

Friebel, Guido; Raith, Michael

Conference Paper

Managers, Training, and Internal Labor Markets

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel -
Session: Organizational Economics: Authority in Organizations, No. B23-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Friebel, Guido; Raith, Michael (2016) : Managers, Training, and Internal Labor Markets, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Organizational Economics: Authority in Organizations, No. B23-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/145666>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

University of Rochester

William E. Simon Graduate School of
Business Administration

The Bradley Policy Research Center
Financial Research and Policy
Working Paper No.
FR 13-31

December 19, 2013

Managers, Training, and Internal Labor Markets

Guido Friebe
Goethe University Frankfurt; Center for Economic Policy Research (CEPR); Institute for the
Study of Labor (IZA)

Michael Raith
University of Rochester – Simon Business School

This paper can be downloaded from the
Social Science Research Network Electronic Paper Collection:
http://papers.ssrn.com/paper/taf?abstract_id=2363492

Managers, Training, and Internal Labor Markets

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

December 19, 2013

¹Contact: gfriebel@wiwi.uni-frankfurt.de; michael.raith@simon.rochester.edu. We would like to thank Ricardo Alonso, Guy Arie, 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, USC, and the Spring 2012 NBER Organizational Economics Working Group meeting for their helpful comments and suggestions.

Abstract

We propose a theory that emphasizes the role of managers for the production and allocation of human capital in firms. Managers invest time to train junior employees, and acquire information about the juniors' abilities that is valuable for job assignments. This dual role of managers matters especially in a multidivisional firm, whose 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 manager's training incentives, and may encourage her to misrepresent the information she provides about her workers. We show that because of these agency problems, implementing a lattice to achieve better job assignments also leads to higher wage costs for the firm. As a result, either silos or a lattice can be optimal. Our comparative-statics analysis suggests that the recent trend for firms to facilitate cross-divisional mobility may be caused by product market competition, a tight managerial labor market, and skill-biased technological change.

Keywords: human capital, internal labor markets, agency problems, multi-divisional firm, intra-firm mobility, careers, training, war for talent

JEL-codes: D2, D8, L2, M5

1 Introduction

Human capital is one of the most important drivers of economic growth (e.g., Gennaioli et al., 2013). Aside from schools, firms are the main institutions in which human capital is created; and it is in firms that human capital is used for production. Following Becker (1964), economists have distinguished between investments in human capital that are made by employees themselves, and those made by the firm. The role of managers, or “bosses,” however, has been largely overlooked.

But managers play a crucial role both in the production and the allocation of human capital within firms, and this dual role leads to agency problems in the operation of internal labor markets that have received little attention in economics. We show that these agency problems matter most in the modern, multidivisional firm, and help 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 aims to extend the economics of internal labor markets to questions that many of today’s large companies grapple with.

Our starting point is the observation that a significant share of training in firms is carried out by managers who may have their own agendas. While formalized training may be overseen by Human Resources departments, its implementation is often in the hands of employees’ direct supervisors. For the purposes of this paper, training encompasses *mentoring* by bosses, which is important especially for the development of less tangible skills (Whittaker and Marchington 2003, Perry and Kulik 2008). It takes many forms, such as imparting leadership skills, providing advice on career decisions, giving psychological or social support (Scandura 1992), 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). It is inconceivable that this broad range of training activities could be centralized, because much of it occurs in the course of day-to-day interaction between managers and their subordinates. Our arguments apply without change to managers’ investments in the *recruiting* of junior employees.

These activities all have in common that they are important to the firm but costly to managers, which creates an incentive problem because both training and its outcomes are difficult to measure. The main payoff from training an employee, typically, is that a good employee raises 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. Learning about employees' abilities plays an important role in some economic theories, e.g., Gibbons and Waldman (1999). In reality, however, it may not be "the firm" that learns, but individual managers, who may be inclined to use their information to their own advantage, 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 training, and adverse selection with respect to information managers provide 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 firms, however, are increasingly concerned with filling positions with the best people from *anywhere* in the organization. Since the days of Doeringer and Piore, many firms have become bigger, more multidivisional, and more global. Managerial skills have become less industry-specific (Murphy and Zabojnik 2004). Moving employees through traditional job ladders is increasingly perceived as inefficient, and job ladders are increas-

ingly 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 (Card and Di Nardo, 2002; Gabaix and Landier, 2008). This evidence, too, points to an increased importance of filling top positions in firms with top talent.

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 some 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 to 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 training is to have a productive employee working for her, then reassigning talented employees to another part of the firm directly undermines training incentives.⁴ Second, in line with Drucker’s

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).

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 are aware of this problem: “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).

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 a lattice, instead, requires different 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. The manager’s ability matters more for production than the worker’s, consistent with other theories and with evidence (see our literature discussion below). Prior to production, the manager invests training 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 on whether or not a worker is transferred 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 from within the firm, or otherwise from outside. With silos (Section 4), only the worker in the same division is eligible for promotion.

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

⁶To 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.

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 training effort, knowing that her worker will stay in her unit; and there is no reason to misrepresent her worker’s ability to anyone. A lattice allows the firm to fill a managerial vacancy with a good worker from the other division, which we assume is preferred to hiring a manager from outside (otherwise silos would always be optimal).⁷

Our analysis proceeds in two steps: we first show (in Sections 4 and 5) that while a lattice leads to better job assignments than silos, under weak conditions a lattice also leads to a higher wage cost for the firm. The firm therefore in general faces a tradeoff in choosing between silos and a lattice, and sometimes silos are optimal. Given this tradeoff, our second set of results (Section 6) examines how different parameters of the model influence the choice between silos and a lattice.

A benchmark result (Section 5.2) highlights the conditions under which a lattice may lead to *lower* agency costs for the firm compared with silos, contrary to our main message. This is the case if all of the following conditions hold (i) the CEO (or a central HR department) can observe the workers’ abilities, (ii) contracts can be contingent on the transfer of a worker to the other division, and (iii) execution effort is unimportant. The optimal incentive contract then consists of a lump-sum referral bonus that is paid if the manager’s (good) worker is promoted to the other division. The resulting wage cost for the firm is *lower* than with silos, and hence a lattice unambiguously dominates silos, because a referral bonus directly rewards the event of developing a good worker through training, whereas with silos, training can be rewarded only indirectly through output-based incentive pay. Interpreting cross-divisional transfers as a form of “selective intervention” in the divisions’ affairs, this benchmark result illustrates that selective intervention need not undermine the division managers’ incentives but can actually strengthen them if they have a stake in the benefits of the intervention.

⁷We 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.

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

First, the more costly is execution effort (Section 5.3), the more the firm must reward own-division performance, which limits its ability to focus training incentives on the transfer of a good worker. Because each manager may end up losing her good worker to the other division, incentives that are sufficient to induce training with silos, are insufficient with a lattice, requiring the firm to pay an additional reward in the event of a transfer.

Second, if the firm is unable to reward transfers directly, for instance because wages can depend only on the divisions' outputs (Section 5.5), then the firm's wage bill is higher with a lattice than with silos. This result holds even if the CEO can observe the workers' abilities and execution effort is unimportant.

Third and most importantly, when managers are privately informed about their workers, then under a weak condition, inducing training is more costly with a lattice than with silos, even if bonuses can depend on worker transfers and if execution effort is irrelevant (Section 5.4). Simple referral bonuses are ineffective 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 that raises the wage cost to the firm above the level of silos. This result echoes a key argument in the recent literature on decisions in firms that centralization is costly if it requires eliciting information from division managers.⁸

We proceed to examine the influence of different parameters on the choice between silos and a lattice (Section 6). Our main result here is that 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

⁸See Alonso et al. (2008), Rantakari (2009, 2011), Dessein et al. (2010), and Friebe and Raith (2010).

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 to facilitate cross-divisional mobility are a relatively recent phenomenon, and why "talent management" within firms appears to go hand in hand with a "War for Talent" between firms (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 train 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 training 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 training by bosses observed by Capelli (2011).

2 Related Literature

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" does the learning. Our theory, in contrast, posits that bosses

usually know more about their workers than do others in the organization.⁹

Second, our theory focuses on managers as agents whose interests may diverge from the firm's. Fairburn and Malcolmson (1994) and Prendergast and Topel (1996) have looked at biased evaluations by managers, and Carmichael (1985) and Friebe 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 and 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. (2012), 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 one vacancy 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 (2008, 2011), Dessein et al. (2010), Friebe and Raith (2010), and Dessein (2012).

Finally, our results that link the choice between silos and lattice to product and labor market conditions links our paper to a diverse literature that studies the interaction

⁹Kim (2011), too, assumes dispersed information about workers. His model studies optimal contracting to achieve efficient evaluation of workers *by peers*.

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).

3 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.

