Authority in Organizations: A Survey*

Patrick Bolton
Columbia University

Mathias Dewatripont
ECARES, Université Libre de Bruxelles

January 2011

Abstract

We survey the contract theory literature on the notion of authority inside firms. This literature is concerned with managers' role inside firms in coordinating production activities involving multiple agents. The exercise of authority by a manager involves at least four related decision rights: i) the power to initiate projects and direct subordinates to take certain actions; ii) the power to exact obedience; iii) the power to ratify and enforce agreements, and iv) the duty to monitor subordinates and the ability to reward them for good perfomance. Some of the central questions concerning authority inside firms are: i) how authority should be allocated among managers; ii) how to define lines of authority; iii) how to circumscribe authority of any given manager, and; iv) how to credibly delegate authority given that ultimately control rights rest with the owners of the firm.

*We are grateful to the editors, Robert Gibbons and John Roberts, for their comments.

1 Introduction

Much of the theory of the firm has been concerned with what determines the boundaries of firms. When does a firm reach its efficient size? When does a merger between two separate entities improve efficiency of production? Which core activities should a firm focus on, and which should it outsource? A central achievement of the property rights theory of the firm (Grossman-Hart-Moore) is that these questions can all be framed in terms of an allocation problem of control rights or property rights over productive assets. That is, all questions relating to the boundaries of the firm can be phrased in terms of who should own which productive asset.

As powerful as this framework is for analyzing the boundaries of firms, it has little to say on what actually goes on inside firms and how firms should be organized internally. When it comes to understanding what managers do it is not very helpful to reduce firms to just a collection of productive assets and a set of owners. One needs to step back from this extreme abstraction and introduce elements from Coase (1937), Simon (1951) and Williamson's (1975) transactions-cost theory of the firm. In his foundational article Coase (1937) envisions firms with internal transactions processes that are different from market processes. On page 392, he describes firms as relying on *authority procedures* in completing production and distribution of goods rather than *negotiation and exchange* as in a market:

> ...the operation of a market costs something and by forming an organisation and allowing some authority (an "entrepreneur") to direct the resources, certain marketing costs are saved.

Thus, what is special about internal transactions in firms is that they are based on a different mechanism which we shall refer to in general terms as *authority*. This mechanism is a central and integral part of the process of production and distribution of goods. The exercise of authority broadly defined is what managers do. As we detail in the definition below, managers' day-to-day job is to execute the broad mandate that is handed to their team by their superiors, to coordinate the execution of this mandate by their subordinates, and to keep tabs on what their subordinates are doing. Although the notion of authority is quite broad and complex it is interesting to observe that almost all of its multiple facets are relevant to firms, how they are organized and how they are run. Following Fama and Jensen (1983), let us define four main different aspects of *authority*:

A] Authority is a supervisor's *power to initiate projects and direct* subordinates to take certain actions.

B] Authority also involves the concomitant power *to exact obedience*. For if the supervisor was unable to get his subordinates to execute the directed actions and had to coax a negotiated agreement from the subordinate each time an action is required there would not be much point in granting a manager the power to direct subordinates.

C] Another decision right associated with managerial authority is *the power to ratify and approve actions* in a pre-determined area of competency. Formal authority in a firm rests with the firm's owners, who have the ultimate right to make decisions and to delegate authority to managers. In owner-managed firms there is full centralization of decision-making authority with the owner(s). In larger firms authority is delegated to a board of directors, and in turn to a professional management team and managerial hierarchy. However, this delegation of authority only stands to if managers' power to

ratify and approve action has been previously delineated.

D] Finally, authority also involves the manager's duty to monitor subordinates and her ability to reward them for good performance.

A central question for the theory of the internal organization of firms is how to allocate authority among its members. How to define lines of authority? How to circumscribe authority of any given agent? How to credibly delegate authority and how to ensure that the exercise of authority is respected? A fundamental theme uniting all the contributions surveyed in this chapter is that the role of authority and its separation from ownership rights arises from the presence of information processing and communication costs. The firm owner's inability to attend to all the decisions involving the operation of the firm, his inability to master all the relevant information to take decisions, and his limited capacity to communicate with all the employees, are the main reasons why authority is separated from ownership. These same reasons are also the source of Coase's *marketing costs,* which authority procedures inside firms seek to economize.

In this chapter we survey the recent economics literature on the internal organization of firms that deals with authority inside firms. This literature builds on the foundational works of Max Weber, Herbert Simon, Alfred Chandler, Anthony Downs, Oliver Williamson among others but seeks to frame some of the core ideas of this literature in the language and formalism of modern contract theory. In the process, this literature offers new perspectives and sharper intuitions on the delegation of authority in organizations. Our goal in this survey is to provide a unified presentation of this contracting literature. A central question this literature is concerned with is how the ultimate holders of formal authority, the owners of the firm, are able to credibly delegate authority to managers further down the organization. A related

question is how authority is optimally allocated inside the firm. To address these questions, it is helpful to begin with a taxonomy of contract-theoretic models, which will allow us to discuss the outline of this survey.

2 A taxonomy of contract models

As a way of organizing the literature it is helpful to distinguish between three different categories of contract models, that differ in the degree of enforceability of contractual agreements. The first, classical, category assumes that essentially every transaction, service or trade that is observable is enforceable. In particular, in the classical contract theory paradigm it is assumed that agreements on both *ex ante* and *ex post action choices* are *enforceable.* Under these broad enforceability assumptions the contracting parties are able to write an optimal comprehensive contract that fully determines how the game between the parties is played. There is then no role for an *authority* mode, as–so to speak–the contracting parties are then simply executing 'the prescribed actions by the *mechanism designer*' under the optimal contract.

A second category of contract models allows for weaker enforceability assumptions whereby some *ex-ante* agreements on *future action choices* are not enforceable. *Ex-post agreements on ex–post action choices,* however, continue to be enforceable. The justification for these assumptions that is commonly given is that while agreements on action choices are easy to describe ex post, they are impossible to describe ex ante. The *incomplete contracts* literature following Grossman and Hart (1986) and Hart and Moore (1990) is the prime example of this category of models. It involves only *partial contracting*–to use the terminology of Aghion, Dewatripont and Rey (2002)–in the sense that the contracts that are *superimposed* on the underlying game between the contracting parties do not fully determine how the game is played. Some

actions typically remain to be determined ex post. This category of contract models introduces the notion of *ownership* as *residual rights of control*, by defining a right for the owner to determine ex post all the actions that have been left unspecified in the ex-ante contract. This is an important advance towards a formalization of *authority*, but as we shall argue, the assumptions on contractual enforcement must still be further weakened to be able to fully formalize the notion of authority. This framework (where only agreements on future action choices are not enforceable) opens the way for a theory of the optimal allocation of *ex post bargaining power.* However, it does not lead to a theory of *authority* over actions.

To be able to model the concept of *authority* the assumptions on contractual enforcement must be weakened still further to allow for the non-enforceability of both *ex-ante and ex-post* agreements on action choice. This third category of contract models can be subdivided further into two subcategories: models where *allocations of control are enforceable* even though agreements over action choice are not, and models where neither agreements over action choice nor allocations of control rights are enforceable.

The category of models assuming contractable-control focus on the *decision rights* facets of authority (parts A and C in our definition) and allow for the formalization of three different notions of *authority,* which we refer to as *horizontal, vertical* and *contingent authority.* Section 3 will address *horizontal authority*, which deals with issues relating to the delineation and assignment of tasks or missions to individual agents; Section 4 will focus on *vertical authority*, which is concerned with the definition of lines of authority, precedence and responsibility; and, Section 5 will deal with *contingent* authority, which addresses the rationale for and merits of temporary assignments of authority over specific issues or projects.

After reviewing these contractable control models, we turn in Section 6 to models with non-contractable control, which focus on the *'interpersonal'* facets of authority (parts B and D in our definition). In these models the agent has effective control over actions (because of the principal's *limited attention*) but the principal will still want to align the agent's objectives through some incentive scheme. He may still attempt to 'get his way' by incentivizing agents to do what he wants (as in classical moral hazard models), by partially *monitoring* agent's actions (as in efficiency wage models), and through *relational contracts*, or '*persuasion*'. We follow Van den Steen's (2007, 2010) terminology and refer to these models as models of *interpersonal authority*.

Finally, section 7 offers concluding comments and suggestions for future research.

