved from http://www.spring-partnerships.com/leftbrain/docs/GC%20

Cleaver, Joanne (2012): *The Career Lattice: Combat Brain Drain, Improve Company Culture, and Attract Top Talent.* McGraw-Hill.

Collings and Mellahi 2009

Conaty, Bill and Ram Charan, *Talent Masters*, Crown Business

Crozier, Michel (1964), *The Bureaucratic Phenomenon,* Chicago University Press

Cuñat, Vicente and Maria Guadalupe (2005), "How Does Product Market Competition Shape Incentive Contracts?", *Journal of the European Economic Association*, 3 (5): 1058-82

Cuñat, Vicente and Maria Guadalupe (2009), "Globalization and the Provision of Incentives Inside the Firm: The Effect of Foreign Competition", *Journal of Labor Economics,* 27 (2): 179-212

Cyert and March 1963

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

Dessein, Wouter (2012): "Incomplete Contracts and Firm Boundaries: New Directions", working paper.

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

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

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): 683-90

Fairburn, James A. and James M. Malcomson (2001), "Performance, Promotion and the Peter Principle" *Review of Economic Studies* 68 (1): 45-66.

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

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

Gersbach, Hand and Armin Schmutzler (2012): "Does Globalization Create Superstars?" Working paper, University of Zurich.

Gibbons, Robert, Richard Holden and Michael Powell (2012): "Organization and Information: Firms' Governance Choices in Rational-Expectations Equilibrium," *Quarterly Journal of Economics* 127 (4): 1813-1841.

Gibbons, Robert and Michael Waldman (1999a), "Careers in Organizations: Theory and Evidence", Chapter 36, *Handbook of Labor Economics*, Volume 3, Part B: 2373–2437, Elsevier.

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

Gibbons, Robert and Michael Waldman (2004): "Task-specific Human Capital", *The American Economic Review Papers and Proceedings*, 94 (2): 203-7

Giroud, Xavier and Holger M. Müller (2010),"Does Corporate Governance Matter in Competitive Industries?", *Journal of Financial Economics*, 95 (3): 312–331

Giroud, Xavier and Holger M. Müller (2011),"Corporate Governance, Product Market Competition, and Equity Prices", *Journal of Finance*, 66 (2): 563–599

Grossman, Gene M. and Elhanan Helpman (2002), "Outsourcing in a Global Economy", *The Review of Economic Studies* , 72 (1): 135-59

Guadalupe, Maria and Julie Wulf (2010), "The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization on Corporate Hierarchies", *American Economic Journal: Applied Economics,* 4 (): 105-127

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

HRWorld (2008): "Employee Referral Bonus Jackpots: 15 Companies with Awesome New-Hire Incentives," retrieved from http://www.hrworld.com/features/referral-bonus-jackpot-031108/ on Nov 5, 2012.

Peter, Lawrence J., and Raymond Hull (1969), "*The Peter Principle: Why Things Always Go Wrong*" (Morrow)

Kim, Jin-Hyuk (2011), "Peer Performance Evaluation: Information Aggregation Approach", *Journal of Economics & Management Strategy*, 20 (): 565–587

Lazear, Edward P. (2004), "The Peter Principle: a Theory of Decline," *Journal of Political Economy* Vol. 112 (S1), pp. S141-S163.

Lazear, Edward P., Kathryn L. Shaw and Christopher T. Stanton (2012), "The Value of Bosses", working paper, Stanford

Legros, Patrick, and Andrew Newman (2013): "A Price Theory of Vertical and Lateral Integration," *Quarterly Journal of Economics* 128 (2): 725-770.

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

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

Malcomson, James (1984): "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy* Vol. 92 (3), pp. 486-507

Marin, Dalia and Thierry Verdier (2003): "Globalization and the new enterprise", *Journal of the European Economic Association*, 1(2-3), 337-344.

Martin et al (2010)

Martin, Stephen (2009): "Microfoundations for the Linear Demand Product Differentiation Model, with Applications", working paper, Purdue University.

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

Montgomery, Cynthia (1999): "Newell Company: Corporate Strategy", HBS case 9-799-139

Murphy, Kevin J. and Jan Zabojnik (2004): "CEO Pay and Appointments: A Market-Based Explanation for Recent Trends", *American Economic Review*, Vol. 94, No. 2, Papers and Proceedings, pp. 192-196

Murphy, Kevin J. and Jan Zabojnik (2007): "Managerial Capital and the Market for CEOs". Working paper http://papers.ssrn.com/sol3/papers.cfm?abstract_id=984376

Nocke, Volker, and Marcus Asplund (2006), "Firm Turnover in Imperfectly Competitive Markets", *Review of Economic Studies* 73 (2), 295-327.

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*, Vol. 19 (2), pp. 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 Organization", *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

Raith, Michael (2003), "Competition, Risk and Managerial Incentives", *The American Economic Review* 93 (4): 1425-36

Raith, Michael (2008): "Specific Knowledge and Performance Mesasurement", RAND Journal of Economics" Vol. 39 (4), pp. 1059–1079.

Rajan, Raghuram G. and Julie Wulf (2006), "The Flattening Firm: Evidence on the Changing Nature of Firm Hierarchies from Panel Data", *The Review of Economics and Statistics*

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 and Conger 2007

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

Symeonidis, George (2000), "Price and Nonprice Competition With Endogenous Market Structure", *Journal of Economics and Management Strategy* 9 (1), 53-83.

Syverson, Chad (2004): "Market Structure and Productivity: A Concrete Example", *Journal of Political Economy* Vol. 112 (6), pp. 1181-1222

Vives, Xavier (2008): "Innovation and Competitive Pressure", Journal of Industrial Economics Vol. 56 (3), pp. 419–469

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

Waldman, Michael (2011), "Theory and Evidence in Internal Labor Markets," Forthcoming in *Handbook of Organizational Economics*

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*, Vol. 25 (3) pp.245-261

WorldatWork (2011): "Bonus Programs and Practices", retrieved from http://www.worldatwork.org/wa