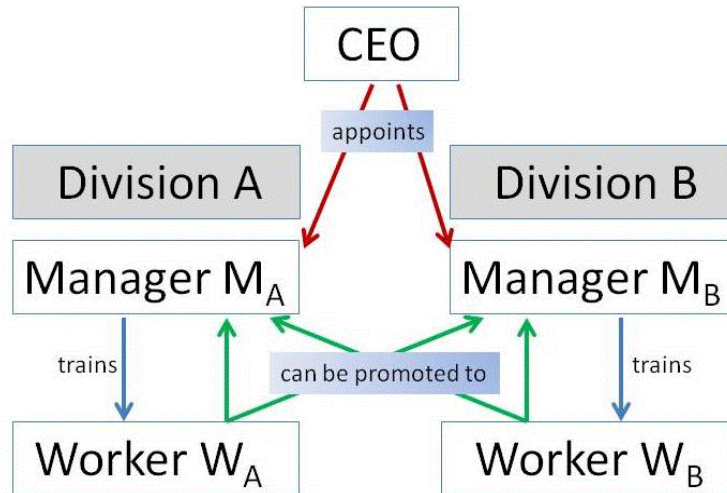


Figure 1: Organizational chart of the firm

3.1 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' training efforts and the CEO's assignment decisions. The “team productivity” of division i is defined as

$$t_i = \kappa_m q_i^m + \kappa_w q_i^w. \quad (1)$$

We assume that $\kappa_m \geq \kappa_w$ 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 $E[y|t] = t$, and a full support $[0, 1]$ for any 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 $\alpha = t/(1 - t)$.

Equation (1) has two implications. First, division i 's expected output is increasing in the productivity of both M_i and W_i , but it is impossible to infer the productivity of M_i or W_i from realized output. It follows that if a manager is compensated based on output (see Section 3.4), it is in her interest to have a good worker. Second, $\kappa_m \geq \kappa_w$ 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.

3.2 Timing

1. The firm appoints managers M_A and M_B (who may be hired from outside or transferred within the firm), 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 $q_m = p_m q_g + (1 - p_m) q_b = q_b + p_m \Delta q$. Managers do not know their type when accepting the position, but learn their type afterwards; a manager's type is therefore best interpreted as reflecting her *match* with the position.

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 training the worker, at cost τe_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 $\Delta p = p_h - p_l$. As explained below, we assume that it is optimal for the firm to induce training effort $e_i = 1$, leading to a probability p_h of training a good worker. Denote the resulting expected productivity of the worker by $q_w = p_h q_g + (1 - p_h) q_b = q_b + p_h \Delta q$.
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 ξx_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 training 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.¹⁰

3.3 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.¹¹ A manager

¹⁰We 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.

¹¹If $\phi < 1$, a worker might be promoted “beyond his level of competence” — as stated by the “Peter Principle,” see Peter and Hull (1969), Fairburn and Malcomson (2001), and Lazear (2004).

hired from outside is good with probability p_o ; denote the expected productivity of such a manager by $q_o = q_b + p_o\Delta q$. 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 we assume is recruited and trained in the same way as the workers at stage 2 of the game. Thus, a replacement worker is good with probability p_h , and his expected productivity is q_w as defined above.

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 (Section 3.5), 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.¹²

To fill a vacancy optimally, the CEO needs to *know* 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.¹³ In contrast, a manager who stays with the firm may have 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 at least prove) their own suitability for a managerial position, see also Friebe and Raith (2004).

In our model, turnover of workers is generated only by manager departures. A bad worker is not fired even after his manager or the CEO has learned his type. This assump-

¹²Because the divisions are symmetric, it is never optimal to transfer a manager laterally to the other division. In contrast, if e.g. division 2 were more important, it might be preferable to transfer a proven M_A to division B instead of promoting a worker.

¹³In 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 incentive to be honest about suitable successors.

tion approximates the real-life regularity that the performance threshold to be considered for promotion is typically much higher than the threshold below which an employee is likely to be terminated.¹⁴

3.4 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 training and execution effort:

$$U_i = w_i - \tau e_i - \xi x_i.$$

A manager accepts to work for the firm if her expected equilibrium utility exceeds a reservation utility level. To simplify the analysis, we assume that the reservation utility is low enough such that the limited-liability constraint is binding and the participation constraint is not. The firm's profit from division i is given by

$$\pi_i = Ry_i - w_i$$

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 $E(y_i|t_i) = x_i t_i$.

We consider two kinds of contracts. First, 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.¹⁵

¹⁴For instance, if employees are evaluated on a forced curve, only those in the top bracket might be considered for promotion, while only those in the bottom bracket would be in danger of losing their job. Those in the (normally largest) middle bracket would just remain in their jobs. Likewise, companies that use the so-called "9 box grid" evaluate employees on both performance and on potential, and it is perfectly possible for employees to be rated effective in their current job but unsuited for higher positions.

¹⁵If Ry represents revenue or gross profit, it might seem more natural for incentive contracts to be

Second, we consider 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 verifiable, it is plausible that wages can be tied to *positions* (a standard argument in the internal labor markets literature) and hence also to observed changes in positions. The only twist here is that the *manager's* wage is tied to changes in her *worker's* position.

Any manager who leaves 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. In general, however, our results do not depend on the details of how replacement managers are paid.¹⁶

Even though the two divisions may be linked through personnel reallocations or the wage contracts or both, 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 possible payments to M_A (the original manager or a replacement) based on division B's output, but does not include payments made to M_B based on A's output.

3.5 Parameter conditions

We assume that with both silos and a lattice, it is optimal for the firm to induce both training and execution effort; that is, $e_i = x_i = 1$. We refrain from stating the relevant

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 6) have no effect on the contract, whereas in the alternative version the bonus parameters would be decreasing in R .

¹⁶Strictly speaking, since our model is static, replacement managers need not be given any training incentives. In an ongoing organization, though, new managers need to train workers too, which is why it is reasonable to assume that they inherit the same wage contracts.

parameter conditions formally; they amount to upper bounds on the effort costs τ and ξ . 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 training effort cost (or low marginal benefit of effort), for instance, the firm would sometimes prefer a lattice with low training effort to silos with high effort. The tradeoffs at work, however, are the same as in our analysis below.

Two parameter conditions ensure that it is optimal to fill vacant manager positions as described above. To state them, denote—exploiting the additive structure of (1)—the productivity of a manager-worker team with *expected* productivities $q_i^m = q_x$ and $q_i^w = q_y$, by $t_{xy} = \kappa_m q_x + \kappa_w q_y$, where possible values of x and y are $x \in \{g, b, m, o\}$ and $y \in \{g, b, w\}$.

The first parameter condition is that the firm prefers to *promote a good worker from the other division*, than to hire a manager from outside. A necessary condition is that the division with the vacancy—whose own worker is bad, or else he would have been promoted—must be better off, which requires

$$\delta\phi t_{gb} + (1 - \delta\phi)t_{bb} > t_{ob} \quad (2)$$

or equivalently, a positive net productivity *gain*:

$$y_G \equiv \kappa_m(\delta\phi - p_o)\Delta q > 0. \quad (3)$$

Condition (3) is equivalent to $\delta\phi \geq p_o$, which means that a good worker from the other division has a greater chance of becoming a good manager than does a random outside hire.

However, the other division loses a good worker and must hire a new one, which leads to a productivity *loss* of

$$y_L = t_{mg} - t_{mw} = \kappa_w(1 - p_h)\Delta q \quad (4)$$

Promoting a good worker from the other division is optimal for the firm, then, if and only if $y_G \geq y_L$, or

$$(C1) \quad \kappa_m(\delta\phi - p_o) > \kappa_w(1 - p_h).$$

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.

The second condition specifies when *hiring from outside* is preferred to promoting a worker. At the very least, an outsider must be preferred to promoting a worker known to be bad. We will make a stronger but plausible assumption, however, that the firm prefers to hire an outsider than promote a worker whom it knows nothing about. Without additional information, the firm's rational expectation about a worker's productivity is the equilibrium value q_w . Taking into account that a promoted good worker becomes a good manager only with probability ϕ , hiring from outside is then preferred if

$$\kappa_m q_o + \kappa_w q_w > \kappa_m [p_h \phi q_g + (1 - p_h \phi) q_b] + \kappa_w q_w,$$

which is equivalent to

$$(C2) \quad p_o > \phi p_h. \tag{5}$$

Below, we assume throughout that (C1) and (C2) hold.

4 Silos

To determine optimal contracts for silos, we begin by constructing the firm's and the managers' payoffs. A quantity of interest that appears in many results is the expected productivity of a division as of stage 1 of the game, taking into account the possible departure of the manager, and the stochastic outcome of the manager's training effort. It can be recursively constructed as a weighted average of the resulting productivity if the manager stays, denoted $t_0(p)$, and if she leaves, $t_r(p)$ where $p \in \{p_h, p_l\}$ is the probability of ending up with a good worker depending on the manager's training effort:

$$t_S(p) = \sigma t_0(p) + (1 - \sigma) t_r(p).$$

Its components are given by

$$t_0(p) = p t_{mg} + (1 - p) t_{mb} = \kappa_m q_m + \kappa_w (q_b + p \Delta q) \tag{6}$$

$$t_r(p) = p [\phi t_{gw} + (1 - \phi) t_{bw}] + (1 - p) t_{ob}. \tag{7}$$

The expression for $t_0(p)$ is simply the productivity of a team whose manager has an expected ability of q_m , while the worker's productivity depends on the manager's training effort. The expression for $t_r(p)$ represents the case where the manager leaves and must be replaced (hence the subscript 'r'). If the manager has a good worker, 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, then the manager's position is filled from outside while the worker stays in his position. Abusing notation, we will use $t_0 = t_0(p_h)$, $t_r = t_r(p_h)$, and $t_S = t_S(p_h)$ whenever we focus on an equilibrium that induces high training effort.