3 Horizontal allocation of authority: multi-task perspectives

While the early contracts literature on the economics of organization based on the *principal-agent* framework by and large simplified the agency problem by assuming that agents are only responsible for a single task (*'hidden effort'*), contributions of the last decade have recognized the need for a broader perspective. Managers, workers and bureaucrats not only need to be motivated to work, but also must be incentivized to optimally allocate their attention among several different activities they are responsible for. Their allocation of attention to the different tasks is determined both by the relative (intrinsic or financial) benefits they derive from the various tasks and by the technological complementarity or substitutability across tasks. This section discusses how the multi-task principal-agent framework can shed light on the

organizational problem of horizontal allocation of authority.[1]

3.1 A framework

Consider the following multitask incentive problem where π_i denotes the level of output on task $i = 1, ..., m$. The distribution of output can be altered through the agent's allocation of effort (or attention) across tasks. The principal has access to a population of identical agents who can exert effort on the various tasks. The principal is interested in maximizing the expectation of an output function ψ, which depends on the vector of statistics π, minus the wage costs of the agents he hires. A special case which we shall refer to for most of the discussion below is the functional form $\psi(\pi) = \sum_{i=1}^{n} \pi_i$. In reduced form, we can define $\Psi(a) = E[\psi(\pi) \mid a]$,where a is the constellation of the agents' effort vectors. For an individual agent, an effort vector $a = (a_1, a_2, ..., a_n)$ has utility cost $c(a)$. An individual agent's objective function is a (linear or concave) separable function U of her wage w, minus her effort cost $c(a)$.

The first-best efficient outcome is achieved if effort a is directly contractable. It is obtained by maximizing the principal's gross payoff $\Psi(a)$ minus the agents' wage bill, which is just sufficient to induce them to accept the employment contract of the principal (participation constraint). The first-best cost of obtaining effort a from an agent is given by the function $U^{-1}(U^* + c(a))$, where U^* is the agents' reservation payoff. Whether it is optimal to rely on one or several agents depends on the shape of this function, which is affected by the level of the reservation utility and the curvature and economies of scale and scope of the cost of effort function $c(.)$.

Non-contractability of effort however means adding to the maximization

[1] See Dewatripont et al. (2000) for additional details.

problem of the principal an incentive constraint on top of the participation constraint. This incentive constraint takes into account the impact of agents' efforts on the level of the contractable variables and on wage levels. What can be contracted upon is only the outcome $F(\pi_1, \pi_2, ..., \pi_m)$, which can be a scalar or a vector. In general, the dimensionality of the contracting problem an organization faces depends on both the number of tasks n and the number of outputs m. The traditional single-task principal-agent problem with output as the only performance measure involves $n = 1$ and $F(\pi_1) = \pi_1$. In this section, we survey a number of approaches on horizontal allocation of authority for the multitask case ($n > 1$). The constrained optimization problem for the multitask case then takes the following form:

$$\max_{a, w(F(\pi_1,\pi_2,...,\pi_m))} \quad \Psi(a) - E(w(F(\pi_1, \pi_2, ..., \pi_m)) \mid a)$$

subject to:

$$E(U(w(F(\pi_1, \pi_2, ..., \pi_m)) \mid a) - c(a) \geq U^*$$

and:

$$a \in \arg\max_{\widetilde{a}} \quad E(U(w(F(\pi_1, \pi_2, ..., \pi_m))) \mid \widetilde{a}) - c(\widetilde{a})$$

3.2 Multiple tasks and effort substitution

Focusing on the two-task case for simplicity, the mainstream multitask problem is based on the observation that encouraging effort on one task may crowd out the effort on the other task when the cross-partial derivative of the effort cost function is strictly positive ($\partial^2 c / \partial a_1 \partial a_2 > 0$). A major theme from this literature is that when a single agent has authority over multiple tasks then the incentive scheme must be designed to *balance* incentives across tasks. In a moral hazard context, if, for example, the performance on task 1 is measured more precisely than that on task 2 ($\pi_1 = a_1 + \varepsilon_1$ and $\pi_2 = a_2 + \varepsilon_2$, where ε_2 is noisier than ε_1), it would seem natural that the incentives be more high-powered on the first task (that is, $dw/d\pi_1 > dw/d\pi_2$), for which

the agent is exposed to low compensation risk. Yet, if $c(a_1, a_2) = c(a_1 + a_2)$ for example, the incentive powers on the two tasks cannot differ without the agent neglecting completely the less rewarded task.[2]

The cost of balancing incentives across tasks is thus that overall the agent is given low-powered incentives even on tasks with well-measured outputs (Holmström and Milgrom, 1991)[3]. This will lead the organization: (i) either to go for a job design which gives authority to agents for tasks that are equally easy or hard to measure (with high-powered incentive schemes given only when a precise measurement is possible); (ii) or to assign authority over the two tasks to two different agents. Indeed, just as a profit-maximizing multiproduct firm must take into account that possible cannibalizing effects of a reduction in price of one product on the sales of other products, a principal must factor in the fact that offering stronger performance incentives along some dimension, while inducing the agent to spend more time on the corresponding task, may also divert attention away from the other tasks.[4]

[2]That is, $dw/d\pi_1 > dw/d\pi_2$ will imply $a_2 = 0$.

[3]Holmström and Milgrom (1991) also discuss incentive pay for teachers: since only a subset of the skills taught to pupils can be measured (e.g. standardized math or language scores), a cost from introducing explicit incentive pay for teachers is that it can lead them to neglect the 'more general education' task in order to train pupils exclusively for the purpose of doing well at the tests.

[4]Other interesting implications from this framework in terms of job design include the following topics: (i) *Helping coworkers in organizations* (Itoh 1991; Auriol et al., 2002): Managers facing high-powered incentives on their own performance will probably spend less time on helping their fellow managers. In academia, individual research incentives may lead to reduced effort in teaching and/or interaction with students and colleagues. (ii) *Multiple principals and exclusivity* (see for example Martimort, 1996). In a situation with multiple principals, effort substitution induces principals to compete for the agent's attention. This leads to incentive schemes that are too high-powered and may justify exclusivity (the agent being forced to deal with a single principal).

3.3 Conflicts between tasks

The optimality of task separation is even clearer in the presence of direct conflicts between tasks. Conflicts between two tasks arise for example when $F(\pi_1, \pi_2)$ is single-dimensional and depends positively on π_1 but negatively on π_2. In this case, it is clearly hard to induce the agent to increase both π_1 and π_2 simply by providing wage incentives on the basis of $F(\pi_1, \pi_2)$. This situation arises in the following circumstances:

- Assume for example as in Dewatripont and Tirole (1999) that π_1 and π_2 represent the amount of verifiable evidence obtained, respectively, on the 'pros' and the 'cons' of a given decision. Efforts spent looking for pros and cons increase, respectively, π_1 and π_2. The decision $F(\pi_1, \pi_2)$ will then depend positively on π_1 and negatively on π_2, while ex ante the principal cares about obtaining the best possible evidence on both sides, so $\Psi(a_1, a_2)$ is increasing. Such an adverserial setup is to be expected in any situation where the firm faces a major decision that is not clear-cut, such as a merger with another firm, or the launch of a new product, and where the decision is improved if it is based on the best available information for or against the decision. As Dewatripont and Tirole (1999) have shown, the principal may then benefit by deliberating with two managers, each with an opposing stake in the decision.

- Another application involves the tasks of ex ante and ex post monitoring of agents in the organization: managers who have to perform both tasks will face a dilemma: uncovering shirking or a faulty product only ex post reveals that the manager has failed to spot the problem ex ante. The manager may then be conflicted ex post in revealing the problem. Separation of ex-ante and ex-post monitoring can then help

avoid 'cover-ups' (see e.g., Boot et al., 1993; and Dewatripont and Tirole, 1994a).[5] Interestingly, such a separation has been implemented in the area of banking regulation in the United States: the OCC (Office of the Comptroller of the Currency) is in charge of the ex-ante monitoring of banks, while the FDIC (Federal Deposit Insurance Corporation) is in charge of ex-post intervention in case the bank is in trouble.

