

POWER IN THE FIRM AND MANAGERIAL CAREER CONCERNS



JAIME ORTEGA

Universidad Carlos III de Madrid
28903 Getafe (Madrid), Spain

More powerful managers make more important decisions. Therefore, firm performance is more informative about the abilities of such managers, who, realizing that they are more visible, are more eager to improve performance. If this reputation effect exists, how should firms allocate power? I analyze the optimal allocation of power and derive implications for several issues that often arise in management practice: the choice of departmentation criteria, the importance given to seniority, and the width of job definitions. Finally, I show that the model is consistent with the empirical evidence on managerial succession.

1. INTRODUCTION

Powerful managers make decisions that have a potentially large impact on firm performance—design and implementation of large investment projects, launching of new products, or management of new processes—and less powerful managers make fewer or less important decisions. A consequence of this simple fact is that managers with more responsibilities tend to be more motivated. Another consequence is that powerful managers are more *visible*: any action they take is more easily observed by outsiders. Some CEOs quickly reach the headlines of the business press or, just as quickly, disappear from them, but less powerful managers are rarely in the spotlight. In this article, I relate these two observations: I argue that, being more visible, more powerful managers realize that they are more accountable, and will try to increase firm value in order to build a good reputation. The article is based on the premise that power is

This paper is based on Chapter 2 of my Ph.D. dissertation written at MIT (Ortega, 1998). I am grateful to Daron Acemoglu, Susan Athey, Bengt Holmstrom, a coeditor, and two anonymous referees for their useful comments. I would also like to thank Alberto Abadie, Sandro Brusco, Marco Celentani, and Adolfo de Motta for their help. This research has been partially funded by the Spanish Ministry of Science and Technology research grant SEC2000-0395 and the Comunidad de Madrid research grant 06/0065/2000.

a limited resource: when a firm wants to reallocate power, it has to take some power away from one manager and give it to another. Therefore, the firm has to trade off the increased incentives of one of the managers against the reduced incentives of the other. I argue that this trade-off depends on the managers' career concerns: when a firm divides power among several managers, it should take into account that power makes managers more visible to the market, and that more visible managers make decisions that are better aligned with shareholders' interests. I show this with a model of team production and career concerns based on Holmstrom (1982a,b) and Jeon (1996). In this model, managers are endowed with (general and specific) human capital, the value of which cannot be perfectly evaluated by their current employer or by the labor market, but can be estimated by using information on firm performance. The managers' choices are influenced by their reputational concerns.

The idea that power motivates managers is not new, but the role of reputation has received little attention.¹ For example, in the incomplete-contracting literature (Grossman and Hart, 1986; Hart and Moore, 1990; Rajan and Zingales, 1998), the role of power is to motivate managers to undertake specific investments that are not contractible. Since the investments are specific, managers are not moved by reputational concerns. In fact, if the investments were not specific, managers would undertake them even if they had very few rights of control, i.e., very little power. Another interesting point of comparison is Aghion and Tirole's (1997) model of formal and real authority, which is also based on the idea that power makes managers undertake uncontractible actions. In their article, powerful managers have better incentives to collect information because, with more information, they can more easily identify the decisions that maximize their private benefits. If there is some congruence between a manager's and his firm's objectives, the firm itself ends up benefiting from the information that the manager collects; but this comes at a cost: when the decisions preferred by the manager are different from the ones preferred by the firm, the manager does what he prefers. In contrast with Aghion and Tirole (1997), I argue that, although managers may use their power inefficiently, competition in the market for executives is strong and limits this type of behavior. In fact, because top managers may have many temptations, we should expect the labor market to make them

1. An exception is Sauer (1988), who studied how the quality of joint publications affected the earnings of professors. He found that the effect of coauthored articles on individual earnings was inversely proportional to the number of coauthors. This evidence is consistent with our theory, because as the number of coauthors increases, each author has less power over the article.

very accountable for their firm's performance. Finally, for similar reasons, the theory that I present is quite different from some theories where the role of power is to protect the managers' careers from potential threats. For example, in Carmichael (1988), tenured professors have an incentive to hire the best assistant professors because, being tenured, they do not feel threatened by them. Friebel and Raith (2001) has a similar flavor: subordinates must communicate with their immediate superior before they can communicate with anyone higher up in the hierarchy. This chain of command gives managers the assurance that their subordinates will not threaten them. While in these papers power protects managers from the (external or internal) labor market, in my paper power exposes them to the labor market.

As a benchmark, I first analyze the optimal allocation of power in a simple case where managers are *ex ante* identical and their actions are not complementary. I show that it is optimal to divide power in an uneven way, and find two reasons for it. First, if *effort* measures the alignment of managers' actions with the firm's interests, then power and effort must be complements: if a manager chooses more effort, it is profitable for the firm to give him more power, because in that way the organization will benefit more from the effort that the manager is already exerting. Furthermore, having more power, the manager realizes that he is more visible to the market and works even harder. As a result, it is profitable to give him even more power. Second, there are increasing returns to power in visibility: as a manager's power increases, his visibility increases at an increasing rate. This result is endogenously derived from the model, and it implies that, as we move away from a perfectly equal distribution of power, the increase in effort of the manager who is made more powerful is larger than the reduction of the other manager's effort. Although the model is based on different assumptions and the trade-offs are of a different nature, this result is consistent with Meyer's (1991) idea that, when the firm learns from coarse information, unequal distributions of power are efficient.

After this basic analysis, I extend the model and derive implications for management practice. First, I analyze the effect of departmentation on incentives in large multidivisional companies. I argue that the practice of organizing divisions along product lines, giving to each division power over all business functions related to its product, is consistent with the objective of making division managers more visible. Second, I argue that the relationship between power and seniority results from a trade-off between human-capital accumulation and implicit incentives. Giving more power to younger managers increases incentives, but it prevents the firm from using the human

capital accumulated by older managers. Third, I show that the model of power and career concerns is consistent with the empirical finding that higher-level jobs are less specialized than lower-level jobs. Specialization is good for human-capital accumulation, but it reduces the managers' visibility and is bad for incentives. As a manager moves up the hierarchy, the visibility effect becomes more important and the firm has an incentive to reduce the degree of specialization. Last, I derive implications for managerial succession, showing that the model is consistent with the evidence found by Parrino (1997) and Huson et al. (2001).

2. THE EFFECT OF POWER ON THE VISIBILITY OF MANAGERS

2.1 A SIMPLE MODEL OF POWER AND CAREER CONCERNS

I use a model of career concerns where managers have an incentive to improve firm performance in order to build a good reputation. I extend Holmstrom's (1982b) original model in two directions. First, I assume that firms employ two managers instead of one. I also assume that the labor market cannot observe their individual levels of performance: only firm performance is observable. A similar extension has already been used by Meyer (1994) and Jeon (1996).² This extension of Holmstrom (1982b) gives rise to a special team production problem, where payoffs depend on the labor market's beliefs about the team members. Team members' incentives are determined by a dynamic learning process. In contrast, in the standard team production problem (see Holmstrom, 1982a) the firm provides incentives by designing pay-for-performance contracts. Second, I define *power* as the right to influence firm performance: a more powerful manager is someone whose decisions have a greater impact on his firm's performance. Furthermore, power is endogenous: the firm can choose how much power to give to each manager. The distribution of power determines the managers' visibility and their incentives to work in the interest of the firm.

Consider a risk-neutral firm (the principal) employing two managers, A and B (the agents), in periods $t \in \{1, 2\}$. The firm maximizes

2. There are differences. First, Meyer's (1994) focus is on efficient task assignment in a context where managers do not need to be given incentives, whereas I am interested precisely in the relationship between power and incentives. Second, Jeon (1996) takes the allocation of power as given and studies the effect of team composition on incentives. Auriol et al. (2002) use a model similar to Jeon (1996).

the discounted sum of expected profits using δ as a discount rate, and its technology in period t has the following form:

$$y_t = \varphi_A(\eta_A + e_{At}) + \varphi_B(\eta_B + e_{Bt}) + \varepsilon_t, \quad (1)$$

where y_t is output, φ_A is manager A 's power, e_{At} is manager A 's effort, η_A is manager A 's ability and φ_B , η_B and e_{Bt} are similarly defined for manager B . The abilities η_A and η_B are constants unknown to everyone (including the managers themselves), and all players have identical prior beliefs with distribution $\eta_i \sim N(\eta_0, \sigma^2)$ for $i \in \{A, B\}$ and $\text{cov}(\eta_A, \eta_B) = 0$. The random variable ε_t is a productivity shock to the firm, with $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ for $t \in \{1, 2\}$. These shocks are assumed to be independently distributed.