With silos, the firm's expected revenue in division A is simply Rt_S . Contracts conditioned on transfers are irrelevant; so let us consider simple output-based contracts (α, β, γ) . It is immediate, however, that setting $\alpha = \gamma = 0$ is optimal: There is no reason to choose $\gamma \neq 0$ because there is no interaction between the divisions. The salary α , in turn, has no effect on the manager's incentives, and because by assumption the managers' participation constraint is slack, it follows from limited liability that $\alpha = 0$ is optimal. The firm's wage cost is therefore simply βt_S .

Two incentive constraints need to be satisfied to induce both training and execution effort. Managers exert execution effort in stage 5 of the game. High execution effort leads to a payoff for the manager of $\beta t_i - \xi$, whereas low effort leads to a payoff of 0. The resulting incentive constraint $\beta t_{bb} - \xi \geq 0$ is most restrictive for the worst possible team with productivity t_{bb} . For this team, inducing execution effort requires

$$\beta \geq \frac{\xi}{t_{bb}} =: \beta^x. \quad (8)$$

The same lower bound β^x for the own-division bonus applies to the case of a lattice analyzed in Section 5 below, regardless of contracting assumptions, because execution effort is exerted after any transfers have taken place.

At stage 2 of the game, M_A invests in training her worker. At this point, she knows her own type $\mu \in \{g, b\}$, but not whether she will stay or leave. If she stays, the expected division output at stage 2 of the game is $t_\mu(p) = pt_{\mu g} + (1 - p)t_{\mu b}$ depending on the

probability p of training a good worker. M_A 's net payoff then is

$$V_A^S(\mu, e, x) = \sigma[t_\mu(p(e))\beta - \xi x] + (1 - \sigma)\underline{U} - \tau e, \quad (9)$$

where the two effort costs enter the payoff differently because the training effort e is invested before, but the execution effort x after, M_A knows whether she will stay with the firm. The resulting training incentive constraint reduces to $\sigma t_\mu(p_h)\beta - \tau \geq \sigma t_\mu(p_l)\beta$, which (noting that $\frac{\partial t_\mu(p)}{\partial p} = \kappa_w \Delta q$ regardless of M_A 's type) is straightforward to solve for β :

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

$$\beta^S = \frac{\tau}{\sigma \kappa_w \Delta p \Delta q}. \quad (10)$$

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 the Appendix. 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 (training) effort τ 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, because 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, κ_w , in the division's production.

Let $\bar{\xi}_x = t_{bb}\beta^S$; it is the value of ξ where $\beta^x = \beta^S$. As indicated in Proposition 1, we will focus on the case $\beta^S \geq \beta^x$ where incentivizing training is the binding constraint, and therefore consider the range of execution effort costs $\xi \in [0, \bar{\xi}_x]$. For larger values of ξ , (8) would become binding, and incentivizing execution effort would automatically incentivize training 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 training incentives irrelevant.

Lower wage cost with lattice			
	Execution effort cost-less	CEO can observe worker abilities	Transfer-contingent contracts
Proposition 2	✓	✓	✓
Higher wage cost with lattice			
Proposition 3	Costly execution effort	✓	✓
Proposition 4	✓	Privately informed managers	✓
Proposition 5	✓	✓	Simple output-based contracts
Proposition 6	✓	Privately informed managers	Simple output-based contracts

Table 1: Overview of results of Section 5

5 Lattice

For the lattice, we begin with deriving payoffs and constraints, and then derive a benchmark result that states conditions under which a lattice unambiguously dominates silos because it leads to better job allocations and lower wage costs. Our subsequent results show that relaxing any one of the conditions overturns the benchmark result by implying a higher wage cost with a lattice, creating a tradeoff for the firm. The relations between this section's results are shown in Table 1.

5.1 Payoffs and effort constraints

With a lattice, a good worker can be promoted to manager in the other division. This can happen only if exactly one manager leaves, and 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, with probability $\sigma(1-\sigma)(1-p_B)p_A$, W_A is promoted to M_B , which leads to a net productivity loss of y_L in division A compared

with silos, as discussed in Section 3.5. And with probability $\sigma(1 - \sigma)(1 - p_A)p_B$, W_B is promoted to M_A , leading to a net productivity gain y_G in division A. Denote by $v = (1 - \sigma)(1 - p_h)$ 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 training effort). Then in equilibrium, when $p_A = p_B = p_h$, both promotion probabilities are equal to $\sigma v p_h$.

The firm's expected wage cost from division A, w_A^L , can be determined by starting from the "silo case" without cross-divisional transfers, and adding and subtracting terms pertaining to a transfer of W_A or W_B :

$$\begin{aligned} w_A^L &= \alpha_n + t_S(\beta_n + \gamma_n) \\ &\quad + \sigma v p_h \{ \alpha_t + t_{mw}\beta_t + [\delta\phi t_{gb} + (1 - \delta\phi)t_{bb}]\gamma_t - \alpha_n - t_{mg}\beta_n - t_{ob}\gamma_n \} \\ &\quad + \sigma v p_h (y_G\beta_n - y_L\gamma_n). \end{aligned} \tag{11}$$

The first line of (11) corresponds to the silo case, allowing for an other-division bonus 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 wage cost, net of what it would be in the corresponding contingency in the silo case. The third line of (11) 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_L in division B, because of the assumption that the new M_A is paid according to the contract terms $(\alpha_n, \beta_n, \gamma_n)$ since *her* worker is not transferred.

The manager's incentive constraint for execution effort is the same as for silos:

$$\beta_n, \beta_t \geq \frac{\xi}{t_{bb}} = \beta^x. \tag{12}$$

M_A 's training incentive constraint can be derived from her expected payoff at stage 2 of the game, as function of training effort e and assuming that in equilibrium, $p_B = p_h$ and $x_A = x_B = 1$:

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

The first line of (13) is the same as (9), with an additional possible bonus $t_S\gamma_n$ based on division 2's output. The second line states the probability of a promotion of W_A to M_B , 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 $\delta\phi t_{gb} + (1 - \delta\phi)t_{bb}$. The last three terms in $\{\}$ subtract what M_A would be paid if her worker were not transferred, to correct for terms already included in the first line of (13). The resulting training incentive constraint is

$$\begin{aligned} \text{(TIC)} \quad \psi \equiv \sigma \Delta p \{ & \kappa_w \Delta q \beta_n + v[\alpha_t - \alpha_n + t_{\mu w} \beta_t - t_{\mu g} \beta_n] \\ & + (\delta\phi t_{gb} + (1 - \delta\phi)t_{bb})\gamma_t - t_{ob}\gamma_n \} \geq \tau. \end{aligned} \quad \text{for } \mu \in \{g, b\} \quad (14)$$

Condition (TIC) is by symmetry the same for manager M_B , and thus describes the condition for $e_A = e_B = 1$ to be a Nash equilibrium of stage 2 of the game. In contrast to the silo case, (TIC) depends on the manager's own type because of the terms $t_{\mu w} \beta_t - t_{\mu g} \beta_n$ in ψ . Specifically, if $\beta_n > \beta_t$, ψ is smaller, and thus (TIC) more restrictive, for $\mu = g$. Intuitively, 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 $\beta_t > \beta_n$, (TIC) is more restrictive for $\mu = b$.

5.2 A benchmark result with costless execution effort, observable worker types and transfer-contingent contracts

Suppose 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 7). If that is the case, division managers are no longer privately informed about their workers, but the firm must still give the managers incentives to train their workers.

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

$$\min_{\zeta} w_A^L(\zeta) \text{ s.t. TIC}(\zeta) \text{ for } \mu \in \{g, b\}, (12), \text{ and } \alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t \geq 0. \quad (15)$$

A useful benchmark result holds if execution effort is irrelevant:

Proposition 2 *Suppose $\xi = 0$. Then with a lattice and observable worker types, an optimal contract that induces $e_A = e_B = x_A = x_B = 1$ is given by*

$$\alpha_t = \frac{\tau}{\sigma v \Delta p} \quad (16)$$

and $\alpha_n = \beta_n = \beta_t = \gamma_n = \gamma_t = 0$. The firm's expected wage bill is smaller than with silos, and therefore a lattice dominates silos.

Proposition 2 states that the optimal contract that incentivizes training consists of a simple “referral bonus” $\alpha_t > 0$ 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.¹⁷ The intuition is that in contrast to noisy outputs, the transfer of a good W_A amounts to a noiseless signal about the worker's good type. Because both the firm and the managers are risk neutral, it is therefore optimal to put all incentives on this event, irrespective how rare it is (α_t is inversely proportional to the probability of a transfer). Note that even though the worker's type is not verifiable, and the CEO cannot commit to any personnel decisions, the ability to tie wages to the event of a transfer in effect makes makes a worker's good type verifiable with some probability.

Having a noiseless performance measure eliminates any rent that needs to be paid to the manager, and thus (dramatically) lowers the wages bill. Since a lattice also leads to more productive teams and hence a higher expected output than with silos, a lattice unambiguously dominates. While the details of this result are specific to our model, the broader message is very general: if coordinating the affairs of two businesses (or divisions) creates added value, and contracts can be designed to give the managers a stake in that added value, then managerial incentives may become stronger rather than weaker. We will return to this point in Section 5.6.