3.4 Implicit incentives and missions

Multitask problems giving rise to a horizontal authority allocation problem are also present when there are no (or not only) explicit incentives in place to motivate managers, such as when managers are mainly motivated by *career concerns* or their *reputation* as in Holmström (1982b). In the standard career concerns model the managerial labor market tries to infer a manager's talent from his past performance. The manager's incentives to expend effort on one task then depend on how the market infers the manager's talent based on the manager's performance in undertaking that task. Such 'career concerns' can be captured in our framework by assuming that $F(\pi_1, \pi_2, ..., \pi_m) = (\pi_1, \pi_2, ..., \pi_m)$, and that the strength of the reward connecting $F(\pi_1, \pi_2, ..., \pi_m)$ to wages for an agent is not contractually chosen but is determined by the labor market. The agent's effort level in a rational-expectation equilibrium is then determined by

$$a \in \arg\max_{\widetilde{a}} E(U(w_0 + w(F(\pi_1, \pi_2, ..., \pi_m), a^*)) \mid \widetilde{a}) - c(\widetilde{a}) \text{ and } \ a = a^*.$$

In the simplest version of this problem, π_i depends positively on the agent's effort a_i on task i and on her talent θ (which neither the agent nor the

[5] Of course, separation of tasks also involves duplication of effort and moral hazard in teams à la Holmström (1982a).

market knows a priori) and on noise. Upon observing the vector of performances $(\pi_1, \pi_2, ..., \pi_m)$, the market tries to infer the talent parameter θ given the expectation of effort a^*, which is assumed to be correct in equilibrium. In a later period, the agent is then rewarded, through promotion, new job offers, etc, according to her perceived talent, which is embodied in the formulation $w(F(\pi_1, \pi_2, ..., \pi_m), a^*)$. Note that the agent's incentive to work here is driven solely by her desire to influence the market's perception of her talent, and her current-period wage w_0 is by assumption constant. In equilibrium, she does not succeed in 'fooling' the market, however she is still driven to perform to avoid 'disappointing' the market: she expends effort ex post because the market expects it. Whether she ex ante benefits from this incentive depends on whether the initial wage w_0 reflects the market's expectations about equilibrium effort.

Dewatripont et al. (1999) study how equilibrium effort depends on the number of tasks the agent has authority over. For example, they establish conditions under which reducing the number of tasks n the agent has authority over increases her total effort $(a_1 + a_2 + ... + a_n)$ and the principal's total profit. The reduction in tasks tends to increase overall performance whenever the link between performance and perceived talent is enhanced with fewer tasks, due to less noisy performance measures (higher signal-to-noise ratio). Note that lowering n has more drastic consequences here than in the standard multitask problem with explicit incentives: it is not a matter of less effort substitution but of *an increase in total effort*!

This benefit from 'focus' is consistent with Wilson's (1989) celebrated study of bureaucracy, where he argues that successful Government agencies are the ones that manage to concentrate on reasonably focused 'missions'. Focus means that 'accountability' improves because performance is easier

to evaluate. The agency in the end benefits from this situation if better accountability translates into enhanced autonomy.

Wilson's book is replete with examples of processes by which bureaucracies have transformed their grand-but-vague objectives into more specific and more operational missions. For example, the Department of Agriculture's mission of 'feeding the Nation' has been transformed into mainly helping farmers, and the Department of Transportation's mission of 'improving the safety of transportation' has traditionally been confined to improving car safety.

The three approaches detailed above all point to the cost of multiple tasks and the advantage of splitting them between several agents. They also imply predictions concerning optimal task clustering when each agent has to be allocated several tasks: *i)* The traditional multitask approach suggests that agents have to be given authority over tasks that have similar degrees of measurability. Agents allocated easily measurable tasks can be given more powerful financial incentives, while others must find less powerful incentives. *ii)* The conflicting-task approach argues for avoiding 'conflicts of interest' in job design. This will lead to specialization of management in pursuing narrow goals and reliance on competition between specialized managers to ensure overall balance. *iii)* Finally, the focused-mission approach pleads for 'functional specialization'. That is, for specializing agents in tasks that require similar talents, so as to keep the inference process between overall performance and talent strong.

4 Vertical allocation of authority

As Alfred Chandler has emphasized, most large firms are run by *professional managers*, and shareholders–the ultimate owners of the firm–have relatively

little say in firms' day-to-day operations or in firms' strategic decisions. In other words, in most larger firms the owners *delegate authority* to the CEO and the management team. The CEO in turn delegates authority further down the hierarchy to divisional managers, product managers, etc. A central question for the theory of organizations is how can owners credibly delegate authority to managers? In other words, how can owners transfer residual rights of control to non-owners? Can they formally transfer authority without transferring ownership rights?

One basic difficulty for owners in delegating authority to managers is that it is very hard to formally describe the manager's area of competency and duties to shareholders. Thus, it is tricky for owners to commit to delegating discretion to the CEO over some decisions such as, the firm's marketing strategy, outsourcing policy, product design, etc. A large owner, who has board representation can always meddle in the managerial decision process and it is difficult for the CEO to ignore such a shareholder if he has the power to fire the manager. Thus, for example, when there is a transfer of ownership of a newspaper or a television channel, to what extent can the new owner credibly delegate editorial policy to the managing editor? This is always a critical question when there is a change in ownership of a media company.

One answer to this question is that the transfer of authority can be made more credible if there are higher costs for the owner of firing the manager. There is then a tradeoff for the principal between more effective (or credible) delegation of authority and better incentives for the agent. One extreme way of delegating authority–as in academia or in the judiciary–is to grant *tenure* to the agent. Although there are undoubtedly incentive reasons for the tenure system, as Carmichael (1988), Waldman (1990), and Friebel and Raith (2004) have notably argued, another important reason for the system

is surely to be able to credibly delegate hiring decisions to departments and to guarantee academic freedom or the independence of the judiciary.

Another answer, which has received more attention in the literature, is that the credibility of delegated authority rests on the superior knowledge or information of the agent. In other words, if shareholders of large firms do not have the skills or knowledge to run a firm efficiently–which is likely to be the case when ownership is dispersed and no single owner has a strong incentive to become more knowledgeable about the firm's business–then the delegation of authority to a team of professional managers is credible.

In this Section we first review and discuss the fundamental contribution of Aghion and Tirole (1997) which has spawned a large literature. This article directly addresses the question of how a principal can use an agent's informational advantage to credibly commit to an *arm's length incentive scheme* or to credibly delegating authority to an agent. As we detail in subsection 4.1, Aghion and Tirole show how "real decision-making" depends on individual costs of information acquisition, the degree of congruence of the parties involved in the relationship and also the allocation of formal authority. We also discuss the credibility of the allocation of formal authority, with special emphasis on the case of academic contracts.

Aghion and Tirole explain how, in some cases, parties not endowed with formal authority can, when they enjoy an informational advantage, "have their way" by communicating to the other party what to do. In settings more general than theirs, this leads to the question of strategic information transmission a la Crawford-Sobel (1982). The desirability of delegation in such a context has been studied by Dessein (2002), and we discuss this work in subsection 4.2 as well as extensions in subsection 4.3, which focuses on the interplay between delegation and adaptation to an uncertain environment.

4.1 Formal and real authority

4.1.1 The Aghion-Tirole model

Aghion and Tirole (1997) focus on a noncontractable action, which in their setup amounts to "choosing one project" that the agent has to work on. There are initially $N \geq 3$ potential projects. Project $k \in \{1, 2, ..., N\}$ gives the principal a private benefit B_k and the agent a private benefit b_k. Initially, these various projects are indistinguishable from one another. At least one of them is sufficiently bad that choosing one project at random is worse for both parties than undertaking no project (with associated payoff normalized to zero in this case). On the other hand, the parties know in advance that the best project for the principal gives him $B > 0$ while it gives the agent $\beta b > 0$, and that the best project for the agent gives her $b > 0$ while it gives the principal $\alpha B > 0$. Call α and β the "congruence parameters", and assume that they are positive but smaller than one. The higher these parameters, the more congruent the preferences of the two parties. Note that partial congruence is built into the setup in any case, since one has assumed that: (i) the two parties agree that choosing a project at random is worse than undertaking no project; (ii) they each prefer to allow the other party to choose his/her favorite project rather than undertaking no project (since α and β are assumed to be positive).

These assumptions about congruence are crucial given the *information acquisition* technology: While both parties are initially uninformed about the private benefits associated with individual projects, they can each exert effort to improve their information. Specifically, the principal can at cost $\psi_P(E)$ (increasing and convex in E) become *fully informed* with probability E about project benefits, while he remains *fully uninformed* with probability $1 - E$. Similarly, the agent can at cost $\psi_A(e)$ (increasing and convex in e)

become fully informed with probability e about project benefits, while she remains fully uninformed with probability $1 - e$.