The key variable in (1) is power. Since power is defined as the right to affect firm performance, any increase in the power of a manager increases the marginal effect of his effort on firm performance. First, note that power does not reduce the variance of a manager's performance. This is different from Cerasi and Daltung's (2000) argument that if powerful managers make many decisions, then, by the law of large numbers, their individual levels of performance must have a low variance. However, it is the case that, in our model, more powerful managers make larger contributions to firm performance and, for this reason, their individual levels of performance are estimated with higher precision. This is similar to Cerasi and Daltung (2000). Second, note that power provides no private benefits in our model, as opposed to Rotemberg (1993) and Aghion and Tirole (1997). This implies that managers are not willing to *pay* for power, i.e., to accept lower salaries in return for power.

As in the incomplete contracting literature (see Hart and Moore, 1990, and Aghion and Tirole, 1997), I assume that power is a limited resource for the firm:

ASSUMPTION 1: $\varphi_A + \varphi_B \leq 1$.

This restriction is realistic if power is a precise set of tasks or responsibilities. For example, Prendergast's (1995) definition of *responsibility* as the number of tasks performed by an employee fits this definition. In his model, there is a continuum of tasks defined over the interval $[0, 1]$, and each task is assigned to one of two agents. In the particular case where all tasks are identical, both in the effort they require and in their marginal effect on performance, we may interpret φ_i as the number of tasks that agent i is responsible for. My definition of power can also be identified with one of French and Raven's (1959) categories. Their seminal paper distinguished four

types of power: coercive, legitimate, expert, and referent. Legitimate power is the power that comes from the position of a person in the formal hierarchy of an organization. This matches the definition I use in the model: legitimate power determines how much effect a manager can have on firm performance; it is a limited resource (Assumption 1); and it is visible to outsiders. The other types of power do not satisfy these conditions.

The following example illustrates our assumptions. Suppose two managers are in charge of launching a new product and have to decide how to use their budgets. Each manager is in charge of a different aspect of the project: manufacturing or marketing; and, for simplicity, assume these two aspects are independent (there is no complementarity). This assumption can be easily relaxed and will be relaxed later on. The company has to assign a budget I_i (for $i \in \{A, B\}$) to each manager. The new product generates returns measured by V , which depend on the two managers' choices. Let V_A and V_B (with $V_A + V_B = V$) denote the returns due to A and B respectively. Ideally, the firm would like to be able to decompose V into V_A and V_B , i.e., to know what part of the returns produced by the new product is attributable to the marketing and manufacturing decisions respectively. However, this is usually impossible. Formally, with $I = I_A + I_B$, the rate of return of the project is

$$\varrho = \frac{V - I}{I} = \frac{I_A}{I} \frac{V_A - I_A}{I_A} + \frac{I_B}{I} \frac{V_B - I_B}{I_B}.$$

The company would like to infer separate rates of return ϱ_A and ϱ_B for each manager based on V , I_A , and I_B . The power of a manager, in this case, is the relative size of his budget, $\varphi_i = I_i/I$, because we can write

$$\varrho = \varphi_A \varrho_A + \varphi_B \varrho_B,$$

where $\varrho_i = (V_i - I_i)/I_i$ depends on manager i 's effort.

Throughout this paper, "effort" has a broad meaning. Any action that a manager takes is *effort* as long as (a) it is more costly to the manager than to the firm; (b) it has a nonstochastic, positive effect on current performance; and (c) the manager knows this effect better than the firm. Internal budgeting decisions, for example, usually satisfy these conditions: the effects of budgetary choices are difficult to assess by persons who do not have as much information as the relevant managers; and their costs are often different for the managers and for the firm (managers might prefer some decisions that benefit their own department very directly but hurt the firm as a whole).

Another interesting example is the allocation of time or attention. A manager may choose to devote his time to improving a particular project in one of a number of dimensions. Some dimensions might be very important to the firm but, at the same time, uninteresting to the manager. Other aspects might be very rewarding to the manager, but irrelevant to the firm's interests. In this model, managers are considered to exert more effort when the actions they take are more in accordance with corporate objectives.

2.1.1 LABOR-MARKET COMPETITION. The labor market is competitive and observes both firm performance and the distribution of power. It does not observe the individual performance of each manager. Labor-market competition is important because part of the managers' human capital is general:

ASSUMPTION 2: *A proportion γ of a manager's ability is general human capital.*

If the labor market believes that manager i 's ability is η_i , the market value of that manager is $\gamma\eta_i$. With competition, this is the wage that every firm will be willing to pay to that manager. A key assumption is that the managers, the firm, and the market have exactly the same information at every point in time: the firm does not have more information about its managers than the labor market or the managers themselves. If the firm were better informed about its own managers than the rest of the labor market, the fear of transmitting too much information to competitors would determine its distribution of power. The firm would take into account that a manager to whom it has given more power could use his promotion to obtain a higher wage. This problem has been studied by Waldman (1984), Ricart-i-Costa (1988), and Bernhardt (1995), and is not considered here.

2.1.2 UTILITY FUNCTIONS. Both managers have the same differentiable, strictly increasing, convex disutility of effort $c(e)$, with $c(0) = 0$, $c'(0) = 0$, and $c'(e)$, $c''(e) > 0$ for $e > 0$. They also have the same risk-neutral utility function,

$$U(w_{i1}, w_{i2}; e_{i1}, e_{i2}) = w_{i1} - c(e_{i1}) + \delta[w_{i2} - c(e_{i2})], \quad (2)$$

where w_{it} is the wage paid to manager $i \in \{A, B\}$ in period t . The reservation utility is also the same for both managers and is equal to 0. The discount rate is δ for the two managers and for the principal.

2.1.3 TIMING. There are only two periods, and decisions are made in the following order:

- The market offers each manager i a wage w_{i1}^m for period 1. The principal offers each manager i a position φ_i for both periods, and a wage w_{i1} for period 1. Managers accept or reject the principal's offers.
- First-period production. Managers choose e_{A1} and e_{B1} . The productivity shock ε_1 is realized, and all players observe y_1 .
- The market offers each manager a second-period wage w_{i2}^m . The principal offers each manager a second-period wage w_{i2} . Managers accept or reject the principal's offers.
- Second-period production. Managers choose e_{A2} and e_{B2} . The productivity shock ε_2 is realized, and all players observe y_2 .

A key assumption, as far as the timing is concerned, is that the firm commits to keep in period 2 the allocation of power that it has previously chosen in period 1. This assumption is correct if the frequency with which the firm reallocates power is lower than the frequency with which it receives information about the managers and updates its beliefs about their abilities. For example, if managers learn by doing, it takes them time to learn about their new responsibilities, and job reassignments must not be too frequent.

2.2 THE EFFECT OF POWER ON VISIBILITY

Let e_{it}^* (for $i \in \{A, B\}$ and $t \in \{1, 2\}$) be the equilibrium levels of effort chosen by the managers. The normality assumptions imply that the posterior beliefs about ability are a weighted average of today's signal and all past signals (see De Groot, 1970). In particular, at $t = 1$

$$\eta_{A1} = (1 - \alpha_{A1})\eta_0 + \alpha_{A1} \left(\frac{z_1}{\varphi_A} - \frac{\varphi_B}{\varphi_A} \eta_0 \right), \quad (3)$$

where

$$\alpha_{A1} \equiv \frac{\varphi_A^2 \sigma_A^2}{\varphi_A^2 \sigma_A^2 + \varphi_B^2 \sigma_B^2 + \sigma_\varepsilon^2} \quad (4)$$

and $z_1 = y_1 - \varphi_A e_{A1}^* - \varphi_B e_{B1}^*$. A similar formula applies to manager B . In the above expression $\eta_{A1} = E(\eta_A | y_1)$, the posterior beliefs about manager A 's ability conditional on period-1 performance. The coefficient α_{A1} measures the rate at which the principal updates his beliefs about A : the higher α_{A1} , the more the principal blames this manager for a bad result or rewards him for a good one. Hence α_{A1} measures accountability.

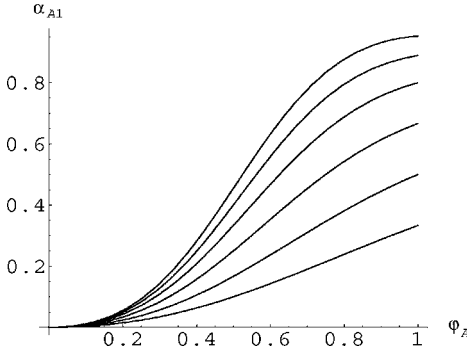


FIGURE 1. UPDATING COEFFICIENT FOR MANAGER A

LEMMA 1: If $\sigma_A^2 = \sigma_B^2 = \sigma^2$, then (i) α_{i1} is increasing in φ_i ; (ii) there exists $\tilde{\varphi}_i > 0$ such that α_{i1} exhibits increasing returns to power if $\varphi_i < \tilde{\varphi}_i$ and decreasing returns to power if $\varphi_i > \tilde{\varphi}_i$; and (iii) $\tilde{\varphi}_i \leq 1$ if $\sigma_\varepsilon \leq \sigma$ and $\tilde{\varphi}_i > 1$ if $\sigma_\varepsilon > \sigma$.