5.3 Costly execution effort

In the next three subsections, we show that reversing any of the key assumptions behind Proposition 2 suffices to create a tradeoff between a lattice and silos, in the sense that

¹⁷This 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 .

a lattice leads to higher productivity but also a *higher wage cost* than with silos. First, suppose that worker types are still observable, but that execution effort is sufficiently costly:

Proposition 3 *Suppose $\xi = \bar{\xi}_x$. With a lattice and observable worker types, an optimal contract that induces $e_A = e_B = x_A = x_B = 1$ is given by $\alpha_n = 0$,*

$$\alpha_t = \frac{1 - p_h}{\sigma \Delta p} \tau, \quad (17)$$

$\beta_n = \beta_t = \beta^S$, and $\gamma_n = \gamma_t = 0$. The firm's expected wage bill is higher than with silos.

In contrast to the previous result, the own-division bonuses β_n and β_t must now be at least β^x to induce execution effort. However, when β_n and β_t equal $\beta^x = \beta^S$ (by definition of $\bar{\xi}_x$), both rewards combined would suffice to induce training 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 training effort in a lattice. An additional reward $\alpha_t > 0$ is necessary, leading to a wage bill strictly higher than with silos.

5.4 Privately informed managers

Now suppose only M_i knows the productivity of W_i . The CEO can then 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 truthtelling even though M_A does not know W_B , because 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

$$(\text{TTg}) \quad \alpha_t + \beta_t t_{\mu w} + \gamma_t [\delta \phi t_{gb} + (1 - \delta \phi) t_{bb}] \geq \alpha_n + \beta_n t_{\mu g} + \gamma_n t_{ob}. \quad (18)$$

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. Like the training incentive constraint (TIC), (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, who is more likely to produce high output in her own division and hence less willing to give up a good worker.

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

$$(TTb) \quad \alpha_n + \beta_n t_{\mu b} + \gamma_n t_{ob} \geq \alpha_t + \beta_t t_{\mu w} + \gamma_t t_{bb}. \quad (19)$$

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, while in division B both the manager and the 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.

By inspection, the contracts of Propositions 2 and 3 both violate (TTb). Intuitively, when the contract pays an own-division output bonus and a simple referral bonus, 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 to some extent emulates the use of contracts such as those derived next. The new optimization problem (denoting again $\zeta = (\alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t)$) is

$$\min_{\zeta} w_A^L(\zeta) \text{ s.t. TIC}(\zeta), \text{TTg}(\zeta), \text{and TTb}(\zeta) \text{ for } \mu \in \{g, b\}, \text{ and } \alpha_n, \beta_n, \gamma_n, \alpha_t, \beta_t, \gamma_t \geq 0. \quad (20)$$

Proposition 4 *Consider a lattice with privately informed managers and suppose $\xi = 0$. Then one of three contracts is optimal (i.e., a solution to (20)), each of which is characterized by $\alpha_n > 0$, $\alpha_t = 0$, $\gamma_n = 0$, $\gamma_t > 0$ and $\beta_n, \beta_t \in \{0, \beta^S\}$ with $\beta_n \geq \beta_t$; the expressions are stated in the proof. Regardless of which contract is optimal, if*

$$(C3) \quad v \leq \frac{\kappa_w q_b}{\kappa_m q_g},$$

then the firm's wage cost is higher than with silos.

Some basic features of the contracts described in Proposition 4 resemble those of general situations in which an agent must exert effort to produce a good outcome, but is to truthfully report a (privately observed) bad outcome too, such as in Levitt and Snyder (1997). Here, an incentive payment $\gamma_t > 0$ is offered as reward for training a good worker who is promoted to manager. Satisfying (TTb) to get a manager to report if she has a bad worker, in turn, that is, is incentivized through a lump-sum payment $\alpha_n > 0$.

As long as v is not too large, however, it is actually more efficient to satisfy (TTb) through $\beta_n > 0$, because β_n has a positive effect on training incentives whereas α_n has a negative effect. Thus, in two of the contracts mentioned in the proposition, β_n is set to the maximum possible value consistent with (TTg), which happens to be the silo wage β^S . More generally, any contract that satisfies (TIC) 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 training 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).

Moreover, 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, $\alpha_n > 0$ is still required to satisfy (TTb), which leads to a wage bill strictly higher than with silos.

Only if v is relatively large (or q_b/q_g very small), it is possible that the lattice wage

cost is smaller than with silos. Intuitively, as $q_b \rightarrow 0$, β_n ceases to prevent a bad manager from reporting a bad worker as a good one because she has nothing to lose. In this case, the optimal contract has $\beta_n = \beta_t = 0$. Training is incentivized through $\gamma_t > 0$, and truthtelling about a bad worker through $\alpha_n > 0$. In the limit, the contract converges to the one of Proposition 2, which is why the wage cost can be lower for a lattice.

Before moving on, let us report briefly what happens when instead of transfer-contingent contracts, *message-contingent* contracts are feasible, of the form $w_i = \alpha_{\hat{\theta}} + \beta_{\hat{\theta}} y_i + \gamma_{\hat{\theta}} y_j$ for reports $\hat{\theta} \in \{g, b\}$ about the worker's type (our analysis is available upon request). With silos, message-contingent contracts cannot improve over the simple contract of Proposition 1, because it is impossible to meaningfully separate a manager with a good or a bad worker, respectively, when the reports have no allocational consequences. With a lattice, separation is possible. Like in Proposition 4, both truthtelling constraints are binding if v is small enough. The wage cost is again higher than with silos, which can be shown directly but already follows from Proposition 4 because message-contingent contracts are conditioned on coarser information, and are hence strictly less powerful than transfer-contingent contracts.

5.5 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 (16)). 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 either create perverse incentives for managers, or 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.

Whatever the reasons, suppose that conditioning contracts on transfers is infeasible, and consider contracts that depend on the divisions' outputs only: $w_i = \alpha + \beta y_i + \gamma y_j$. The expression for $\pi_A^L(p_A, p_B)$ in (11) simplifies to

$$w_A^L(p_h, p_h) = \alpha + [t_S + \sigma v p_h (y_G - y_L)](\beta + \gamma), \quad (21)$$

and M_A 's training incentive constraint simplifies to

$$\psi \equiv \sigma \Delta p \{(\kappa_w \Delta q - v y_L)\beta + v y_G \gamma\} \geq \tau. \quad (22)$$

Like Proposition 4, the next result relies on a weak assumption that places an upper bound on v . Economically, it implies 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.

Proposition 5 *Consider a lattice with observable worker types, assume that only output-based contracts are feasible, and assume that*

$$(C3') \quad v \leq \frac{\kappa_w}{\kappa_m + \kappa_w}.$$

Then the optimal contract that induces $e_A = e_B = x_A = x_B = 1$ is given by $\alpha = \gamma = 0$ and

$$\beta = \frac{\tau}{\sigma \kappa_w \Delta p \Delta q [1 - v(1 - p_h)]} > \beta^S.$$

Intuitively, both β and γ provide incentives for training, but as intuition would suggest, it is less costly to incentivize training through β provided that (C3') holds, i.e., if the probability of a relevant vacancy is not too large. The expression for β equals β^S divided by the term in $[]$ -brackets, which raises β above β^S . This increase in the necessary bonus reflects the prospect of losing a good W_A to division B, while because of $\gamma = 0$, M_A does not have a stake in the associated gain for the firm.¹⁸

Proposition 5 establishes a third reason that can overturn the benchmark result of Proposition 2: even if there is no adverse-selection problem and if execution effort does not matter, wage costs are higher with a lattice than with silos if contracts cannot (sufficiently) reward managers for the benefits of a lattice.

If the managers have privately information about their workers, the truthtelling constraint (TTg) from Section 5.4 becomes

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

which can be expressed more simply as $\gamma y_G \geq \beta y_L$. This leads to a lower bound for γ as fraction of β :

$$\gamma \geq \gamma_g \beta \quad \text{for } \gamma_g = \frac{y_L}{y_G} = \frac{\kappa_w(1 - p_h)}{\kappa_m(\delta\phi - p_o)}. \quad (23)$$

Notice that $\gamma_g < 1$ per condition (C1). For a manager with a bad worker, the constraint (TTb) from Section 5.4 becomes

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

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

$$\gamma \geq \gamma_b \beta \quad \text{for } \gamma_b = \frac{t_{mw} - t_{mb}}{t_{ob} - q_b} = \frac{\kappa_w p_h}{\kappa_m p_o}, \quad (24)$$

with $\gamma_b < 1$ per condition (C2). Thus, with simple linear contracts, both truthtelling constraints can be satisfied by giving a manager a large enough stake in the other division.

¹⁸With risk-averse managers (and unlimited liability) and linear contracts, γ would be positive due to the Informativeness principle. At least for σ high enough, however, the wage bill would still be higher with a lattice than with silos because γ is paid as unconditional bonus, whereas the firm realizes the gain y_G only if a worker is transferred.

From Proposition 5, the only reason to choose $\gamma > 0$ is to induce truthtelling. Therefore, it is optimal to set $\gamma = \max\{\gamma_g, \gamma_b\}\beta$.

It is straightforward to verify that $\gamma_g \geq \gamma_b$ if and only if $p_o \geq \delta\phi p_h$, which in turn is implied by (C2). Thus, if the firm prefers an outside manager over an internally promoted worker of unknown quality, it follows that preventing managers from hoarding good workers (and forcing the firm to hire from outside) is a more restrictive constraint than preventing managers from kicking upstairs bad workers.