The timing of the game is as follows: In stage 1, the parties contract; in stage 2, they exert effort to acquire information about individual project payoffs; finally, in stage 3, a decision can be taken on which project to undertake, if any. Assuming that efforts are privately chosen and that parties only care about their private benefits[6], the contract only consists in an allocation of authority for stage 3. As Aghion and Tirole stress, what can be contractually allocated is solely "formal" authority, that is, who has the *right* to take the decision. This differs from "real" authority, that is, who *actually* takes the decision. Indeed, given the partial congruence built into the model, a party endowed with formal authority chooses to undertake a project only if he/she is informed about project benefits; otherwise, he/she *transfers* authority to the other party (or equivalently, asks the other party for a recommendation and follows it). In turn, this other party chooses a project/makes a recommendation only if he/she is informed about project benefits. Otherwise no project is undertaken.

Given this continuation equilibrium in stage 3, if the contract allocates formal authority to the principal, the payoffs for the principal and the agent (U_P and u_A respectively) upon choosing their effort levels are:

$$U_P = EB + (1 - E)e\alpha B - \psi_P(E)$$

$$u_A = E\beta b + (1 - E)eb - \psi_A(e).$$

These conditions reflect the fact that the principal chooses his favorite action whenever he is informed about individual project payoffs, while the agent's

[6] This is assumed for the sake of presentation and can be generalized (see Aghion and Tirole (1997)).

information matters only when the principal is uninformed. In stage 2, simultaneous effort choice leads to the following first-order conditions:

$$(1 - \alpha e)B = \psi'_P(E)$$

$$(1 - E)b = \psi'_A(e).$$

As indicated by the second first-order condition, higher effort E by the principal crowds out effort e by the agent, who understands that her effort matters with a lower probability. There may therefore be a gain for the principal to commit to exerting lower effort, for example by choosing an agent who is more congruent with him (that is, an agent "with a higher α")[7]. In this case indeed, the principal exerts less effort, as indicated by the first first-order condition: He is less worried about being uninformed, because the project chosen by the agent when she is the only one informed leads to a lower relative loss for the principal.

Another way for the principal to induce the agent to work harder is to *delegate* her formal authority on project choice. In this case, given the continuation equilibrium in stage 3, the payoffs for the principal and the agent upon choosing their effort levels are:

$$U_P = e\alpha B + (1 - e)EB - \psi_P(E)$$

$$u_A = eb + (1 - e)E\beta b - \psi_A(e).$$

The agent now chooses her preferred action whenever she is informed, while the principal's information matters only when the agent is uninformed. Simultaneous effort choice now leads to the following first-order conditions:

$$(1 - e)B = \psi'_P(E)$$

[7] Aghion and Tirole also look at increases in the "span of control" of the principal as a way for him to commit to spend less effort per agent, since he has more projects to acquire information about.

$$(1 - \beta E)b = \psi'_A(e).$$

A comparison of these first-order conditions with the ones where the principal has formal authority indicates that the agent exerts more effort and the principal exerts less effort under delegation. Indeed, the two effort levels are strategic substitutes, which reinforces the fact that the individual endowed with formal authority has more incentives to exert effort, *ceteris paribus*, since he/she can have the first go at choosing the action.

Using a very simple model, Aghion and Tirole are able to generate interesting predictions about the allocation of authority in organizations, (for example, as Burkart et al. (1997) argue, that shareholder dispersion acts as a commitment device to "empower" management). These predictions are in line with some prominent cases of internal organization described by Roberts (2004) such as BP's (or Johnson and Johnson's and ABB's) organizational strategy to empower front-line managers by lightly staffing its corporate headquarters. Using a similar model, Rotemberg and Saloner (1993) have also been able to generate simple predictions on what traits a CEO should have to be an effective leader in motivating agents under his authority. They show that a more *empathic* CEO (that is, a CEO that puts more weight on maximizing his subordinate's utility) may be able to induce greater incentives for his subordinates' to exert effort to come up with new ideas, even if this comes at the cost of implementing too many bad ideas ex post. In general, Rotemberg and Saloner show that a somewhat empathic CEO is optimal for the firm: he provides the agent with some incentive to exert effort, but also rejects the worst ideas she proposes.

4.1.2 Credible delegation?

While the key innovation behind the Aghion-Tirole model concerns the concept of *real authority*, it has also generated debate as far as *formal authority* is concerned. At some level, one can interpret delegating formal authority simply as the equivalent, in the Grossman-Hart-Moore approach, of *selling the underlying asset* necessary to get the project going. While this interpretation is straightforward, it is not that interesting if one focuses on authority *within organizations*. In such a case, delegation of formal authority is problematic because, at least in the US, the "business judgment rule" typically leads courts not to enforce contracts written between parties within a single firm. It thus seems that, as far as authority within organizations is concerned (e.g. between the firm and one of its divisions), one can choose the allocation of real authority but not of formal authority: In the end there is no credible way for the boss to commit not to "take back" formal authority whenever it is in his interest to do so. This can be circumvented in a repeated-game setting, where allocation of formal authority does matter for efficiency, as has been analyzed by Baker et al. (1999) (see also section 6.2 of this chapter), but these arguments do not work in a one-shot setting.

Note, however, that there are exceptions to this business judgment rule. First of all, it does not apply with financial investors: as has been discussed in the financial contracting field, there are enforceable allocations of formal authority among the various investors that have put money in the firm. Another exception, already mentioned in the introduction of this section, concerns some employment contracts: as academics, we are well aware of contracts that grant us tenure but also *academic freedom*. Aghion et al. (2008) have built on this observation to discuss the relative merits of academia and the private sector in the innovation process. In fact, as a modeling

strategy, they keep at a minimum the differences between academia and private-sector, profit-maximizing firms, focusing solely on the allocation of formal authority between the researcher and the organization.

They consider an economically viable product (e.g. a new drug) which starts with an idea I_0, that can be built on by researchers, leading to idea $I_1, I_2, ...,$ until idea I_k which generates economic value V. Assume that, for each of the k stages, one researcher can work on the idea (this can be generalized to more than one researcher). In academia, the researcher is free to pursue his own strategy. He can therefore choose the "practical strategy", which yields a probability p of being successful, that is, of moving to the next stage. However, he can also choose the "alternative strategy", which yields a probability 0 of moving to the next stage.

Why would the researcher go for the alternative strategy? Because first (and this "makes it hard" for academia to emerge as an efficient organization), one assumes that financial incentives are not doable, and only fixed wages are possible; and second because, while the alternative strategy involves zero disutility for the researcher, the same is not true for the practical strategy: with probability α, it does generate zero disutility, but with probability $1-\alpha$, the researcher will have a disutility z from the practical strategy. Moreover, at the time of being hired, the researcher does not know what his preferences will be.

Consequently, when the researcher is protected by academic freedom, that is, can choose his research strategy, he will only pursue the practical strategy with probability α (this also "makes it hard" for academia to emerge as an efficient organization: in reality, the alternative strategy will not have zero value). Instead, in the private sector, the researcher's boss can direct his research, that is, impose the practical strategy. This raises the probability

of success, from αp to p, which is the benefit of private research. However, there is also a cost: since having a boss imposes an ex-ante disutility of $(1-\alpha)z$ on researchers, and if researchers have an outside option whose value is R, equilibrium academic wages will be R and private-research wages will be $R+(1-\alpha)z$. This is in fact consistent with evidence presented by Stern (2004) on academic versus private-sector wage offers given to the same individuals.

While the cost and benefit of academia are intuitive, they generate an interesting prediction in the context of a sequential innovation process. Indeed, in the last stage of research, that is, when one already has idea I_{k-1}, the private sector is the more efficient way of organizing research if and only if:

$$\pi(k) = pV - R - (1-\alpha)z > \alpha pV - R.$$

Assume this first condition is satisfied. Then, in the next-to-last stage, the private sector is the more efficient way of organizing research if and only if:

$$p\pi(k) - R - (1-\alpha)z > \alpha p\pi(k) - R.$$

Interestingly, since $\pi(k) < V$, this second condition is harder to satisfy than the first one. Intuitively, the idea is that, if as assumed the wage difference between academia and the private sector is (roughly) constant across stages, then academia looks worse and worse when one moves downstream in the innovation process, because the cost of a lower probability of moving to the next stage represents a bigger and bigger financial stake. As the two conditions above indicate, in the last stage the stake is V but in the next-to-last stage, the stake is only $\pi(k)$, the value of moving to the last stage. Indeed, the main result is that the efficient sequence will have academia perform the earlier stages and the private sector the later stages of the innovation process.