Proof. See Appendix. □

The first part of the lemma states that more powerful managers are held more accountable, because the labor market can have more precise information about their abilities. The second part of the lemma is illustrated in Figure 1, where α_{A1} is plotted as a function of φ_A for a given value of σ_ε^2 ($\sigma_\varepsilon^2 = 1$) and different values of $\sigma_A^2 = \sigma_B^2$. Lower curves correspond to higher prior variances. There are increasing returns to power when power is low, and decreasing returns to power when power is high.

Despite this ambiguity, increasing returns to power dominate as far as *team* incentives are concerned. Before we present a formal proof, let us explain the intuition for it. First, note that we are always more likely to update our beliefs about a manager when our prior beliefs about him are weak. Second, the same logic applies to the team of managers considered as a whole: the weaker our prior beliefs about the team, the more quickly we will update them. Third, our prior beliefs about the team are always more precise when we divide power more equally: in that case we have a more diversified portfolio of managers, and there will be less uncertainty about the team's performance. Therefore, if power is distributed more equally between the managers, the beliefs about the team will be updated more slowly. Finally, if the updating is slower, then the incentives of the team as a

whole will also be lower. This means that there are increasing returns to power because, if we move away from the perfectly even distribution of power, the increase in one of the managers' incentives will be greater than the reduction in the other manager's incentives.

More formally, note that the sum of the updating coefficients is

$$\alpha_{A1} + \alpha_{B1} = \frac{(\varphi_A^2 + \varphi_B^2)\sigma^2}{(\varphi_A^2 + \varphi_B^2)\sigma^2 + \sigma_\varepsilon^2} = \frac{\sigma_J^2(\varphi_A)}{\sigma_J^2(\varphi_A) + \sigma_\varepsilon^2},$$

where $\sigma_J^2(\varphi_A) = (\varphi_A^2 + \varphi_B^2)\sigma^2$ is the prior variance of the team's productivity, and is minimized at $\varphi_i = \frac{1}{2}$. This sum determines the incentives of the team as a whole. Now suppose that the principal had hired only one manager, with prior ability $\varphi_A\eta_0 + \varphi_B\eta_0 = \eta_0$ and prior variance $\sigma_J^2(\varphi_A)$. According to the Bayesian updating rules, the principal would use a rate $\alpha_{A1} + \alpha_{B1}$ to update his beliefs about this imaginary manager.³ But this rate is higher the higher the prior uncertainty about the manager:

LEMMA 2: *The sum $\alpha_{A1} + \alpha_{B1}$ is maximized at $\varphi_i^* \in \{0, 1\}$ and minimized at $\varphi_i^* = \frac{1}{2}$. Furthermore, it is increasing in σ^2 and decreasing in σ_ε^2 .*

The rate of updating for the imaginary manager, $\alpha_{A1} + \alpha_{B1}$, is maximized at the corners ($\varphi_i^* \in \{0, 1\}$), because this is where the prior uncertainty about his ability is highest. Hence, as far as the team of managers is concerned, there are increasing returns to power in learning and incentives.

3. THE OPTIMAL ALLOCATION OF POWER: A BENCHMARK

Suppose that managers are *ex ante* identical ($\eta_{A0} = \eta_{B0} = \eta_0$ and $\sigma_A^2 = \sigma_B^2 = \sigma^2$). To analyze the optimal distribution of power, let $w_{it}(\eta_{it-1}, e_{it}^*)$ be the wage that the firm offers to manager i in period t given his expected ability (η_{it-1}) and effort (e_{it}^*). Similarly, let $w_{it}^m(\eta_{it-1})$ denote the wage offered to manager i by the labor market in period t . Labor-market competition imposes the constraints

$$\begin{aligned} w_{it}(\eta_{it-1}, e_{it}^*) - c(e_{it}^*) &= w_{it}^m(\eta_{it-1}), \\ w_{it}^m(\eta_{it-1}) &= \gamma\eta_{it-1} \end{aligned} \tag{5}$$

for $i \in \{A, B\}$ and $t \in \{1, 2\}$: because of labor-market competition, managers obtain a salary raise every time they are believed to have

3. In the single-manager case, with $y_t = \eta + \varepsilon_t$, $\eta \sim N(\eta_0, \sigma^2)$, and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, the posterior ability is $\eta_t = (1 - \alpha_t)\eta_0 + \alpha_t y_t$ and the updating rate is $\alpha_t = \sigma^2/(\sigma^2 + \sigma_\varepsilon^2)$. Hence, when there is a single manager for whom prior beliefs have a variance of $\sigma_J^2(\varphi_A)$, the rate of updating will be $\sigma_J^2/(\sigma_J^2 + \sigma_\varepsilon^2)$.

a higher ability. Note that more powerful managers are paid higher salaries because they have to be compensated for their higher levels of effort. This would not be so if power produced private benefits for managers, as in Rotemberg (1993) and Aghion and Tirole (1997).

Managers have an incentive to choose high levels of effort in the first period: by doing so they can improve performance, make the market believe that their ability is high, and earn higher wages in period 2. However, such incentives do not exist in the last period, and $e_{A2}^* = e_{B2}^* = 0$. The incentive-compatibility constraint for period 1 is

$$e_{i1}^* = \arg \max_{e_{i1}} [\delta \gamma \eta_{i1} - c(e_{i1})], \quad (6)$$

which leads to $c'(e_{i1}^*) = \delta \gamma \alpha_{i1}$; and the individual rationality constraints are

$$\begin{aligned} w_{i1} + \delta E(w_{i2}) - c(e_{i1}) &\geq 0, \\ w_{i2} &\geq 0. \end{aligned} \quad (7)$$

When posterior abilities happen to be negative, the second constraint in (7) will be binding. However, following a simplification that is standard in this literature, we ignore this possibility. This is a good approximation only if the prior mean η_0 is high and negative posterior values for η_i happen with very small probabilities.

Under these assumptions, the optimal allocation of power is uneven, as shown by the following result:

PROPOSITION 1: *If managers are ex ante identical, it is optimal to distribute power in an uneven way ($\varphi_A^* \neq \varphi_B^*$). Moreover, if $c(e_{ii}) = e_{ii}^2/2$, the optimal allocation of power is a corner solution ($\varphi_A^* \in \{0, 1\}$).*

Proof. See Appendix. □

There are two reasons for this result. First, the effect of power on learning and incentives is not linear: there are increasing returns to power (Lemmas 1 and 2). Second, power and effort are complements in (1), and this complementarity plays a key role. In fact, suppose that power is equally distributed between *A* and *B*, and imagine that *A*, for an exogenous reason, decides to choose a higher level of effort. The firm will have an incentive to give *A* more power, in order to benefit more from his higher effort. Having more power, *A* will have stronger career concerns and will work even harder, making it profitable for the firm to give him even more power. Meanwhile, as *A*'s power increases, *B* has less and less power, and chooses to exert less and

less effort. However, although this is costly to the firm, this cost is smaller and smaller, because B has less and less power, and therefore his effort has less and less importance. Thus very small differences in the managers' propensities to work can lead to very large power differences.

Meyer (1991) proposed a different reason why learning could lead to unequal distributions of power. She analyzed a promotion contest as a device used by the firm to learn about its promotion candidates. She found that the introduction of a bias toward the favorite candidate would improve learning. Proposition 1 is different from Meyer's (1991) result in two respects. First, her article identifies pure learning effects and does not link these learning effects to the incentives of managers, as opposed to Proposition 1. Second, in her model the firm can only use order statistics to learn about the employees. In my model, the firm also uses coarse information, but the coarseness is of a different kind: team performance is observed, but individual performance is not. However, despite these two differences, the existence of increasing returns to power in my model is consistent with Meyer's (1991) idea that, when the firm is learning from coarse information, it is not profitable to give the same power to all the employees that it wants to learn about.

The result in Proposition 1 is robust to the introduction of explicit incentive schemes. If the firm can offer linear incentives (as in Gibbons and Murphy, 1992) and incentives are based on firm performance, the firm will still choose to divide power in an unequal way, as the following proposition shows.