Proposition 6 *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 high training and execution effort is given by $\beta = \beta^S$ and $\gamma = \gamma_g\beta^S$. The firm's expected wage bill is higher than with silos.*

Since $\gamma_g \geq \gamma_b$, (TTg) is binding with an optimal contract, but (TTb) is not. The result $\beta = \beta^S$ then follows for the same reason we already encountered in Section 5.4: 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.

To conclude, if wages can depend only on outputs, it is best to incentivize training 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 *without* any direct compensation. With privately informed managers, both the no-hoarding and the no-kicking-upstairs truthtelling constraints 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.

5.6 Discussion: “selective intervention” in internal labor markets

Let us interpret our results in the context of Williamson’s (1985) “selective-intervention puzzle”. Even without any other synergies between divisions A and B, selectively transferring employees between divisions is a form of value-increasing “selective intervention”, much like other forms of resource reallocation.¹⁹ Williamson and others have argued that intervention from the top tends to undermine division managers’ incentives, either directly (as in Aghion and Tirole, 1997) or because of deliberate provision of weaker incentives. As Proposition 2 has shown, however, this argument is incomplete and relies on (possibly overly restrictive) assumptions about contracting constraints. If division managers’ incentives can be sufficiently tied to the value-increasing interventions, then incentives can in fact be strengthened, in which case integration is unambiguously optimal.

Our subsequent results illustrate why this result may not hold. The most obvious reason is that the contracting space may in fact be restricted, possibly for reasons discussed in Section 5.5. In that case, moral hazard in training alone may lead to higher costs of a lattice (Proposition 5) because managers bear the costs of losing a good worker with little or no stake in the benefit.

Costly execution effort (Proposition 3) has a similar but more subtle effect: even when transfer-contingent contracts are feasible, the requirement to incentivize own-division performance limits the firm’s ability to target incentives at the event of a worker transfer, again leading to higher costs of a lattice if execution effort is sufficiently important. Thus, additional incentive constraints can have the same effect as outright restrictions on the contracting space.

Proposition 4, finally, reinforces a key argument of Friebe and Raith (2010): when selective intervention relies on information that resides with the division managers, then the costs of establishing truthful upward communication are genuine costs of integration. Here, this result holds even when contracts can be targeted at the event of a transfer

¹⁹General Electric is a good example of a conglomerate held together significantly by synergies resulting from an actively managed corporate internal labor market; see Linebaugh (2012).

because private information introduces a *two-sided* adverse-selection problem with (so to speak) upward and downward truthtelling incentive constraints that are both binding.

A practically important implication of our results is that a lattice cannot be implemented simply by changing decision rights over personnel allocations without changing the incentive system as well: 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' training effort and create incentives for them to misrepresent information about their people. Costly changes in the incentive system are necessary to get a lattice to work at all.

6 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. All results stated in this section hold for both transfer-contingent and for simple output-based contracts.

Proposition 7 *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 4 and 6, 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 ($\kappa_m > \kappa_w$), then a higher value of output implies not just a higher value of having good employees in general, but a higher 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). 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 7 thus suggests a link between product market competition and the organization of internal labor markets. Intuitively, when increasing quality or lowering costs is a (hierarchical) team production process in which higher-level employees are more important, then greater competition increases the value of moving talented employees into higher-level positions. Practitioners, too, have suggested a link between competition and the value of searching for talent internally:

“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 depends on the external market for managers:

Proposition 8 *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. Intuitively, while expected output is increasing in p_o

for both silos and a lattice, the effect is smaller for a lattice, and hence the profit *difference* decreasing in p_o , because with a lattice there is less hiring from the external market for managers. A higher p_o also increases the incentive to hoard a good worker (tightens (TTg)), and lowers training incentives (tightens (TIC)) 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.

Proposition 8 thus establishes a link between the “war for talent”—businesses’ perception that recruiting good people externally has become harder—and the increase in firms’ attention towards “talent management”. Tighter labor markets, too, have been linked to greater product market competition; for a theoretical analysis see Gersbach and Schmutzler (2012). Propositions 7 and 8 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.²⁰

Proposition 9 *The difference in profit between a lattice with private information, and silos, $\pi^{LP} - \pi^S$, is increasing in κ_m with simple output-based contracts, and with transfer-contingent contracts when contracts (1) or (3) of Proposition 4 are optimal, and may be increasing or decreasing in κ_m if contract (2) is optimal.*

There is again a simple direct effect of κ_m on the profit difference: by assumption, it is preferable to promote good workers from another division than to hire from outside. This difference in productivity is scaled by κ_m ; thus, the more important managers are for production, the more important it is to get the best people into managerial positions, which raises the profitability of a lattice relative to silos. At the same time, a higher κ_m generally relaxes all incentive constraints by putting a larger weight on efficient managerial staffing decisions. An exception occurs when contract (2) of Proposition 4 is optimal, in

²⁰The 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 Zabochnik 2004), 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.

which case a higher κ_m can tighten (TTb) and require an increase in α_n , which can dominate the other effects.

Proposition 9 suggests that firms' recent efforts to establish greater cross-divisional mobility of employees is partly caused by a skill-biased technological change that has increased the returns to talent in higher managerial positions; see Kaplan and Rauh (2013). Scale and "superstar" effects, partly a result of skill-biased technological change, thus not only have profound effects on managerial labor markets but likely have effects on the internal organization of firms as well.

Proposition 10 *The difference in profit between a lattice with private information and silos, $\pi^{LP} - \pi^S$, is increasing in δ and ϕ .*

In words, both greater "human capital similarity" between divisions, and a greater upward transferability of skills, favor cross-divisional mobility. The net output gain (proportional to $y_G - y_L$) is obviously increasing in δ , but is increasing in ϕ too because there are more frequent promotions with a lattice. Moreover, increases in δ or ϕ relax both (TIC) and (TTg) without tightening (TTb), which results in lower wages, cf. Propositions 4 and 6.

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. Nevertheless, we obtain an interesting link between corporate strategy and internal labor markets, and a potentially testable prediction: other things equal, cross-divisional mobility is more likely in firms whose divisions are more similar.

The result for ϕ implies that greater vertical mobility of employees should also tend to favor cross-divisional mobility. 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.

7 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' training 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.

Direct monitoring of training effort: All solutions discussed so far provide incentives for training that are based on outcomes. Some firms choose to measure and reward training effort directly, such as instrument maker Agilent, whose top managers receive a third of their compensation based on HR development efforts (Conaty and Charan 2010, page 160). One way in which to measure training effort is through 360-degree reviews, in which subordinates are asked about their boss's training 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)²¹. Most strikingly, these forms

²¹We 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.

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 training.

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. above. The primary objective of job rotation is usually to develop junior employees’ skills and to prepare them for higher-level positions (e.g. Conaty and Charan 2010, 228-229). One might think that job rotation also solves several problems emphasized in this paper: employees can develop their skills through the rotation program rather than having to rely on training by their boss; bosses can no longer “hoard” good people; and talented employees become more visible to senior managers in different units, reducing an adverse-selection problem.

Only the last point—greater visibility—is a true advantage, however. Job rotation does not diminish the importance of training by managers; it only spreads the training responsibility across multiple managers. While alleviating the adverse-selection problem, it makes the moral-hazard problem worse: managers have even less incentive to train someone who will leave their unit not with some probability but for sure! Managers may no longer be able to hoard good people, but hoarding is only a symptom of a problem; the cause is managers’ concern for getting an adequate return on their training efforts.

Another downside of job rotation, though not captured in our model, is that while employees acquire general managerial skills, they may lack the time and motivation to acquire division- (or industry-)specific expertise.²² After many years of emphasizing general managerial skills, General Electric recently decided to keep senior managers in their divisions longer to help them acquire the expertise needed to compete in their

²²For a classical description about the effects of cross-firm rotation of the French engineering elite from the Ecole Polytechnique, see Crozier (1964).

industries (Linebaugh 2012). To conclude this point, job rotation has its advantages but its costs as well, and is therefore a good solution for some firms but not others.

Other leadership development initiatives: Other initiatives, too, simultaneously foster human capital development and generate firm-wide information about talented employees. General Electric’s Management Development Institute in Crotonville, NY, 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 and Schmidt 2010). Likewise, Procter & Gamble fosters the development of social networks to make sure that managers can rely on their colleagues’ help 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 private information about top talent may be only small. Efforts for everyone to get to know everyone else are costly, however, and can realistically cover only the top tier of managers, perhaps 100-200 people. It follows that even though private information may not be an issue at the very top of a company, moving down the ranks it will eventually begin to matter (recall our remark in Section 3 that the tiers in our model may represent any adjacent tiers in a larger hierarchy).

In conclusion, job rotation and other leadership development initiatives have clear benefits but have disadvantages or costs 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 trains and holders of private information.

8 Conclusion

Traditional vertical job ladders in firms—the subject of a large economics literature—have recently been giving way to active “talent management” aimed at optimally matching people with positions, which includes efforts to promote cross-divisional mobility of employees. We have argued in this paper that such efforts create agency costs for firms because managers play a key role as gatekeepers in internal labor markets. Because managers prefer to have good employees working for them, efforts to increase employees’

cross-divisional mobility undermine managers' incentives to invest in training, and create incentives to use private information about employees strategically, either by "hoarding" good employees or by "kicking upstairs" bad ones. Our model captures the contractual origins of these agency problems: team production in firms, and an inability to measure training 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 prevalence 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 train 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 or labor 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 training 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 observe a decline in managerial training (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 executives found it tough to move up when young presidents stood in the way, and tougher still to jump over to a separate company [within J&J]." 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 motivate them 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 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 (Conaty and Charan, 2010). The second departure 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 their firm’s. This approach, we hope, advances the agenda proposed by Gibbons (2013) for organizational economics to “focus on what managers actually do,” in the spirit of Cyert and March’s (1963) observation that “managers devote much more time and energy to the problems of managing their coalition than they do to the problems of dealing with the outside world” (p. 205-6).