And the efficient transition from academia to the private sector occurs earlier the lower the level of α or z and the higher the level of V. The model therefore generates academia as an endogenously efficient institution even if the assumptions make academic freedom costly, by not allowing (explicit or implicit) incentive schemes and by assuming away any value for the alternative strategy except in terms of the researcher's disutility.

As shown by Aghion et al. (2008), real authority can also be brought into the picture: if in order to exert her authority the boss of a private firm needs to expend effort to be informed about the specifics of the practical strategy, it is easy to see that the return to such effort, and therefore its intensity, will grow as research progresses. The model therefore generates the prediction that the private firm becomes "bossier", and pays higher wages, the closer one gets to the final product.

4.2 Strategic information transmission and delegation

The Aghion-Tirole model stresses the importance of information as the source of *real* authority and casts the problem of optimal allocation of *formal* authority in terms of an incentive problem to obtain information. An extension of this model by Dessein (2005) shows that relinquishing formal authority may be a way for the principal to *signal congruence* with the agent and therefore to increase his real authority. This can make sense in the setting considered by Dessein where there is a probability that the principal, even when keeping formal authority and being informed about his favorite project, cannot 'get his way' because his control over the agent is imperfect. In such a case, the agent's behavior when uninformed will depend on her belief about the congruence between the two parties. Relinquishing formal authority can then be a credible way for the principal to signal such congruence, since it is

less costly for him to do so if the parties are actually more congruent. Dessein argues that this setting can help explain how venture-capital contracts are structured and how formal authority is allocated between the venture capitalist and entrepreneur. In a somewhat similar vein, Aghion at al. (2004) discuss a model where a principal also faces an agent over which he has only partial control, and may initially want to relinquish partial (formal) authority in order to learn about her type, in this case her degree of trustworthiness.

A key assumption of the Aghion-Tirole model, is that communication between the parties takes an extreme form: when uninformed, a principal with formal authority 'blindly' follows the instructions of the agent. In general, one would expect that the principal would, in such cases, try to elicit information from the agent and then make the best decision from his point of view, as in the *cheap talk* model of Crawford and Sobel (1982). Of course, the informed agent may then respond by attempting to 'manipulate' the principal. As shown by Crawford and Sobel, equilibrium communication will then inevitably be noisy communication, with a higher level of noise when the principal and agent's objectives are less congruent. Dessein (2002) builds on the Crawford-Sobel model by adding the possibility for the principal to *delegate* the decision altogether to the agent (and giving up on communication). In Dessein's expanded problem, the principal has to decide between *informed-but-biased* decisions (under delegation) and *noisy-but-unbiased* decisions (under no delegation).

For the sake of brevity, let us focus here only on a very simple example, where one project $k \in \{1, 2, 3, 4, ...\}$ has to be chosen, and where the loss function of the principal is $(s - k)^2$, while the loss function of the agent is $(s + b - k)^2$. The variable s is the value of the profit-maximizing project. It can take values 1, 2 or 3, with respective probabilities p_i, $i = 1, 2, 3$. The

agent has a 'bias' $b > 0$ in favor of 'bigger' projects. This would be the case for example if the agent is an 'empire-builder'.

Dessein compares two organization structures: delegation and no delegation. Under delegation, of course, the higher the bias b, the worse the outcome for the principal: for $b < 1/2$, the agent always chooses the project the principal prefers, leading to a loss of 0 for him; for $1/2 < b < 3/2$, the agent always chooses a project one unit higher than what the principal prefers, leading to a loss of 1 for him; for $3/2 < b < 5/2$, the agent always chooses a project two units higher than what the principal prefers, leading to a loss of 4 for him; and so on.

Without delegation, we are in the Crawford-Sobel setting, where communication involves the agent sending (costless) messages to the principal about s. Without loss of generality, given rational expectations, we can assume that these messages are truthful (with noise), but obviously the amount of information they reveal will depend on parameter values. Just as with delegation, the principal will be worse off when the bias rises, because communication will become less revealing. But here, since the principal can always choose the best outcome from his point of view (conditional on the information he has), the loss is bounded: at worst, he can disregard any information sent by the agent and choose the loss-minimizing project, that is, choose the k which minimizes $\sum_k p_i(s-k)^2$. For example, if the principal chooses $k = 2$, his loss is limited to $p_1 + p_3$, which is less than 1. Clearly, delegation is therefore dominated in this example whenever $b > 1/2$.

The next question is: how well can the principal do without delegation?

1. For $b > 3/2$, he can do no better than disregard information sent by the agent: since the agent knows that the principal will never choose any project greater than 3, the principal cannot hope to receive information

leading him to choose a project lower than 3. Indeed, in this case, even if $s = 1$, the agent prefers the principal to choose $k = 3$, so that the only equilibrium involves full pooling, i.e. no meaningful communication.

2. Can the principal hope to obtain some information for lower levels of the bias b? In fact, for $b < 1/2$, full revelation is possible, since the agent prefers the principal to choose $k = s$ rather than a higher k. Here we have equivalence between delegation and no delegation (zero loss for the principal) even if, in the absence of delegation, less than full communication is also an equilibrium: as is well-known, cheap-talk games also admit 'babbling equilibria' with coordination failures, namely the agent sends an uninformative message because she expects the principal not to pay attention, and this is self-fulfilling.

3. Finally, for $1/2 < b < 3/2$, the possibility of partial communication depends on parameter values: (i) since the principal will never choose $k > 3$, the principal cannot hope to distinguish between $s = 2$ and $s = 3$; (ii) similarly, if $p_2 > p_3$, it will be impossible for the principal to distinguish between $s = 1$ and $s > 1$: indeed, the latter case would imply $k = 2$, which is then attractive for the agent if she knows $s = 1$; in this case, we are back to no information revelation; (iii) on the other hand, if $p_2 < p_3$, it becomes possible for the principal to distinguish between $s = 1$ and $s > 1$: indeed, the latter case would now imply $k = 3$, which is unattractive for the agent if she knows $s = 1$, because she prefers $k = 1$ to $k = 3$; in this case, the loss for the principal is only p_2, and he does strictly better than under delegation.

This example delivers a number of general lessons: first, it illustrates the general tradeoff for the principal between informed-but-biased decisions and

noisy-but-unbiased decisions; second, it shows that an increase in the bias hurts the principal both under delegation (because the decision becomes more and more biased) and under no delegation (because it becomes noisier and noisier); third, it shows that delegation becomes worse when the bias is high enough, because without delegation the loss for the principal is bounded.

One dimension for which the answer delivered by the example is special concerns the potential optimality of delegation when the bias b is low (so much so that communication is possible under no delegation). In the original Crawford-Sobel example (which allows for continuous values for k and s, and assumes a uniform distribution for s on a bounded interval), Dessein shows that delegation dominates no delegation for all parameter values for which communication becomes possible under no delegation, which is quite a striking result. Dessein shows moreover that his result does depend on the uniform-distribution assumption. And indeed, here, there exists an equilibrium where no delegation strictly dominates delegation while involving communication when $1/2 < b < 3/2$ and $p_2 < p_3$. On the other hand, and in keeping with the Crawford-Sobel setting, when $p_2 > p_3$, delegation does as well as no delegation whenever communication is possible under no delegation, and one can argue that it is 'safer' to delegate in this case, since this guarantees that a babbling equilibrium is avoided.

4.3 Coordination, Adaptation, Costly Communication and Decentralization

We now turn to another rationale for delegation or *decentralization* inside an organization: economizing on the costly communication, *red tape*, and costly *bureaucracy* that coordination of multiple agents' actions entails under a centralized command. We discuss this theme in the context of Dessein and

Santos' (2003) model of an organization, where the benefit of multiple agents working together in an organization arises from individual agents' ability to increase their productivity through specialization.[8] To reap the benefits of specialization however, requires coordination. In particular, in a changing environment the organization must be able to adapt to new circumstances, which means that individual agents' tasks must be continuously redefined and coordinated to achieve maximum productivity. This coordination may involve more or less high communication costs, depending on how centralized the organization is. While Dessein and Santos assume away incentive problems in their model for simplicity (in this respect following the *team theory* literature), we continue our discussion by exploring how incentives may be introduced in their setting, and the implications for *delegation* that follow from such an extension, using an approach developed by Dewatripont (2006) (itself based on Dewatripont and Tirole, 2005).