PROPOSITION 2: *Suppose the two managers are ex ante identical, the firm can offer linear contracts $w_{it} = a_i + b_{it}y_t$ ($i \in \{A, B\}$, $t \in \{1, 2\}$), and $c(e_{it}) = e_{it}^2/2$. Then the optimal allocation of power is a corner solution ($\varphi_A^* \in \{0, 1\}$).*

Proof. See Appendix. □

When managers are risk-neutral, optimal explicit incentives make them behave as residual claimants, and efforts move up to first-best levels. But first-best efforts are increasing functions of power: since the effort of more powerful managers has a larger effect on firm performance, in a first-best world they should work harder. In fact, the first-best level of effort is $e_{i1} = \varphi_i$. Since first-best efforts are increasing functions of power, effort and power are complements, and the optimal allocation of power is uneven.

4. EXTENSIONS AND MANAGERIAL IMPLICATIONS

The main strength of our model lies in its ability to produce qualitative predictions on how to distribute power in different contexts. The effect of power on the managers' visibility matters because it imposes constraints on several personnel and organizational decisions. The following extensions of the model are intended to illustrate this point. For simplicity, a quadratic cost function [$c(e_{it}) = e_{it}^2/2$] is used throughout this section.

4.1 POWER AND DEPARTMENTATION

Multidivisional companies usually combine two basic criteria for departmentation: a functional criterion and a market-based criterion. Functional departmentation creates units responsible for each of the business functions (finance, operations, marketing, and personnel). Market-oriented departmentation creates units responsible for a product line, a physical area, or some other market-dependent variable. According to Mintzberg (1979), "divisions are created according to markets served and are then given control over the operating functions required to serve these markets" (p. 381). For example, in many companies each division is responsible for a different product line, but its internal departmentation follows functional lines.

This pattern of departmentation has two interesting features. First, there is one market measure that makes it possible to assess the performance of each division, and one manager (the head of the division) is responsible for that measure of performance. This is consistent with our previous results: if an observable measure of performance is available, it is profitable to make a single manager responsible for it. In this way, the division manager is more visible than he would be if his division were responsible for a business function rather than a product. Second, consider the managers who are one level below the division head. Each of them has responsibility over one of the business functions, and, as a consequence, their responsibilities are complementary. This provides division heads with more power. If a product manager is given power over finance, marketing, and operations, but not personnel, he can always argue that the progress of his division is hindered by inadequate personnel policies. This division head would be more visible (and more accountable) if he were made responsible for all business functions, not just because he would have more power, but because any of his decisions would be more fully implemented and would have a larger effect on his division's performance. If the company decided to give him more power by giving him control of an activity unrelated to his other responsibilities, his

visibility would not increase as much. Hence the most effective way to increase a manager's power is to give him more decision rights on dimensions that complement the ones he is already in charge of.

Another interesting issue is how the firm should divide power between the managers who belong to the same division and are immediately below the head. If each of these managers is responsible for one of the business functions, their efforts are complementary, and our initial production function, given by (1), has to be redefined. We can take intradivisional complementarities into account by assuming

$$y_t = \varphi_A(\eta_A + e_{At}) + \varphi_B(\eta_B + e_{Bt}) + \zeta \varphi_A e_{At} \varphi_B e_{Bt} + \varepsilon_t, \quad (8)$$

where the importance of complementarities is measured by ζ . This parameter measures to what extent responsibilities are divided along complementary lines. With $c(e_{it}) = e_{it}^2/2$, the incentive compatibility constraint (for manager A) is

$$e_{A1} = \delta \gamma (\alpha_{A1} + \zeta \varphi_B e_{B1}), \quad (9)$$

and in a Nash equilibrium

$$e_{A1}^* = \delta \gamma \alpha_{A1} \frac{1 + \delta \gamma \zeta \varphi_B \alpha_{B1}}{1 - \delta^2 \gamma^2 \zeta^2 \varphi_A \varphi_B \alpha_{A1} \alpha_{B1}}. \quad (10)$$

PROPOSITION 3: *Suppose that the managers' actions are complementary as in (8), with the degree of complementarity being measured by ζ . Then there is a value $\zeta_0 > 0$ such that for every $\zeta > \zeta_0$ the even distribution of power $\varphi_A = \frac{1}{2}$ is preferred to the uneven distribution of power $\varphi_A \in \{0, 1\}$; and for every $\zeta < \zeta_0$ the uneven distribution $\varphi_A \in \{0, 1\}$ is preferred to the even distribution $\varphi_A = \frac{1}{2}$.*

Proof. See Appendix. □

Consider the managers who are immediately below the division head. If each of these managers is responsible for one of the business functions, then $\zeta > 0$. According to Proposition 3, power should be divided on a rather equal basis among these managers, even taking into account that this will make them less visible. Thus our model provides a rationale for the type of organizational structure where each division is responsible for all the business functions related to a specific product and managers responsible for the various business functions have similar amounts of power. First, the division head is made more visible in this way, because there exists an observable (but noisy) measure of his performance. Second, the division head has more influence on his division's performance, because he can use all

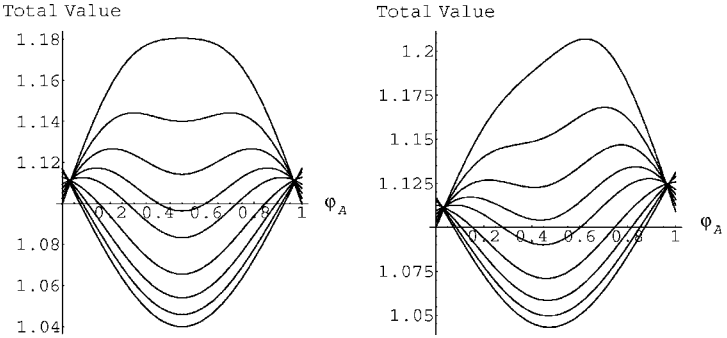


FIGURE 2. TOTAL EXPECTED VALUE WITH COMPLEMENTARITIES

business functions to implement his decisions. Third, by giving similar amounts of power to all functional managers, the lack of visibility of these managers is compensated by the fact that their complementarities will be more fully exploited.

The model also shows some determinants of power differences among managers. Figure 2 plots total expected value as a function of φ_A for different values of ζ .⁴ The left-hand side represents the symmetric case, where managers are a priori identical, while the right-hand side shows the case where prior information about manager A is less precise than that about manager B ($\sigma_A^2 > \sigma_B^2$). In the left graph, higher plots correspond to higher ζ 's: as complementarities become stronger, the optimal allocation of power becomes more equal. In the plot on the right, the asymmetry leads to giving more power to manager A than to manager B : when differences in ability are not substantial, the firm should give more power to younger managers who are eager to prove themselves. Asymmetries in prior abilities yield similar results: holding everything else equal for the two managers, if $\eta_{A0} > \eta_{B0}$ then more power should be given to A than to B . Similarly, if A 's job has a higher content of general human capital (higher γ for A than for B) then A should have more power than B .

4. More precisely, we have plotted $E_0(y_1 - c(e_{A1}) - c(e_{B1}))$. In fact, the expected total value is $E_0(y_1 - c(e_{A1}) - c(e_{B1}) + \delta(y_2 - c(e_{A2}) - c(e_{B2})))$, but $e_{A2} = e_{B2} = 0$ independently of the allocation of power; and we are also considering $\eta_{A0} = \eta_{B0} = \eta_0$. Hence there is no loss of generality from looking at the first-period value only. To generate the plot we have chosen $c(e) = 0.5e^2$, $\eta_{A0} = \eta_{B0} = 1$, $\gamma = \delta = 0.5$, $\sigma_A^2 = 1$, $\sigma_B^2 = 1$, and $\zeta \in \{0, 1, 2, 3, 4, 4.5, 5, 5.5, 6\}$. In the left-hand side $\sigma_A^2 = \sigma_B^2 = 1$, and in the right-hand side $\sigma_A^2 = 1.3 > \sigma_B^2$. These results are robust to different parameter values.

4.2 POWER AND SENIORITY

A traditional feature of internal labor markets is the high positive correlation between power and seniority, i.e., the fact that an employee's progression within an organization depends crucially on his in-the-firm tenure. Referring to blue-collar workers employed in manufacturing, Doeringer and Piore (1971) observed that "promotions to nonentry jobs are generally determined by fixed standards of seniority and ability." Seltzer and Merrett (2000), referring to white-collar jobs in a bank, reported that "even when forthcoming, promotion was typically slow." However, both Seltzer and Merrett (2000) and Baker et al. (1994) find evidence of *fast tracks*—quick paths of promotion.