Appendix: Proofs

Proof of Proposition 1: As noted in the text, (9) leads to the training incentive constraint $\sigma t_\mu(p_h)\beta - \tau \geq \sigma t_\mu(p_l)\beta$ or

$$\sigma \kappa_w \Delta p \Delta q \beta \geq \tau \tag{25}$$

which leads to the stated expression for β^S . If $\beta^S \geq \beta^x$, then β^S satisfies both (25) and (8), and thus the optimal contract is $(\alpha, \beta, \gamma) = (0, \beta^S, 0)$. QED

Proof of Proposition 2: We will derive an optimal contract for the general case $\xi \in [0, \bar{\xi}]$. The proof proceeds in several steps.

1. First, we will show that for the left-hand side of M_A ’s training incentive constraint

(TIC), ψ , and the firm's wage cost w_A^L , we have

$$\frac{\partial\psi/\partial\beta_n}{\partial w_A^L/\partial\beta_n} < \frac{\partial\psi/\partial\alpha_t}{\partial w_A^L/\partial\alpha_t} \quad \text{and} \quad \frac{\partial\psi/\partial\beta}{\partial w_A^L/\partial\beta} < \frac{\partial\psi/\partial\alpha_t}{\partial w_A^L/\partial\alpha_t}, \quad (26)$$

where the derivatives in the second inequality are evaluated for $\beta_n = \beta_t \equiv \beta$. From (14) and (11), the right-hand side of (26) is easily computed as $\Delta p/p_h$. For the left-hand side of the first inequality, we have

$$\frac{\partial\psi}{\partial\beta_n} = \sigma\Delta p(\kappa_w\Delta q - vt_{\mu g}) < \sigma\Delta p(\kappa_w\Delta q - v\kappa_w\Delta q) = \sigma\Delta p(1-v)\kappa_w\Delta q,$$

where the inequality follows because $t_{\mu g} \geq t_{bg} = (\kappa_m + \kappa_w)q_b + \kappa_w\Delta q$; and

$$\frac{\partial w_A^L}{\partial\beta_n} = t_S + \sigma vp_h(y_G - t_{mg}) > \sigma t_0 - \sigma vp_h t_{mg} = \sigma(1-v)t_0 + \sigma v(1-p_h)t_{mb} > \sigma(1-v)t_0.$$

Overall, therefore, we have

$$\frac{\partial\psi/\partial\beta_n}{\partial w_A^L/\partial\beta_n} < \frac{\sigma\Delta p(1-v)\kappa_w\Delta q}{\sigma(1-v)t_0} = \Delta p \frac{\kappa_w\Delta q}{t_0} < \frac{\Delta p}{p_h},$$

where the last inequality follows from $t_0 > p_h t_{mg} = p_h(\kappa_m q_m + \kappa_w q_g) > p_h \kappa_w \Delta q$. This proves the first inequality in (26). For the second inequality, we have

$$\frac{\partial\psi}{\partial\beta} = \sigma\Delta p[\kappa_w\Delta q - v(t_{\mu g} - t_{\mu w})] = \sigma\Delta p[1 - v(1 - p_h)]\kappa_w\Delta q,$$

and

$$\begin{aligned} \frac{\partial w_A^L}{\partial\beta} &= t_S + \sigma vp_h(y_G - t_{mg} + t_{mw}) > \sigma t_0 - \sigma vp_h(t_{mg} - t_{mw}) \\ &= \sigma[1 - v(1 - p_h)]t_0 + \sigma v(1 - p_h)(t_0 - p_h \kappa_w \Delta q) > \sigma[1 - v(1 - p_h)]t_0, \end{aligned}$$

and the final step is the same as above.

2. Constraint (12) restricts attention to contracts with $\beta_n, \beta_t \geq \beta^x$. Next, α_n and γ_n have no effect on execution effort incentives and a negative effect on training incentives, cf. ψ in (14). Therefore, set $\alpha_n = \gamma_n = 0$. Next, (13) and (14) lead to

$$\frac{\partial w^L/\partial\gamma_t}{\partial w^L/\partial\alpha_t} = \frac{\partial\psi/\partial\gamma_t}{\partial\psi/\partial\alpha_t} = \delta\phi t_{gb} + (1 - \delta\phi)t_{bb}.$$

It follows that training 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).

3. Next, we show that an optimal contract must satisfy $\beta_t = \beta_n$. To see this, observe that

$$\frac{\partial w^L / \partial \beta_t}{\partial w^L / \partial \alpha_t} = t_{mw} \quad \text{and} \quad \frac{\partial \psi / \partial \beta_t}{\partial \psi / \partial \alpha_t} = t_{\mu w} \quad (27)$$

If $\beta_t \geq \beta_n$, in which case (14) is (weakly) more restrictive for $\mu = b$, then (27) implies that it is less costly to provide training incentives through α_t than through β_t ; hence an optimal contract must have β_t at the lowest possible level, which is $\beta_t = \beta_n$ in order to satisfy both $\beta_t \geq \beta_n$ and (12). Conversely, if $\beta_t < \beta_n$, (14) is (strictly) more restrictive for a $\mu = m$, and (27) implies that is optimal to incentivize training through β_t rather than α_t , while per step 1, providing incentives through α_t is less costly than providing incentives through β_n . It follows that it is less costly to incentivize training through β_t than through β_n , in which case a contract with $\beta_t < \beta_n$ cannot be optimal. Overall, it follows that $\beta_t = \beta_n$.

4. According to step 1, it is less costly to incentivize training 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 (14) given the values of the other contract variables. The solution is

$$\alpha_t = \frac{\tau}{\sigma v \Delta p} - \frac{1 - v(1 - p_h)}{v q_b} \frac{\kappa_w}{\kappa_m + \kappa_w} \Delta q \xi. \quad (28)$$

The optimal contract is not unique: a different, cost-equivalent contract is one with $\alpha_t = 0$ and γ_t chosen to satisfy (14). Any convex combination of these two corner solutions is optimal too.

5. For silos, the expected wage cost is $t_S \beta$, which using $\beta = \beta^S$ from Proposition 1 equals

$$\frac{t_S}{\sigma \kappa_w \Delta p \Delta q} \tau > \frac{\sigma t_0}{\sigma \kappa_w \Delta p \Delta q} \tau = \frac{\kappa_m q_m + \kappa_w (q_b + p_h \Delta q)}{\kappa_w \Delta p \Delta q} \tau > \frac{p_h}{p_h - p_l} \tau. \quad (29)$$

For a lattice, the wage cost for the case $\xi = 0$ is obtained from (11) for $\alpha_n = \beta_n = \beta_t = \gamma_n = \gamma_t = 0$, and is simply $\sigma v p_h \alpha_t = (p_h / \Delta p) \tau$, which in conjunction with (29) proves that the wage cost is lower with a lattice. QED

Proof of Proposition 3: If $\xi = \bar{\xi}_x$, then by definition $\beta_n = \beta_t = \beta^S$. Solving (28) for $\xi = \bar{\xi} = t_{bb} \beta^S$ leads to $\alpha_t = (1 - p_h) \tau / (\sigma \Delta p)$, as stated. (For larger values of ξ , 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 training and execution effort.) To evaluate the wage cost, observe in (11) 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 second and third terms in (11) are both strictly positive, because $\alpha_t > 0$ and the β -terms reduce to $\sigma v p_h (y_G - y_L) \beta > 0$. It follows that the expected wage cost with a lattice is strictly higher than with silos. 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. For any optimal contract, (TTb) must be binding. Solving the problem without it would lead to the contract of Proposition 2 (or a profit-equivalent contract with $\gamma_t > 0$; see step 2 of the proof of Proposition 2). But any such contract, with $\beta_n = \beta_t = \beta^x$, and $\alpha_n = \gamma_n = 0$, would violate (TTb).

2. 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

$$\frac{\partial w_A}{\partial \alpha_t} = \sigma v p_h \text{ and } \frac{\partial w_A}{\partial \gamma_t} = \sigma v p_h [\delta \phi t_{gb} + (1 - \delta \phi) t_{bb}],$$

the contract $\zeta' = (\alpha_n, \beta_n, \gamma_n, \alpha_t - \Delta, \beta_t, \gamma_t + \frac{\Delta}{\delta \phi t_{gb} + (1 - \delta \phi) t_{bb}})$ leads to the same wage cost; similarly, it can be shown that both (TIC) and (TTg) hold for ζ' iff they hold for ζ . The same substitution, however, relaxes (TTb) by reducing the right-hand side by Δ times $1 - \frac{t_{bb}}{\delta \phi t_{gb} + (1 - \delta \phi) t_{bb}}$, which is positive. All constraints are thus satisfied, and it follows that if ζ is optimal, then ζ' must be optimal too. But because of step 1, ζ' cannot be optimal because (TTb) is slack. It follows that ζ cannot be optimal either, proving by contradiction that $\alpha_t = 0$.