4.3.1 The Dessein—Santos model

In their model, Dessein and Santos (hereafter DS) endogenize the degrees of centralization and adaptiveness of an organization with an arbitrary number N of agents as well as the quality of its internal communication. For consistency and simplicity we shall only consider a simplified version of their model, with only two agents and two actions: a *primary* action a and a *complementary* action b. Action a should ideally be as close as possible to local information θ, a random variable with mean θ_0 and variance σ_θ^2. One can interpret θ_0 as the status quo and the realized θ as the change in the organization's environment. Also, action b should ideally be as close as possible

[8] The notion that gains from specialization are the source of value added of an organization is, of course, not new to Dessein and Santos (2003). This idea can be traced back at least to Adam Smith and also underlies recent models of optimal team organization following Radner (1993). See e.g. the chapter by Garicano and Van Zandt (2011).

to action a (think of b as an "input" that must "fit" with a). The expected mis-adaptation and mis-coordination cost to the organization is given by $E\left[\phi(a-\theta)^2+\beta(b-a)^2\right]$, where ϕ is the weight given to mis-adaptation and β the weight given to mis-coordination. With two agents there can only be two possible organizational forms, *decentralization* (or delegation) where each agent is assigned decision rights over one action, and *centralization* where a single agent (the boss) has authority to command both actions (and the other agent simply executes what he is told to do). Under *delegation* each agent is more specialized and therefore more productive, but the organization faces greater coordination and adaptation problems. Under centralization there is better coordination and adaptation but there are lower returns to specialization.

Decentralized organization Decentralization means that two individuals are hired and each individual controls one and only one action. We assume that the individual who chooses a is able to first observe θ but that the individual who chooses b observes neither θ nor a. Let the individual who chooses a be the *sender* (S, "she"), because she must send a message about a to the other individual, called the *receiver* (R, "he"). Communication is assumed to be imperfect: S's message is received by R only with probability p, while R learns nothing with probability $1-p$.

The timing of the game is assumed to be as follows: In stage 1, S observes θ, chooses a, and communicates it to R; in stage 2, R receives S's message with probability p and sets b. Given that we take a team-theoretic perspective, both parties want to minimize mis-coordination and mis-adaptation costs $E\left[\phi(a-\theta)^2+\beta(b-a)^2\right]$, and so R sets $b=a$ whenever he learns the value of a and sets $b=\theta_0$ otherwise. Consequently, S chooses a to minimize

$\phi(a-\theta)^2 + \beta(1-p)(\theta_0 - a)^2$ for each realization of θ, or:

$$a = \theta_0 + \frac{\phi}{\phi + \beta(1-p)}(\theta - \theta_0).$$

The ratio $\phi/\left[\phi + \beta(1-p)\right]$ can be interpreted as the degree of *adaptiveness* of the organization, and also as the *level of discretion* enjoyed by S. It grows with ϕ and p and goes down with β. Substituting for the value of the optimal action a into $E\left[\phi(a-\theta)^2 + \beta(b-a)^2\right]$ yields the equilibrium expected cost for the decentralized organization:

$$\frac{\phi\beta(1-p)}{\phi + \beta(1-p)}\sigma_\theta^2.$$

As is intuitive, equilibrium costs are increasing with the variance σ_θ^2 (the change in the environment), the importance of mis-adaptation measured by ϕ and the importance of mis-coordination measured by β, and decreasing with the quality of communication measured by p. If the quality of communication can be improved at a cost (i.e. if the firm can invest in better internal communication), DS ask what determines the optimal choice of p. They allow for a continuous p and a continuous convex cost function but here, for simplicity, we shall only allow for two levels: $p_L = 0$ and $p_M > 0$, where p_M is obtained at a cost F for the organization. With an endogenous p, the expected cost of the organization under decentralization is then:

$$EDC = \min\left\{\frac{\phi\beta}{\phi+\beta}\sigma_\theta^2, \frac{\phi\beta(1-p_M)}{\phi+\beta(1-p_M)}\sigma_\theta^2 + F\right\}.$$

Centralized organization Mis-coordination costs can be reduced under centralization, as a single individual commands which actions all the agents in the organization must take. However, centralization comes at the cost of more costly communication, red tape, and less individual discretion (or initiative), which lowers the beneficial scope of specialization. In our simple

two-agent version of the DS model we can only represent centralization and the benefits and costs it entails very crudely.

We shall assume that centralization is essentially a more extreme form of decentralization, with higher quality and costs of communication. That is, under centralization the quality communication p_H and cost K is such that $1 \geq p_H > p_M$ and $K > F$. Moreover under centralization a single individual commands both actions a and b when communication is effective (with probability p_H). When communication is ineffective, the second agent does not receive her order from the boss and makes a "best effort" to coordinate by choosing $b = \theta_0$. Under centralization, the expected cost of the organization is then

$$ECC = \frac{\phi\beta(1-p_H)}{\phi+\beta(1-p_H)}\sigma_\theta^2 + K.$$

The choice of centralization or decentralization for the firm then reduces to a comparison of EDC with ECC. When is decentralization likely to be better than centralization?

Results The comparison of ECC and EDC illustrates the following results derived in a more general setting by DS:

- Higher uncertainty, σ_θ^2, or higher mis-adaptation costs ϕ raise the benefits of adaptiveness, which in turn raises the benefits of communication/centralization. Thus, in highly volatile environments it is best not to decentralize. A decentralized organization would either incur high costs of mis-coordination or high costs of mis-adaptation, as the agent who observes the realization of θ would choose to be less responsive to improve coordination. The DS model thus captures in a simple way why *disaster relief* is best done by highly centralized organizations. [9]

[9] A related but somewhat orthogonal explanation by Bolton and Farrell (1990) is that a

- Higher mis-coordination costs β have ambiguous effects: for β close to 0 or ∞, *decentralization* with $p = 0$ is optimal, but for intermediate values of β this is not necessarily the case. Indeed, a rise in β from a very low level raises the value of communication in order to keep achieving *ex-post coordination*. But when β is very large, the best solution is to have $a \to \theta_0$, and thus avoid mis-coordination costs. In this case the organization relies on *ex-ante coordination*, with very little communication and responsiveness to local information. Decentralization is then optimal, in order to reap the Smithian returns from specialization.

- Finally, lower communication costs F under decentralization may result: (i) in a move away from centralization (with a move from p_H to p_M) and therefore in *decreased adaptiveness*; or (ii) in a *rise in adaptiveness* (with a move from $p_L = 0$ to p_M) while the organization remains decentralized. In a *symmetric multitask* setting DS also show that lower communication costs—by raising the attractiveness of adaptiveness—may at times imply more (partial) centralization with more task bundling and less specialization.

The DS model offers a rich setting, which our two agent illustration cannot fully do justice to. It sheds light on the recent moves toward higher adaptiveness as well as increased communication and task bundling in organizations. In particular, their analysis helps explain the empirical findings of the simultaneous shift toward more communication and more decentralization in organizations of Caroli (2001), Caroli and Van Reenen (2001), and Bresnahan et al. (2002).

centralized organization can respond more quickly to a change in the environment (there is no delay) and can achieve better coordination (there is no duplication of actions). However, this comes at the cost of worse adaptation (under decentralization there is a better utilization of local information).

The DS model (at least in our two-agent formulation) is also a good starting point to introduce incentive considerations. Assuming away internal incentive problems is a useful first step in the study of organization design, but it is only a first step. To better capture the internal organization design problems firms face in practice, it is important to also allow for the incentive problems associated with hidden information and/or hidden actions. As far as communication is concerned, incentive problems can hinder efficient communication in two main respects. First, there can be a *lack of congruence* of objectives between the sender and the receiver with respect to the decision to be taken. Second, there can also be a *moral hazard in communication* problem, which is really a form of *moral hazard in teams problem*, as: (i) the *sender* must spend time, attention, and other resources to communicate her knowledge effectively; and (ii) the receiver must spend time by paying attention, decode, understand, and rehearse the acquired knowledge. In a nutshell, when it comes to communication, "it takes two to tango".