The weight given to seniority is usually explained by the importance of human-capital accumulation, and the existence of fast tracks is usually related to sorting: on the one hand, if higher-level jobs require substantially more human capital, then seniority must be strongly correlated with power. On the other hand, the firm should also be able to promote more quickly employees who have greater abilities and therefore learn more quickly. Thus, fast tracks are also needed. However, although fast tracks can be viewed as a sorting mechanism, they also have clear effects on incentives: fast tracks produce younger managers, and younger managers have different incentives than older managers.

A simple consequence of our model is that younger managers have stronger career concerns than older managers, and therefore will choose higher levels of effort than the latter. However, this is so because our model does not take human-capital accumulation into account. Let us now assume that there is a rate λ of on-the-job learning, and as a consequence of that the prior ability of an employee i with τ periods of tenure is $\theta_{i0} = (1 + \lambda)^\tau$. To simplify, consider the case without complementarities ($\zeta = 0$) where, by Proposition 1, the optimal allocation is either $\varphi_A = 1$ or $\varphi_B = 1$, and comparisons are particularly simple. On the one hand, more-senior managers have more human capital, and therefore a higher prior ability ($\theta_{A0} > \theta_{B0}$). On the other hand, the firm knows their abilities much better ($\sigma_A^2 < \sigma_B^2$) and these managers will have less incentives to prove their value. Hence, giving more power to young managers is more advantageous if the abilities of junior and senior managers are not too different, as shown by the following result.

PROPOSITION 4: *Suppose manager A's tenure is $\tau > 1$ periods and manager B's tenure is one period. Suppose managers accumulate human capital at a rate λ per period, and suppose that the precision about the managers'*

abilities increases at a rate l per period. Then for every $l < \tau^{1/(\tau-1)} - 1$ there exists a $\lambda_0(l) > 0$ such that $\varphi_A^* = 1$ if $\lambda > \lambda_0(l)$ and $\varphi_B^* = 1$ if $\lambda < \lambda_0(l)$. Furthermore, $\lambda_0(l)$ is increasing in l .

Proof. See Appendix. □

The interest of this simple result lies in its implications for the allocation of power in different environments. First, consider new vs. traditional lines of activity. In newer activities senior employees are not necessarily better trained than junior employees. Hence, power should be given to the latter. This is the case, for example, in jobs requiring high information-technology skills: young employees in software companies can be promoted as quickly as older employees, because the knowledge needed for higher-level jobs is very new. Second, some jobs require employees to set objectives or to design plans, while other jobs require employees to implement the plans that others have established. The former jobs require mostly experience, and the latter require a lot of time and effort, but not necessarily a lot of expertise. According to our model, junior managers can be given considerable responsibility in activities requiring a large amount of time or attention and little expertise. Third, there are activities where the knowledge that is needed is cumulative in nature and activities where it is not. For example, legal or medical jobs require qualifications that depend heavily on experience. Similarly, supervision jobs require experience, because a supervisor needs to know the characteristics of the jobs that he is supervising. In these cases, seniority should be a more important determinant of an employee's position. In other types of activities, the qualifications that are required do not depend very much on previous experience, but depend to a large extent on the employee's effort. In these cases, seniority should be a less important determinant of power.

4.3 POWER AND SPECIALIZATION

A stylized fact in organizations is that higher-level jobs are wider, i.e., include more tasks or more diverse ones, than lower-level jobs. Top managerial jobs are usually seen as prototypes of nonspecialization: for example, according to Mintzberg (1979, pp. 79–80), “these roles managers perform are so varied, and so much switching is required among them in the course of any given day, that managerial jobs are typically the least specialized in the organization.”

Narrow job definitions have costs and benefits. Specialized managers can be more productive than nonspecialized managers because

they are able to concentrate their learning on a specific issue, of which they become experts. However, good firm performance always results from a variety of complementary decisions for which different types of expertise are needed. Therefore, generalists are more visible to the outside labor market than specialists: a generalist is more likely to be held accountable than a specialist, who is only responsible for a limited (and difficult to isolate) aspect of the problem. Thus, the firm has to choose between developing specialists, who will become experts but will be less visible to the outside market, and developing generalists, who will not be experts but will be more visible to the outside market.

Note that, in the model, any of the corner solutions ($\varphi_i = 1$) maximizes autonomy and minimizes the degree of specialization; and the middle solution, $\varphi_A = \varphi_B = \frac{1}{2}$, maximizes the degree of specialization of the team.⁵ Because the technology given by (1) does not take into account the benefits of specialization, we now consider the following alternative function:

$$y_t = \sqrt{\varphi_A}(\eta_A + e_{At}) + \sqrt{\varphi_B}(\eta_B + e_{Bt}) + \varepsilon_t. \quad (11)$$

This new production function takes into account the trade-off between specialization and autonomy: more autonomy makes a manager more visible and hence generates incentives; but it also reduces specialization and makes the manager less productive. With this technology, the updating coefficient is simply

$$\alpha_{A1} = \frac{\varphi_A \sigma_A^2}{\varphi_A \sigma_A^2 + \varphi_B \sigma_B^2 + \sigma_\varepsilon^2}. \quad (12)$$

Figure 3 shows expected total value as a function of φ_A for different values of σ_ε^2 .⁶ The upper curve corresponds to the case where firm performance is measured with high precision (σ_ε^2 is low): in this case, the benefits of autonomy are higher and the power allocation should be relatively uneven. The lower curve, where equality is optimal, corresponds to a case where performance measurements are difficult (σ_ε^2 is high).

5. If φ_A is very small, manager A will be very specialized, but in that case φ_B will be very large, and manager B will not be specialized. Maximum specialization is reached when power is divided equally.

6. We have again plotted $E_0(y_1 - c(e_{A1}) - c(e_{B1}))$ using $c(e) = 0.08e^2$, $\eta_{A0} = \eta_{B0} = 1$, $\gamma = \delta = 0.5$, $\sigma_A^2 = \sigma_B^2 = 1$, and $\sigma_\varepsilon^2 \in \{0.05, 1.5\}$. As in previous plots, these results are robust and the particular parameter values have only been chosen to make the figure illustrative.

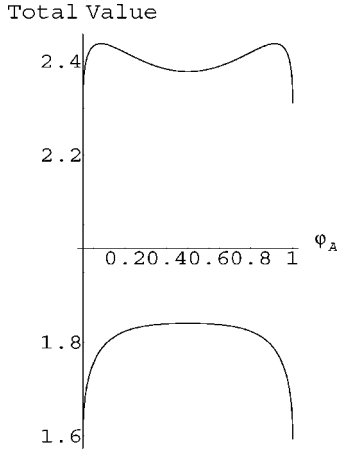


FIGURE 3. AUTONOMY VS. SPECIALIZATION

PROPOSITION 5: *If managers are a priori identical and there are benefits to specialization (given by (11)), then there exists $\sigma_0^2(\sigma^2) > 0$ such that: (i) if $\sigma_\varepsilon^2 < \sigma_0^2$, then $0 < \varphi_A^* < \frac{1}{2}$ or $\frac{1}{2} < \varphi_A^* < 1$; and (ii) if $\sigma_\varepsilon^2 > \sigma_0^2$, then $\varphi_A^* = \frac{1}{2}$. Furthermore, $\sigma_0^2(\sigma^2)$ is decreasing in σ^2 .*

Proof. See Appendix. □

This proposition has interesting implications for job design: suppose we consider equation (11) at different levels of the hierarchy. As we approach the top levels, managers become more and more visible to the market, and σ_ε^2 becomes smaller. This is so independently of whether the manager is a specialist or a generalist: no matter what the manager does, if he is at a higher level, then the σ_ε^2 corresponding to his level must be smaller. At any given level, though, the firm can choose to use narrow or wide job definitions. According to Proposition 5, the firm should choose wider jobs as σ_ε^2 diminishes. Hence, as managers move up the hierarchy, they should become generalists. This happens because at higher positions the performance of managers is more easily assessed by the market. Proposition 5 shows that the firm has an incentive to take advantage of this increased visibility and enlarge job definitions at those levels. On the contrary, employees are less visible at lower levels of the hierarchy, and the firm has little to gain by using wide job definitions. Hence, the benefits of

specialization are greater in this case, which is consistent with the stylized fact we mentioned above.

4.4 PATTERNS OF MANAGERIAL SUCCESSION

According to Parrino (1997), “the availability of a strong outside candidate is an important consideration in the decision to replace a poor CEO.” In his study of CEO departures in the US, he found that “when accounting performance is not substantially below that of the industry, but has been declining, the potential benefits from an outside appointment are not sufficient to outweigh the costs. Consequently, when boards decide to replace the CEO in such situations they tend to appoint insiders.” Boards are more likely to appoint outsiders when the firm’s performance is substantially below the industry average. A more recent paper by Huson et al. (2001) finds similar evidence.