3. 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 (TIC) and (TTg) is for $\mu = b$, whereas the most restrictive version of (TTb) is for $\mu = g$. In this case,

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

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

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

Finally, ζ' leads to a decrease in the wage cost by $\sigma v p_h(t_{mw} - t_{bw})$. 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 (TIC) and (TTg) is for $\mu = g$, whereas the relevant version of (TTb) is for $\mu = b$.

4. Next, if ζ is optimal, then $\gamma_n = 0$. Suppose instead that γ_n is positive. Then the contract $\zeta' = (\alpha_n + \Delta t_{ob}, \beta_n, \gamma_n - \Delta, \alpha_t, \beta_t, \gamma_t)$ would continue to satisfy (TTb) for $\mu = b$, and (TIC) and (TTg) for $\mu = g$. The wage cost, in turn, changes by Δ times

$$-\sigma[t_0 - t_{ob}] - (1 - \sigma)[t_r - t_{ob} - \sigma p_h(1 - p_h)y_L]. \quad (30)$$

The first term in (30) is negative because $t_0(p_h) > t_{mb} > t_{ob}$, where the first inequality follows from (6) and the second from the assumption $p_m \geq p_o$. For the second term in (30), we have

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

and from (7) we have

$$t_r - t_{ob} = p_h[\phi t_{gb} + (1 - \phi)t_{bb} - t_{ob}] = p_h \kappa_m(\phi - p_o)\Delta q,$$

which implies that the second term in (30) is positive too. It follows that ζ' is feasible and dominates ζ . Therefore ζ with $\gamma_n > 0$ cannot be optimal.

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

$$\begin{aligned} \alpha_n &= \frac{t_{bb}}{\sigma v \delta \phi \kappa_m \Delta p \Delta q} \tau - \left[\frac{(1 - v)\kappa_w - v\kappa_m}{v \delta \phi \kappa_m} + 1 \right] t_{bb} \beta_n \\ &\quad - \left[\frac{1 - \delta \phi}{\delta \phi} t_{bb} - \kappa_w p_h \Delta q \right] \beta_t \text{ and} \\ \gamma_t &= \frac{1}{\sigma v \delta \phi \kappa_m \Delta p \Delta q} \tau - \frac{(1 - v)\kappa_w - v\kappa_m}{v \delta \phi \kappa_m} \beta_n - \frac{1}{\delta \phi} \beta_t. \end{aligned} \quad (31)$$

Any optimal contract must satisfy (31).

6. Restricting attention to corner solutions, it is optimal to choose β_n either as small or as large as possible, subject to constraints. A lower bound to β_n is given by β^x ; an upper

bound is given by (TTg). If (TTg) is binding, then $\beta_n = \beta^S$: Substituting the right-hand side of (TTg) for $\alpha_t + \beta_t t_{\mu w} + \gamma_t [\delta \phi t_{gb} + (1 - \delta \phi) t_{bb}]$ in (14) cancels the entire $[-]$ -term in (14) and leaves the same effort incentive constraint as for the silo case. Therefore, $\beta_n = \beta^S$. For any optimal contract, we therefore have $\beta_n \in \{\beta^x, \beta^S\}$, which in combination with $\beta_n \geq \beta_t$ leads to exactly three contracts: (1) $\beta_n = \beta_t = \beta^S$, (2) $\beta_n = \beta^S, \beta_t = \beta^x$, and (3) $\beta_n = \beta_t = \beta^x$, with α_n and γ_t given by (31) and $\alpha_t = \gamma_n = 0$. Each can be optimal for different parameter values. For the first two contracts and the case $\xi = 0$ and hence $\beta^x = 0$, the expressions for α_n and γ_t obtained from (31) are as follows:

$$\begin{aligned} (1) \beta_n = \beta_t = \beta^S: \quad & \alpha_n = \frac{\kappa_m(q_b + \delta \phi p_h \Delta q) + \kappa_w q_b}{\sigma \delta \varphi \kappa_m \kappa_w \Delta p \Delta q} \tau, \quad \gamma_t = \frac{1}{\sigma \delta \varphi \kappa_m \Delta p \Delta q} \tau \\ (2) \beta_n = \beta^S, \beta_t = 0: \quad & \alpha_n = \frac{(1 - \delta \phi) \kappa_m + \kappa_w}{\sigma \delta \varphi \kappa_m \kappa_w \Delta p \Delta q} t_{bb} \tau, \quad \gamma_t = \frac{\kappa_m + \kappa_w}{\sigma \delta \varphi \kappa_m \kappa_w \Delta p \Delta q} \tau. \end{aligned}$$

7. It remains to evaluate the wage cost. From (11), using $\gamma_n = \alpha_t = 0$, the expected wage cost with a lattice is

$$\alpha_n + t_S \beta_n + \sigma v p_h \{t_{mw} \beta_t + [\delta \phi t_{gb} + (1 - \delta \phi) t_{bb}] \gamma_t - \alpha_n - t_{mg} \beta_n + y_G \beta_n\}. \quad (32)$$

For contracts (1) and (2) of step 6, (TTg) is binding and $\beta_n = \beta^S$. In this case, the term $t_S \beta_n$ in (32) equals the wage cost with silos. With a lattice, therefore, the wage cost differs by the expression

$$\Delta w = \alpha_n + \sigma v p_h \{t_{mw} \beta_t + [\delta \phi t_{gb} + (1 - \delta \phi) t_{bb}] \gamma_t - \alpha_n - t_{mg} \beta_n + y_G \beta_n\}.$$

From (TTg) for $\mu = g$, we have

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

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

$$\Delta w = \alpha_n + \sigma v p_h [(1 - p_m)(\beta_n - \beta_t) + (\delta \phi - p_o) \beta_n] \kappa_m \Delta q,$$

which is positive because $\beta_n \geq \beta_t$ from step 3, and $\delta \phi > p_o$ from (C1). Since $\alpha_n > 0$, it follows that $\Delta w > 0$.

For contract (3), wage costs are lowest if $\xi = 0$ and hence $\beta_n = \beta_t = 0$. In this case, the wage cost difference compared with silos is

$$\begin{aligned} \Delta w &= (1 - \sigma v p_h) \alpha_n + \sigma v p_h [\delta \phi t_{gb} + (1 - \delta \phi) t_{bb}] \gamma_t - \beta^S t_S \\ &= [t_{bb} + \sigma v p_h \delta \phi (t_{gb} - t_{bb})] \gamma_t - \beta^S t_S, \end{aligned} \quad (33)$$

where the second equality follows from $\alpha_n = t_{bb}\gamma_t$ according to (31). Because $\delta\phi$ appears in the denominator of γ_t , the first term in (33) is decreasing in $\delta\phi$. So we have (using (31) and (10))

$$[t_{bb} + \sigma v p_h \delta\phi(t_{gb} - t_{bb})]\gamma_t \geq [t_{bb} + \sigma v p_h(t_{gb} - t_{bb})]\frac{1}{\sigma v \kappa_m \Delta p \Delta q} \tau \geq t_{bb} \frac{\kappa_w}{v \kappa_m} \beta^S,$$

whereas t_S in the second term of (33) cannot exceed t_{gg} . It follows that a sufficient condition for $\Delta w > 0$ is

$$t_{bb} \frac{\kappa_w}{v \kappa_m} > t_{gg},$$

which is equivalent to condition (C3) stated in the proposition. QED

Proof of Proposition 5: If $\xi = 0$, there is only the training incentive constraint $\psi > 0$ to satisfy, and an optimal corner solution will have either $\beta > 0$ and $\gamma = 0$, or $\gamma > 0$ and $\beta = 0$. The first case obtains if and only if

$$\frac{\partial\psi/\partial\beta}{\partial w_A/\partial\beta} > \frac{\partial\psi/\partial\gamma}{\partial w_A/\partial\gamma}. \quad (34)$$

Since $\partial w_A/\partial\beta = \partial w_A/\partial\gamma$ (cf. 21), (34) holds if and only if $\partial\psi/\partial\beta > \partial\psi/\partial\gamma$, or

$$\kappa_w \Delta q - v y_L > v y_G.$$

The left-hand side is $\kappa_w \Delta q - v y_L = \kappa_w [1 - v(1 - p_h)] \Delta q > \kappa_w (1 - v) \Delta q$, and the right hand side is $v y_G = v \kappa_m (\delta\phi - p_o) \Delta q < v \kappa_m (\delta\phi - p_o) \Delta q$. Under condition (C3'), we have $(1 - v) \kappa_w \geq v \kappa_m$, and therefore $\partial\psi/\partial\beta > \partial\psi/\partial\gamma$. It follows that training is least costly to incentivize with $\beta > 0$ and $\gamma = 0$, and solving (22) for this case leads to the expression for β stated in the proposition. QED

Proof of Proposition 6: As argued in the text, under (C3'), the cost-minimizing value of γ that satisfies both (TTg) and (TTb) is $\gamma = \min\{\gamma_g, \gamma_b\} = \gamma_g$ because of (C2). In this case, $\gamma = \gamma_g \beta$ as stated. Substituting this expression for γ in (22) and solving for β leads to $\beta = \beta^S$. QED