4.3.2 Cheap talk versus costly communication

Imperfect communication with incentives is largely an unexplored topic.[10] One exception is Dewatripont and Tirole (2005; hereafter DT) who discuss a model that considers both imperfect congruence and moral hazard in teams. It distinguishes in particular between *costless but soft communication*—cheap talk, where the receiver must decide whether to trust the sender's message or not—and *hard-but-costly communication*—which the receiver knows he can trust but which entails a cost for both parties.

To introduce communication incentives in the DS framework, suppose

[10] There is a literature on costly state verification, starting with Townsend (1979), which considers costly audits. This literature concerns unilateral information acquisition and thus differs from the model presented here, where communication is costly for both parties.

first that achieving hard-but-costly communication (with probability p_M) requires *the principal and agent* to expend one unit of effort at cost $F/2$. That is, any single effort is wasted if the other party does not also expend effort. Second, to make the problem interesting, suppose also that the two parties have different objectives. Indeed, if they had the same objective and since soft communication about the value of a can occur at no cost, then there is no role for costly communication.

Specifically, assume that (as in DS) S has payoff $U_S \equiv -sE(\phi(a-\theta)^2 + \beta(b-a)^2)$, but that R has payoff $U_R \equiv -r(\alpha E\left|b-a\right| + (1-\alpha)E\left|b-\theta_0\right|)$. With probability α it is then in R's interest to align his action on S by picking $b=a$, and with probability $(1-\alpha)$ it is in R's interest to instead stay with the status quo ($b=\theta_0$). Thus, α is a measure of the degree of *congruence* between the two parties. Assume, finally, that neither S nor R know the realization of R's objective function, but that α is common knowledge.

The timing of the game is then as follows: In stage 0, each party privately decides whether or not to expend one unit of effort, at an individual cost of $F/2$. In stage 1, S observes θ, chooses a, and—if she has expended effort—can at no further cost try to engage in hard-but-costly communication. In stage 2, if both parties have expended effort, then R learns the outcome of hard-but-costly communication. Cheap talk is also always possible, independently of communication efforts. Finally, in stage 3, R sets b.

What is the difference between cheap talk and hard-but-costly communication in this setting? We assume that hard-but-costly communication enables R to learn his objective function. In contrast, under cheap talk, R has to trust S about a and cannot learn his objective function (which S does not know and therefore cannot convey).[11] For simplicity,we further assume

[11] Note, however, that S has an incentive to tell the truth about a.

that cheap talk conveys the value of a to R with probability 1.

Under these assumptions, it is easy to see that one of two outcomes will obtain.

- Under *high-congruence* (which occurs when $\alpha > 1/2$), S has "real authority" in the sense of Aghion and Tirole (1997): she can pick $a = \theta$ and announce it to R using cheap talk while not expending effort on costly communication since, faced with the impossibility of learning about his objective function, R cannot do better than choosing $b = a$.

- Under *low-congruence* (which occurs when $\alpha < 1/2$), R will not select $b = a$ unless he becomes convinced that his objective function is $-rE|b - a|$. In this case, cheap talk is useless and the only way to induce R not to pick $b = \theta_0$ is to have successful costly communication.[12] For a given α, an equilibrium of costly communication (with probability p_M) exists when the cost of communication F is low enough and when *both* parties' stakes in the relation are high enough (i.e. when $\min\{s, r\}$ is high enough).

An increase in the congruence parameter α between S and R therefore leads at some point to the breakdown of costly communication when S can start counting on R to *rubber-stamp* her (costless) recommendation. This breakdown leads to an upward (resp. downward) jump in S's (resp. R's) expected payoff. When this point is reached, centralization is of course not attractive for the sender. On the other hand, centralization might be attractive for $\alpha < 1/2$, provided its cost is not too high.

While the DT model allows for costly communication as an alternative to cheap talk, Kartik et al. (2007) and Kartik (2009) have more recently gen-

[12] Even then R picks $b = a$ only with probability α.

eralized the Crawford-Sobel *cheap-talk* paradigm to allow for endogenously costly soft information communication by introducing intrinsic *lying costs* for the sender. Kartik et al. (2007) show that even with low costs of lying it is possible for the sender to truthfully communicate her information (without noise) in a separating (signaling) equilibrium when the sender's type space is unbounded. In a more general model with convex lying costs Kartik (2009) also shows that signaling equilibria involve *language inflation* in the sense that equilibrium messages overstate the truth in the direction of the sender's bias. Moreover, when lying costs decrease the equilibria with language inflation involve more inflation and become less informative. A natural question worth exploring in such a model is when delegation of authority from the receiver to the sender is a superior outcome to centralization with inflated and distorted (costly) communication.

5 Contingent allocation of authority

Another common form of authority allocation in organizations is a temporary, contingent, allocation in response to major changes in the environment, unexpected events, or the advent of special problems. There may be several reasons why control needs to be shifted in this way.

A first reason may be the limited attention of the party in control. When major new changes occur the controlling party may not have the time to deal with these new problems or opportunities. As is often the case in government, one response is to delegate decision-making authority over these new tasks to a new agent or committee. Even if the original party retains formal control rights, the committee gains real authority as it is able to make better informed decisions, as Aghion and Tirole (1997) have stressed.

A second reason why control needs to be allocated to a new party (typi-

cally appointed from outside the organization) is that the existing holder of authority may no longer be best placed to deal with the new situation. Thus, a common practice in many corporations is to replace managers of underperforming units and to temporarily appoint so called *turnaround managers*, who are specialized in restructuring of failing businesses. These managerial services are often provided by major consulting firms, or by buy-out firms who then also acquire the unit and run it until it is turned around and resold. Also, large distressed corporations entering chapter 11 often appoint a specialized crisis manager referred to as *chief restructuring officer* (CRO) to provide expertise in reorganizing the firm while it operates under bankruptcy protection.

A third reason why control is removed from the agent in charge when there is significant underperformance is to protect other agents in the organization against decisions that may go against the interests of the organization. The manager in charge of the failing unit may, for example, be tempted to *gamble for resurrection*, by taking excessively risky actions which could restore his reputation and further his career in the event of a good outcome.

Although the contracting literature has not analyzed the *turnaround management* problem *per se*, there are at least two closely related literatures that can shed light on this problem. One is the literature on referrals in hierarchies initiated by Garicano (2000) and the other is the literature on state-contingent allocations of control following Aghion and Bolton (1992). We shall discuss the main themes of these two strands of literature in turn and explain how they shed light on the turnaround management problem.

Taking the referral, or management-by-exception, perspective of Garicano (2000), the management turnaround intervention can be seen as an optimal institutional arrangement designed to address unusual problems requiring

extraordinary skills. Thus, according to Garicano it would be too costly to train all managers to be able to rectify the operations of their divisions when they seriously underperform, as in normal circumstances divisions are expected to perform properly and turnaround of operations requires more involved and costly training. In an effort to minimize managerial training costs it may be more desirable to let a few talented managers specialize in turnaround skills (the management consultants and private equity shops) and let them take over control of a division or the entire firm in the event of major underperformance. It is worth noting that this is an efficient organizational arrangement even in situations considered by Garicano, where all managers' objectives are aligned.

On the other hand, when managers' objectives cannot be perfectly aligned with the firm's overall objectives, the choice between an internal or external turnaround manager may no longer be a matter of indifference. When the divisional manager has private information on when it is optimal to allocate authority to a turnaround manager he may choose to delay reporting of damaging information to hold off the change in control. In such situations it may be optimal to give the divisional manager either a stake in the success of the turnaround, or some alternative job protection or career path. Although these incentives have the negative effect of dulling the divisional manager's incentives to run the divisions' day-to-day operations they have the advantage of avoiding the worst underperformance and of permitting more timely interventions. However, as suggested by Garicano and Santos (2004) in their analysis of optimal referrals under asymmetric information, to make way for such incentive adjustments, a more inconspicuous internal turnaround procedure may be desirable than the very visible outsourcing to an external turnaround expert.

Related to the Garicano (2000) analysis of ad-hoc efficient interventions is the analysis of ex ante contingent allocation of authority or control in Aghion and Bolton (1992) and the subsequent literature on financial contracting. However, unlike in the literature on hierarchies and referrals the basic reason why a change of management may be desirable in response to underperformance is not so much to appoint a turnaround manager with special skills, than to ensure that divisional managers' objectives are best aligned with the firm's objectives.