This finding can easily be explained with our model. Suppose that the firm chooses an allocation of power φ_{i1} for period 1 and can change it after having observed first-period performance. Let φ_{i2} be the allocation of power in the second period. Suppose also that after observing first-period performance the firm has the option to hire a new manager (an outsider). Furthermore, suppose that insiders accumulate human capital during the first period, while outsiders do not. As in Section 4.2, let λ be the rate of human-capital accumulation of insiders: then the posterior abilities, after first-period performance has been observed, are $(1 + \lambda)\eta_{A1}$ for manager A , $(1 + \lambda)\eta_{B1}$ for manager B , and η_0 for every outsider.

Without loss of generality, suppose $\varphi_{A1} > \varphi_{B1}$. Then, if $\varphi_{A2} > \varphi_{B2}$, we will consider that there is no succession in period 2; if $\varphi_{A2} < \varphi_{B2}$, we will call this an inside succession; and if the firm fires A and B and hires an outsider for period 2, we will call this an outside succession.

PROPOSITION 6: *Suppose the firm can hire new managers at the end of period 1. Suppose insiders accumulate human capital at a rate λ and outsiders do not accumulate human capital. Then for every λ there exists $\underline{z}(\lambda) < \eta_0$ such that (i) if the first-period performance is sufficiently low ($z_1 < \underline{z}(\lambda)$), then outside succession takes place in period 2; (ii) if the first-period performance is intermediate ($\underline{z}(\lambda) < z_1 < \eta_0$), then inside succession takes place in period 2; and (iii) if the first-period performance is sufficiently high ($z_1 > \eta_0$), no succession takes place. Furthermore, $\underline{z}(\lambda)$ is decreasing in λ .*

Proof. Without loss of generality, suppose that $\varphi_{A1} > \varphi_{B1}$. In period 2, only one of the following three choices can be optimal: (a) to give all the power to A (no succession); (b) to give all the power to B (inside

succession); or (c) to give all the power to a third person (outside succession). In case (a), the second-period expected surplus will be equal to $S_2^a = (1 + \lambda)\eta_{A1}$; in case (b) it will be equal to $S_2^b = (1 + \lambda)\eta_{B1}$; and in case (c) it will be equal to $S_2^c = \eta_0$.

First, option (a) will be chosen if $S_2^a > S_2^b$ and $S_2^a > S_2^c$, which is equivalent to $z_1 > \eta_0$. Second, option (b) will be chosen if $S_2^b > S_2^a$ and $S_2^b > S_2^c$. After some manipulation, it can be found that this is equivalent to $z_1 < \eta_0$ and

$$z_1 > \frac{\eta_0}{\varphi_{B1}(1 + \lambda)\sigma^2} [\varphi_{B1}^2\sigma^2 + \varphi_{A1}\varphi_{B1}(1 + \lambda)\sigma^2 - \lambda\varphi_{A1}^2\sigma^2 - \lambda\sigma_\varepsilon^2] \equiv \underline{z}(\lambda), \quad (13)$$

which completes the proof. \square

This proposition is consistent with the evidence found by Parino (1997) and Huson et al. (2001). These authors find that the probability of a fired CEO being replaced by an outsider is more sensitive to firm performance, measured by industry-adjusted return on assets, than the probability of a fired CEO being replaced by an insider. This means that fired CEOs are more likely to be replaced by insiders when firm performance is close to industry performance; and CEOs are more likely to be replaced by outsiders when firm performance is substantially below the industry average.

5. CONCLUSIONS

I have argued that the allocation of power at top levels of the hierarchy affects the visibility of managers, and, if labor markets are competitive, their visibility has an effect on incentives. Hence, if labor markets are competitive, one way to align the objectives of top managers with those of shareholders is to define their responsibilities in a way that maximizes their visibility. This article makes two basic contributions in this direction. First, it shows that if firms want to make their managers more visible, they will distribute power in an unequal way. This is not an obvious result, because making a manager very powerful (very visible) implies making other managers less powerful (less visible). Second, the article goes beyond this basic result and analyzes how companies can design jobs in order to increase their managers' visibility. I argue that the criteria of departmentation, the importance given to seniority, and the width of job definitions are some of the variables that the firm can use for that purpose.

APPENDIX

A.1 PROOF OF LEMMA 1

Differentiating α_{A1} with respect to φ_A , we find $\partial\alpha_{A1}/\partial\varphi_A > 0$, and $\partial^2\alpha_{A1}/\partial\varphi_A^2 > 0$ or $\partial^2\alpha_{A1}/\partial\varphi_A^2 < 0$ depending on whether φ_A is low or high.

A.2 PROOF OF PROPOSITION 1

Since managers are *ex ante* identical, they have the same prior ability, and the distribution of power only affects output through effort. Hence the optimal allocation of power maximizes

$$S_1(\varphi_A) = \varphi_A e_{A1}^* + \varphi_B e_{B1}^* - c(e_{A1}^*) - c(e_{B1}^*),$$

where $c'(e_{A1}^*) = \delta\gamma\alpha_{A1}$ and $c'(e_{B1}^*) = \delta\gamma\alpha_{B1}$. The function $S_t(\varphi_A)$ gives the total surplus in period t . As of period 0, total surplus in period 2, $S_2(\varphi_A)$, is not a function of φ_A and therefore does not need to be considered.

First, I show that $\varphi_A = \frac{1}{2}$ is a local *minimum*: the first-order condition is

$$e_{A1}^* - e_{B1}^* + \varphi_A \frac{\partial e_{A1}^*}{\partial \varphi_A} + \varphi_B \frac{\partial e_{B1}^*}{\partial \varphi_A} - \gamma\delta\alpha_{A1} \frac{\partial e_{A1}^*}{\partial \varphi_A} - \gamma\delta\alpha_{B1} \frac{\partial e_{B1}^*}{\partial \varphi_A} = 0,$$

using the incentive compatibility constraints. Since the two managers are *ex ante* identical, it is easily verified that $\varphi_A = \frac{1}{2}$ satisfies this condition. On the other hand, the second-order condition evaluated at $\varphi_A = \frac{1}{2}$ is, after some manipulation,

$$\left(4 - 2\gamma\delta \frac{\partial \alpha_{A1}}{\partial \varphi_A}\right) \frac{\partial e_{A1}^*}{\partial \varphi_A} > 0, \tag{14}$$

which is satisfied because

$$\frac{\partial \alpha_{A1}}{\partial \varphi_A} = \frac{2\sigma^2}{\sigma^2 + 2\sigma_\varepsilon^2}$$

at $\varphi_A = \frac{1}{2}$. Hence $\varphi_A = \frac{1}{2}$ is a local minimum and cannot be a global maximum.

Second, I show that if $c(e_{it}) = e_{it}^2/2$, then the optimal allocation is the corner solution $\varphi_A \in \{0, 1\}$. With a quadratic cost function, we have

$$S_1(1) = \delta\gamma \left[\frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} - \frac{\delta\gamma}{2} \left(\frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \right)^2 \right] \quad (15)$$

and

$$S_1(\varphi_A) = \delta\gamma \left(\frac{(\varphi_A^3 + \varphi_B^3)\sigma^2}{(\varphi_A^2 + \varphi_B^2)\sigma^2 + \sigma_\varepsilon^2} - \frac{\delta\gamma}{2} \frac{(\varphi_A^4 + \varphi_B^4)\sigma^4}{[(\varphi_A^2 + \varphi_B^2)\sigma^2 + \sigma_\varepsilon^2]^2} \right). \quad (16)$$

To prove that $S_1(1) > S_1(\varphi_A)$ for every $\varphi_A \in (0, 1)$, it suffices to prove that the following stronger condition is satisfied for every $\varphi_A \in (0, 1)$ and every δ and γ :

$$\begin{aligned} & \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} - \frac{1}{2} \left(\frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \right)^2 \\ & > \frac{(\varphi_A^3 + \varphi_B^3)\sigma^2}{(\varphi_A^2 + \varphi_B^2)\sigma^2 + \sigma_\varepsilon^2} - \frac{1}{2} \frac{(\varphi_A^4 + \varphi_B^4)\sigma^4}{[(\varphi_A^2 + \varphi_B^2)\sigma^2 + \sigma_\varepsilon^2]^2}. \end{aligned} \quad (17)$$

Developing this inequality, we obtain the following equivalent condition:

$$\frac{\sigma^4 + 2\sigma^2\sigma_\varepsilon^2}{2(\sigma^2 + \sigma_\varepsilon^2)^2} \geq \frac{[2(\varphi_A^5 + \varphi_B^5) - (\varphi_A - \varphi_B)^2]\sigma^4 + 2(\varphi_A^3 + \varphi_B^3)\sigma^2\sigma_\varepsilon^2}{2[(\varphi_A^2 + \varphi_B^2)\sigma^2 + \sigma_\varepsilon^2]^2}. \quad (18)$$

The left-hand side of this inequality is not a function of φ_A , and the right-hand side is U-shaped with a minimum at $\varphi_A = \frac{1}{2}$. Moreover, at $\varphi_A = \frac{1}{2}$ condition (18) is satisfied with equality. Since (18) implies (17), we have proved that the optimal allocation of power is a corner solution. This completes the proof of Proposition 1.