Proof of Proposition 8: For transfer-contingent contracts, according to Proposition 4, $\beta_n = \beta^S$ and $\gamma_n = 0$, like in the silo case. The profit difference therefore equals (see (11))

$$\Delta\pi = -\alpha_n + \sigma v p_h \{ (R - \beta_t) t_{mw} - [\delta\phi t_{gb} + (1 - \delta\phi) t_{bb}] \gamma_t + \alpha_n - (R - \beta_n) t_{mg} + (R - \beta_n) y_G \} \quad (35)$$

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 output-based contracts, using $\pi^S = (R - \beta^S)t_S$, (21), and $\beta = \beta^S$ according to Proposition 6, we have

$$\Delta\pi = [t_S + \sigma v p_h(y_G - y_L)](R - \beta - \gamma) - t_S(R - \beta) \quad (36)$$

$$= \sigma v p_h(y_G - y_L)(R - \beta) - [t_S + \sigma v p_h(y_G - y_L)]\gamma \quad (37)$$

$$= \sigma v p_h(y_G - y_L)(R - \beta - \gamma) - t_S\gamma, \quad (38)$$

where the first term in (37) 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 $\frac{\partial}{\partial p_o}[t_S + \sigma v p_h(y_G - y_L)] = v\kappa_m(1 - \sigma p_h)\Delta q > 0$. QED

Proof of Proposition 9: For transfer-contingent contracts, there are four effects of κ_m on $\Delta\pi$ in (35): first, in the expression in $\{\}$, the last term $(R - \beta_n)y_G$ is increasing in κ_m via y_G . Second, the derivative of $(R - \beta_t)t_{mw} - (R - \beta_n)t_{mg}$ with respect to κ_m is $(\beta_n - \beta_t)q_m \geq 0$ in all three versions of the optimal contract according to Proposition 4. Third, for $x \in \{\kappa_w, \kappa_m + \kappa_w, \kappa_w/v\}$ depending on the version of the optimal contract, we have

$$-[\delta\phi t_{gb} + (1 - \delta\phi)t_{bb}]\gamma_t = -[t_{bb} + \delta\phi\kappa_m\Delta q]\frac{x}{\delta\phi\kappa_m}\beta^S = -\left[\frac{t_{bb}x}{\delta\phi\kappa_m} + x\Delta q\right]\beta^S, \quad (39)$$

which is increasing in κ_m . The fourth effect is on $-(1 - \sigma v p_h)\alpha_n$ which is positive for contracts (1) and (3) because $\partial\alpha_n/\partial\kappa_m < 0$. For contract (2), however, $\partial\alpha_n/\partial\kappa_m < 0$ holds only if $\kappa_m^2(1 - \delta\phi) < \kappa_w^2$, which may or may not hold. If $\partial\alpha_n/\partial\kappa_m > 0$, then $\partial\Delta\pi/\partial\kappa_m < 0$ is possible.

For output-based contracts, the term $\sigma v p_h(y_G - y_L)(R - \beta - \gamma)$ in (38) is increasing in κ_m because y_G is increasing and γ is decreasing in κ_m , and all other variables are unaffected. The derivative of the remaining part, $-t_S\gamma$, is positive because t_S has the form $x\kappa_m + y\kappa_w$, while γ has the form z/κ_m . Therefore $t_S\gamma = (xz + yz\kappa_w/\kappa_m)$ is decreasing in κ_m . QED

Proof of Proposition 10: For transfer-contingent contracts, there are three effects of δ and ϕ on the profit difference, cf. (35): the first is a positive effect equal to $\sigma v p_h(R -$

$\beta_n)\partial y_G/\partial(\delta\phi) = \sigma v p_h(R - \beta_n)(t_{gb} - q_b)$. The second is a profit increase due to a decrease in α_n in all three versions of the optimal contract. The third is through the term

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

for $x \in \{\kappa_w, \kappa_m + \kappa_w, \kappa_w/v\}$ depending on the version of the optimal contract. Either way, this term is increasing in δ and ϕ .

For output-based contracts, the first term in (38) is increasing in δ and ϕ via y_G , provided that a lattice is profitable at all ($R - \beta - \gamma > 0$). The second term, $-t_S\gamma$, is clearly increasing in δ because t_S does not depend on it while γ is decreasing in δ through the effect on γ_g . For ϕ , it takes the form $-t_S\gamma = -(a+b\phi)c/(d\phi-e)$ for positive constants a through e , and the derivative is $c(ad+be)/(e-d\phi)^2 > 0$. QED

9 References

- Aghion, Philippe, and Jean Tirole (1997): “Formal and Real Authority,” *Journal of Political Economy* 105 (1): 1-29
- 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.
- Athey, Susan and John Roberts (2001), “Organizational Design: Decision Rights and Incentive Contracts”, *American Economic Review Papers and Proceedings* 91(2): 200-205
- Baker, George P. (2002): “Distortion and Risk in Optimal Incentive Contracts”, *Journal of Human Resources*, 37 (4): 728-751
- Baker, George P., Michael Gibbs, and Bengt Holmström (1994)

- 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 <https://urldefense.proofpoint.com/v1/url?u=http://partnerships.com/leftbrain/docs/GC%20Review%20July%20Aug%2008.pdf>, 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
- Cappelli, Peter (2008): “Talent management for the twenty-first century” *Harvard Business Review* 86 (3) p74
- Cappelli, Peter (2011): “Why Companies Aren’t Getting the Employees They Need”, *Wall Street Journal*, October 26
- Card, David, and John DiNardo (2002): “Skill-biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles”, *Journal of Labor Economics* 20,
- Carmichael, Lorne H. (1985), “Incentives in Academics: Why Is There Tenure?”, *Journal of Political Economy*, 96 (3): 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 <https://urldefense.proofpoint.com/v1/url?u=http://partnerships.com/leftbrain/docs/GC%20Review%20July%20Aug%2008.pdf>

- Cleaver, Joanne (2012): *The Career Lattice: Combat Brain Drain, Improve Company Culture, and Attract Top Talent*. McGraw-Hill.
- Collings, David G., and Kamel Mellahi (2009): “Strategic talent management: A review and research agenda.” *Human Resource Management Review* 19.4: 304-313.
- Conaty, Bill and Ram Charan, *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): 1058-1082
- 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, Richard M., and James March (1963): *A behavioral theory of the firm*. University of Illinois at Urbana-Champaign’s Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.
- Demougin, Dominique and Aloysius Siow (1994), “Careers in Ongoing Hierarchies”, *The American Economic Review*, 84 (5): 1261-1277
- Dessein, Wouter (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): 683-690
- 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
- Gabaix, Xavier, and Augustin Landier (2008): “Why Has CEO Pay Increased So Much?” *Quarterly Journal of Economics* (2008) 123 (1): 49-100.
- Garicano, Luis (2000): “Hierarchies and the Organization of Knowledge in Production”, *Journal of Political Economy* Vol 108 (5)
- Gennaioli, Nicola; Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer (2013): “Human Capital and Regional Development”, *Quarterly Journal of Economics* 128, 105-164.
- Gersbach, Hans and Armin Schmutzler (2012): “Does Globalization Create Superstars?” Working paper, University of Zurich.
- Gibbons, Robert (2013): “Cyert and March (1963) at Fifty: A Perspective from Organizational Economics”, prepared for NBER Working Group in Organizational Economics
- 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 <https://urldefense.proofpoint.com/v1/url?u=http://www.hrworld.com/bonus-jackpot-031108/5>, 2012.
- Kaplan, Steven N., and Joshua Rauh (2013): “It’s the Market: The Broad-Based Rise in the Return to Top Talent.” *Journal of Economic Perspectives*, 27(3): 35-56.
- 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
- Martin, Jean, and Conrad Schmidt. “How to keep your top talent.” *Harvard Business Review* 88.5 (2010): 54-61.
- Martin, Stephen (2009): “Microfoundations for the Linear Demand Product Differentiation Model, with Applications”, working paper, Purdue University.
- Mellahi, Kamel, and David G. Collings (2010): “The barriers to effective global talent management: The example of corporate élites in MNEs”, *Journal of World Business* 45 (2): 143–149
- 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 Zabojsnik (2004): “CEO Pay and Appointments: A Market-Based Explanation for Recent Trends”, *American Economic Review*, 94 (2), Papers and Proceedings: 192-196
- 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*, 19 (2): 262-273
- Peter, Lawrence J., and Raymond Hull (1969), “ *The Peter Principle: Why Things Always Go Wrong*” (Morrow)
- 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 Measurement”, *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, DA and JA Conger (2007): “Make Your Company a Talent Factory”, *Harvard Business Review* 85(6):68-77
- Rosen, Evan (2010) “Smashing Silos”, *Business Week* February 10, 2010

- Rosen, Sherwin (1981), “The Economics of Superstars”, *American Economic Review* Vol. 71, No. 5, pp. 845-858
- Rosen, Sherwin (1982), “Authority, Control, and the Distribution of Earnings”, *Bell Journal of Economics*, 13 (2): 311-323
- Scandura, Terri A. (1992): “Mentorship and Career Mobility: An Empirical Investigation”, *Journal of Organizational Behavior* 13(2): 169–174
- 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* 112 (6), pp. 1181-1222
- Vives, Xavier (2008): “Innovation and Competitive Pressure”, *Journal of Industrial Economics* 56 (3): 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*, 25 (3): 245-261
- WorldatWork (2011): “Bonus Programs and Practices”, retrieved from <https://urldefense.proofpoint.com>