A.3 PROOF OF PROPOSITION 2

Suppose the firm can offer the managers linear contracts of the form $w_{it} = a_i + b_{it}y_{it}$. First, note that in period 2 it will be optimal to set $b_{A2}^* = b_{B2}^* = 1$, and as a consequence we will have $e_{A2}^* = \varphi_A$ and $e_{B2}^* = \varphi_B$. To find this it suffices to maximize total period-2 surplus, $S_2(\varphi_A)$, with respect to b_{A2} and b_{B2} . The solution is straightforward, because, managers being risk-neutral, it is optimal to make them residual claimants.

Second, note that period-2 salaries will depend on period-1 performance:

$$a_{A2}^* = \gamma \eta_{A1} + c(e_{A2}) - b_{A2}(\varphi_A \eta_{A1} + \varphi_B \eta_{B1} + \varphi_A e_{A2}^* + \varphi_B e_{B2}^*). \quad (19)$$

Third, consider the period-1 problem. Manager A will choose effort so as to maximize his expected stream of income. Hence, if $c(e_{it}) = e_{it}^2/2$ his incentive compatibility (IC) constraint is

$$e_{A1}^* = \frac{\partial w_{A1}}{\partial e_{A1}} + \delta \frac{\partial E(w_{A2})}{\partial e_{A1}}, \quad (20)$$

where $E(w_{A2})$ is conditional on the information that manager A has at the end of period 1. This yields

$$e_{A1}^* = b_{A1} \varphi_A + \delta [\gamma \alpha_{A1} - \varphi_A (\alpha_{A1} + \alpha_{B1})] \quad (21)$$

and an optimal incentive rate

$$b_{A1}^* = 1 + \delta \left(\alpha_{A1} + \alpha_{B1} - \gamma \frac{\alpha_{A1}}{\varphi_A} \right). \quad (22)$$

Substituting this back into (21) yields $e_{A1}^* = \varphi_A = e_{A2}^*$, and similarly $e_{B1}^* = \varphi_B = e_{B2}^*$. As a consequence, it is optimal to choose $\varphi_A \in \{0, 1\}$. This completes the proof of Proposition 2.

A.4 PROOF OF PROPOSITION 3

With complementarities, the total expected surplus is given by

$$S_1^c(\varphi_A) = \varphi_A e_{A1}^* + \varphi_B e_{B1}^* + \zeta \varphi_A e_{A1}^* \varphi_B e_{B1}^* - c(e_{A1}^*) - c(e_{B1}^*), \quad (23)$$

where effort levels are given by expression (9). First, note that $S_1^c(1) = S_1(1)$ and is not a function of ζ :

$$S_1^c(1) = e_{A1}^*(1) - \frac{e_{A1}^{*2}(1)}{2}, \quad (24)$$

where

$$e_{A1}^*(1) = \frac{\delta \gamma \sigma^2}{\sigma^2 + \sigma_\varepsilon^2}. \quad (25)$$

Second, note that

$$S_1^c\left(\frac{1}{2}\right) = e_{A1}^*\left(\frac{1}{2}\right) + \left(\frac{\zeta}{4} - 1\right) e_{A1}^{*2}\left(\frac{1}{2}\right), \quad (26)$$

where

$$e_{A1}^*\left(\frac{1}{2}\right) = \frac{4\delta\gamma\sigma^2(2\sigma^2 + 4\sigma_\varepsilon^2) + 2\delta^2\gamma^2\zeta\sigma^4}{4(2\sigma^2 + 4\sigma_\varepsilon^2)^2 - \delta^2\gamma^2\zeta^2\sigma^4}. \quad (27)$$

Third, by Proposition 1, if $\zeta = 0$ then $S_1^c\left(\frac{1}{2}\right) < S_1^c(1)$.

Fourth, it can also be proved that $S_1^c(\frac{1}{2})$ is increasing in ζ . This is so because total expected surplus is increasing in effort when effort is below the first-best level, and in this case effort is below the first-best level: first-best effort is given by

$$e_{A1}^{FB} = \frac{\varphi_A + \zeta \varphi_A \varphi_B^2}{1 - \zeta^2 \varphi_A^2 \varphi_B^2}, \quad (28)$$

and second-best effort is given by (10). Hence $e_{A1}^* < e_{A1}^{FB}$ if

$$\begin{aligned} & \delta \gamma \alpha_{A1} (1 + \delta \gamma \zeta \varphi_B \alpha_{B1}) (1 - \zeta^2 \varphi_A^2 \varphi_B^2) \\ & < (\varphi_A^2 + \zeta \varphi_A \varphi_B^2) (1 - \delta^2 \gamma^2 \zeta^2 \varphi_A \varphi_B \alpha_{A1} \alpha_{B1}). \end{aligned} \quad (29)$$

Rearranging terms, this condition can be proved to be equivalent to the condition

$$\begin{aligned} & [(\varphi_A^2 + \varphi_B^2) \sigma^2 + \sigma_\varepsilon^2] [(1 + \zeta \varphi_B^2) (\varphi_A^2 + \varphi_B^2) - \delta \gamma \varphi_A (1 - \zeta^2 \varphi_A^2 \varphi_B^2)] \sigma^2 \\ & + (1 + \zeta \varphi_B^2) \sigma_\varepsilon^2 - \delta^2 \gamma^2 \zeta \varphi_A \varphi_B^3 (1 + \zeta \varphi_A^2) \sigma^4 > 0. \end{aligned} \quad (30)$$

This condition is satisfied if $\gamma \leq \frac{1}{2}$, and $\gamma \leq \frac{1}{2}$ must be satisfied, because otherwise the firm would have negative profits.

Finally, $S_1^c(\frac{1}{2})$ is not bounded above. This is proved by noting that, according to (10), when ζ increases, effort levels increase monotonically and with no upper bound. Hence there is a high enough value of ζ for which $S_1^c(\frac{1}{2}) > S_1^c(1)$. This completes the proof of the proposition.

A.5 PROOF OF PROPOSITION 4

First, given the rates of learning λ and l , the prior means and variances in period 0 are $\theta_{A0} = (1 + \lambda)^T \theta_0$, $\theta_{B0} = (1 + \lambda) \theta_0$, $\sigma_A^2 = \sigma^2 / (1 + l)^T$, and $\sigma_B^2 = \sigma^2 / (1 + l)$.

Second, since there are no complementarities in the production function, there are two possible optimal allocations of power: $\varphi_A^* = 1$ and $\varphi_B^* = 1$.

Let e_{A1}^* be the effort level chosen by A when $\varphi_A = 1$, and let e_{B1}^* be the effort level chosen by B when $\varphi_B = 1$. Furthermore, define $g(e) = e - e^2/2$. Then

$$e_{A1}^* = \delta \gamma \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2 (1 + l)^T}, \quad (31)$$

$$e_{B1}^* = \delta \gamma \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2 (1 + l)}, \quad (32)$$

and

$$S_1(0) = \theta_{B0} + g(e_{B1}^*), \quad (33)$$

$$S_1(1) = \theta_{A0} + g(e_{A1}^*). \quad (34)$$

Therefore $S_1(1) > S_1(0)$ if

$$\theta_{A0} - \theta_{B0} > g(e_{B1}^*) - g(e_{A1}^*). \quad (35)$$

Third, the left-hand side of this inequality does not depend on l and is increasing in λ .

Fourth, the right-hand side does not depend on λ , and we can show that it is increasing in l if $l < \tau^{1/(\tau-1)} - 1$. Differentiating the right-hand side with respect to l ,

$$g'(e_{B1}^*) \frac{\partial e_{B1}^*}{\partial l} - g'(e_{A1}^*) \frac{\partial e_{A1}^*}{\partial l}. \quad (36)$$

Since $e_{A1}^* < e_{B1}^*$, we have $g'(e_{A1}^*) > g'(e_{B1}^*) > 0$. Moreover,

$$\begin{aligned} \left| \frac{\partial e_{A1}^*}{\partial l} \right| > \left| \frac{\partial e_{B1}^*}{\partial l} \right| &\Leftrightarrow \left(\frac{\sigma^2 + (1+l)\sigma_\varepsilon^2}{\sigma^2 + (1+l)^\tau \sigma_\varepsilon^2} \right)^2 > \frac{1}{\tau(1+l)^{\tau-1}} \\ &\Leftrightarrow \frac{\sigma^2 + (1+l)\sigma_\varepsilon^2}{\sigma^2 + (1+l)^\tau \sigma_\varepsilon^2} > \frac{1}{\sqrt{\tau}(1+l)^{(\tau-1)/2}} \\ &\Leftrightarrow [\sqrt{\tau}(1+l)^{(\tau-1)/2} - 1]\sigma^2 \\ &\quad + [\sqrt{\tau}(1+l)^{(\tau+1)/2} - (1+l)^\tau]\sigma_\varepsilon^2 > 0, \end{aligned} \quad (37)$$

for which a sufficient condition is $l < \tau^{1/(\tau-1)} - 1$. If this condition is satisfied, expression (37) is positive, and the right-hand side of (35) is increasing in l . This completes the proof of Proposition 4.

A.6 PROOF OF PROPOSITION 5

Let $\widehat{S}_1(\varphi_A)$ denote the total surplus in period 1. Using the incentive-compatibility constraints,

$$\begin{aligned} \widehat{S}_1(\varphi_A) &= \delta\gamma \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \left(\varphi_A^{3/2} + \varphi_B^{3/2} - \frac{\delta\gamma}{2} (\varphi_A^2 + \varphi_B^2) \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \right) \\ &\quad + (\sqrt{\varphi_A} + \sqrt{\varphi_B})\eta_0 \end{aligned} \quad (38)$$

for $\varphi_A \in (0, 1)$, and

$$\widehat{S}_1(1) = \delta\gamma \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \left(1 - \frac{\delta\gamma}{2} \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \right) + \eta_0. \quad (39)$$

Maximizing $\widehat{S}_1(\varphi_A)$ with respect to φ_A , we find the following first-order condition:

$$\begin{aligned} & (\varphi_A^{-1/2} - \varphi_B^{-1/2})\eta_0 \\ &= \delta\gamma \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \left(\delta\gamma \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} (\varphi_A - \varphi_B) - \frac{3}{2} (\varphi_A^{1/2} - \varphi_B^{1/2}) \right). \end{aligned} \quad (40)$$

The analysis of this first-order condition delivers the following results. First, both sides of (40) are strictly decreasing in φ_A in the interval $[0, 1]$ (shown by differentiation). Second, both sides are equal to zero at $\varphi_A = \frac{1}{2}$. Third, the left side tends to plus infinity as φ_A approaches 0 and tends to minus infinity as φ_A approaches 1. Fourth, the right side tends to finite numbers as φ_A approaches 0 or 1. Fifth, the slopes of the left and the right sides increase, in absolute value, as φ_A moves away from $\frac{1}{2}$. Finally, as φ_A moves away from $\frac{1}{2}$, the (absolute value of the) slope of the left side increases more than the (absolute value of the) slope of the right side.

Proposition 5 follows from these results. First, it is never optimal to choose $\varphi_A \in \{0, 1\}$, because \widehat{S}_{φ_A} is increasing in φ_A at $\varphi_A = 0$ and decreasing in φ_A at $\varphi_A = 1$. The optimal allocation of power is always an interior solution.

Second, the second-order condition for a local maximum is satisfied at $\frac{1}{2}$ if

$$\delta\gamma \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \left(\frac{3\sqrt{2}}{2} - 2\delta\gamma \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \right) < 2\sqrt{2}\eta_0, \quad (41)$$

i.e., at $\varphi_A = \frac{1}{2}$ the left side of (40) is steeper than the right side of it. Furthermore, in that case $\varphi_A = \frac{1}{2}$ is the only local maximum. This is so because, as we move φ_A away from $\frac{1}{2}$, the slope of the left side of (40) increases more than the slope of the right side of (40).

Third, the second order condition for a local *minimum* is satisfied at $\varphi_A = \frac{1}{2}$ if at that point the left side of (40) is flatter than the right side of (40), i.e., condition (41) is not satisfied. Furthermore, in that case there will be two and only two local maxima. This is so because, as we move φ_A away from $\frac{1}{2}$, the slope of the left side of (40) increases more than the slope of the right side. Hence the left and right sides of (40) cross again at two points other than $\frac{1}{2}$ (at $0 < \varphi_A < \frac{1}{2}$ and at $\frac{1}{2} < \varphi_A < 1$). These two crossings correspond to local maxima.

Finally, the left side of (41) is increasing in σ^2 and decreasing in σ_ε^2 , and the right side is independent of these variables. This completes the proof of Proposition 5.

REFERENCES

- Aghion, P. and J. Tirole, 1997, "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105, 1–29.
- Auriol, E., G. Friebe, and L. Pechlivanos, 2002, "Career Concerns in Teams," *Journal of Labor Economics*, v. 20, n. 2, pp. 289–307.
- Baker, G., M. Gibbs, and B. Holmstrom, 1994, "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics*, 109, 921–955.
- Bernhardt, D., 1995, "Strategic Promotion and Compensation," *Review of Economic Studies*, 67, 315–339.
- Carmichael, L., 1988, "Incentives in Academics: Why Is There Tenure?" *Journal of Political Economy*, 96, 453–472.
- Cerasi, V. and S. Daltung, 2000, "The Optimal Size of a Bank: Costs and Benefits of Diversification," *European Economic Review*, 44, 1701–1726.
- DeGroot, M., 1970, *Optimal Statistical Decisions*, New York: McGraw-Hill.
- Doeringer, P. and M. Piore, 1971, *Internal Labor Markets and Manpower Analysis*, Lexington, MA: D. C. Heath.
- French, J. and B. Raven, 1959, "The Bases of Social Power," in D. Cartwright, ed., *Studies in Social Power*, Ann Arbor, MI: University of Michigan, Institute for Social Research.
- Friebe, G. and M. Raith, 2001, "Abuse of Authority and Hierarchical Communication," Mimeo, SITE and University of Chicago GSB.
- Gibbons, R. and K. Murphy, 1992, "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 100, 468–505.
- Grossman, S. and O. Hart, 1986, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691–719.
- Hart, O. and J. Moore, 1990, "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98, 1119–1158.
- Holmstrom, B., 1982a, "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 324–340.
- , 1982b, "Managerial Incentive Problems: A Dynamic Perspective," in *Essays in Economics and Management in Honor of Lars Wahlbeck*, Helsinki, Finland: Swedish School of Economics. Reprinted: *Review of Economic Studies*, 96 (1999), 169–182.
- Huson, M., R. Parrino, and L. Starks, 2001, "Internal Monitoring Mechanisms and CEO Turnover: A Long-Term Perspective," *Journal of Finance*, 56, 2265–2297.
- Jeon, S., 1996, "Moral Hazard and Reputational Concerns in Teams: Implications for Organizational Choice," *International Journal of Industrial Organization*, 14, 297–315.
- Meyer, M., 1991, "Learning from Coarse Information: Biased Contests and Career Profiles," *Review of Economic Studies*, 58, 15–41.
- , 1994, "The Dynamics of Learning with Team Production: Implications for Task Assignment," *Quarterly Journal of Economics*, 109, 1157–1184.
- Mintzberg, H., 1979, *The Structuring of Organizations. A Synthesis of the Research*, Englewood Cliffs, NJ: Prentice-Hall.
- Ortega, J., 1998, "Essays on the Economics of Authority and Bureaucracy," Ph.D. Thesis, Massachusetts Institute of Technology.
- Parrino, R., 1997, "CEO Turnover and Outside Succession. A Cross-Sectional Analysis," *Journal of Financial Economics*, 46, 165–197.
- Prendergast, C., 1995, "A Theory of Responsibility in Organizations," *Journal of Labor Economics*, 13, 387–400.
- Rajan, R. and L. Zingales, 1998, "Power in a Theory of the Firm," *The Quarterly Journal of Economics*, 113, 387–432.
- Ricart-i-Costa, J., 1988, "Managerial Task Assignments and Promotions," *Econometrica*, 56, 449–466.

- Rotemberg, J., 1993, "Power in Profit-Maximizing Organizations," *Journal of Economics & Management Strategy*, 2, 165–198.
- Sauer, R., 1988, "Estimates of the Return to Quality and Coauthorship in Economic Academia," *Journal of Political Economy*, 96, 855–866.
- Seltzer, A. and D. Merrett, 2000, "Personnel Policies at the Union Bank of Australia: Evidence from the 1888–1900 Entry Cohorts," *Journal of Labor Economics*, 18, 573–613.
- Waldman, M., 1984, "Job Assignments, Signalling, and Efficiency," *Rand Journal of Economics*, 2, 255–267.