To see why an ex-ante commitment to replace a divisional manager only in some contingencies helps align managerial and firm objectives consider the following stylized example of a division in a firm. The division's lifecycle can be divided into three key dates: date 1 when the division's investment of I is undertaken; date 2 when a new event or state of nature occurs, which calls for new actions on how to run the division, such as whether to let it continue or whether to close down its poorly performing operations; and, date 3 when the final returns from the division's operations are realized. In our simple example we normalize the returns to 0 in the event of "failure" of the division's investment and 1, in the event of "success". We shall also take it that the probability of success or failure of the investment depends both on the (*verifiable*) realization at date 2 of a state of nature θ and on a managerial decision to be taken following the realization of the state. There are only three equally likely states of nature in our example denoted respectively by θ_1, θ_2, θ_3.

As in the previous sections, the role for managerial authority arises from a basic limitation of contract enforceability: the inability to specify a complete state-contingent enforceable action-plan to be carried out by the divisional manager. All the contract with the manager can specify is that he is in charge

of running a division and that he is ultimately responsible for its performance. The contract can also specify an incentive scheme which is based on the division's final performance and, if the division is a wholly owned subsidiary, in what states of nature θ the holding company can remove the divisional manager.

We shall also simplify the managerial decision problem to the extreme and assume that in each state of nature the divisional manager only needs to decide whether to restructure the division or to continue operating as usual. In our example, expected payoffs in each state and for each action choice are then given as follows:

1. *Returns*: In any given state θ the probability of success of the investment is π_R when the manager chooses to restructure. Instead, when the manager chooses to continue operations as usual, the probability of success is π_i in state θ_i, with $\pi_1 < \pi_2 < \pi_3$.

2. *Payoffs*: Company headquarters' objectives for the division are to get the highest possible financial return. The divisional manager, however, also derives on-the-job, private, benefits (in the form of perquisites, prestige, reputation, power, etc.) as long as the division is not restructured. We denote these benefits by $B > 0$. These payoffs capture the idea that divisional managers often don't like to "rock the boat", or are reluctant to fire their employees in a restructuring out of loyalty to them.

3. *Optimal decision*: Assume restructuring is desirable in a first-best world only in state θ_1, that is, $\pi_1 + B < \pi_R < \min\{\pi_2, \pi_3\} + B$. But, from the point of view of company headquarters alone, we assume it would be best to restructure in both states θ_1 and θ_2 as $\pi_2 < \pi_R < \pi_3$.

4. *Participation constraint*: The new investment is worth undertaking as long as the division is restructured in at least state θ_1, since we assume that $(\pi_R+\pi_2+\pi_3)/3 > I$. However, the new investment is unprofitable if the divisional manager has to be compensated for the loss of private benefits in state θ_1, that is, $((\pi_R - B) + \pi_2 + \pi_3)/3 < I$.

In this example the first-best outcome can be achieved if headquarters can commit to only leaving the divisional manager in charge in states θ_2 and θ_3. If the divisional manager could not be dismissed in any state without cause he would only agree to restructure the division in state θ_1 if he gets fully compensated for the loss of private benefits B. But, then the divisional manager would be too entrenched and would make the division unprofitable.

Unfortunately, letting headquarters decide on management changes and on appointments of turnaround managers ex-post would also be inefficient. Indeed, since $\pi_2 < \pi_R < \pi_3$, headquarters would choose to appoint a turnaround manager in both states θ_1 and θ_2, thus rendering the divisional manager's job too precarious. This is why there is generally a benefit of committing ex-ante to precise and objective conditions which can lead to a managerial turnover, or a change in control.

In mature and stable organizations, it may be possible to achieve such a credible ex-ante commitment informally by developing a reputation for predictable human resource-management policies. However, for younger and smaller firms the choice is often only between full entrenchment of divisional management or complete discretion of headquarters. An intermediate solution is also to set up the division as a wholly owned subsidiary and to write an enforceable *management contract*, with contingent control change clauses, between the subsidiary and the holding company.

6 Interpersonal authority

In this section we review models of organizations where contractual enforcement is extremely limited. Ex-ante agreements on action choices between a principal and an agent are not enforceable, and formal agreements on transfers of authority (absent transfers of ownership titles) are not enforceable either. The key question these models are concerned with is how the principal–who does not have formal control rights, or who has to rely on the agent taking a desired action–can nevertheless exercise authority. When can he rely on his superior knowledge–his *intellectual authority*–to get the agent to implement his *recommendations*? When does he need to give a more formal order, how can he get the agent to follow his orders, and, finally, what is the role of financial incentives in helping induce the agent do what the principal wants?

The first model we discuss builds on Aghion and Tirole's 1997 model (AT). Unlike in their model, the principal cannot formally appropriate authority from the agent. The principal may, however, rely on his intellectual authority to get the agent to do what he wants. He has intellectual authority over the agent whenever he is informed but the agent is not. In that case a simple recommendation will be sufficient to get the agent to do what he wants. The principal may also exercise his authority by giving the agent an order. This is a more effective way of getting things done, as the agent then has to do what he is told, even when he prefers some other action. However, to make sure that the agent executes the order, the principal then has to rely on a costly enforcement technology.

6.1 A simple model of *hard* and *soft* authority

The model outlined in this section is concerned with the question of how to best exercise authority. Three main modes of authority are considered and compared: first, *hard authority*, which takes the form of the principal ordering the agent to take a specific action; second, *soft authority*, which can take a weaker and a stronger form. The strong form involves the principal making recommendations rather than issuing orders. Finally, under the weak form the principal gives discretion to the agent to decide what to do.

As the focus of the analysis is on the relative benefits of soft and hard authority, we simplify the model by assuming that the principal is always well-informed about project returns. As in AT, we reduce the authority problem to a one-shot relation, where the principal cannot take the action himself and has to rely on the agent to get things done. The principal's options are:

Hard authority: the principal can at some cost K make his orders *verifiable*, and thus force the agent to follow them.

Soft authority: the principal makes a recommendation or gives advice at some cost $k \in (0, K)$, but ultimately the agent is free to ignore the advice and to take whatever action she prefers.

Discretion: the principal gives the agent *carte blanche* and lets her decide what the best action is.

More formally, as in AT, assume the two parties face $N > 3$ mutually exclusive (ex ante identical) actions (or projects). Ex post, all but two of these projects produce losses for both parties. Of the two remaining projects, one gives the principal a private benefit $B > 0$ and the agent a private benefit βb (with $1 > \beta > 0$). The other project gives the principal a lower private benefit of αB (with $1 > \alpha > 0$) and the agent a higher private benefit of

$b > 0$. Finally, inaction is also an option and gives both parties a payoff of zero.

While the principal always knows the true payoffs of all the projects, the agent only gets this information with probability $p \in (0, 1)$. Ex-ante expected payoffs are then as follows under the different modes of authority.

- Hard authority: the principal is sure to be able to enforce his preferred action and obtains $B - K$, while the agent obtains βb.
- Soft authority: when the principal makes a recommendation it will be followed by the agent if and only if she is uninformed. Thus, the principal can expect to get $(1-p)B + p\alpha B - k$, while the agent expects $(1-p)\beta b + pb$.
- Discretion: the agent chooses her preferred action when informed and otherwise chooses inaction. Thus, the principal expect to get $p\beta B$, and the agent pb.

Although this is a highly stylized model of interpersonal authority it still produces some striking insights. First, if p is small enough, the principal does not need to rely on hard authority, where the agent is required to follow specific orders: all he needs to do is recommend an action (provided k is small enough), and the agent will follow the recommendation with high probability. Second, if p is high, the principal may want to rely on hard authority (provided K is not too large), as otherwise the agent will, with high probability, not follow his recommendations. In other words, the principal may be worse off hiring an agent who knows too much or is *overqualified* (the principal's payoff is not monotonic in p). This may be one simple explanation for the observed difficulty highly-qualified unemployed workers have of

getting employment in jobs requiring lesser skills or experience. Third, when principal and agent objectives are highly congruent (α and β are close to one) then the principal may prefer to give full discretion to agents with a high p.

The insights gained from this simple model can be applied to two important aspects of job design. First, if the amount of discretion granted to the agent can be contracted on, how much discretion should the agent be given? Second, if the principal can give financial incentives to the agent, to what extent should the agent's compensation be based on output performance?

6.1.1 Partial contracting over authority

Suppose that the probability p that the agent learns the true payoffs of all the projects is a function of some prior effort of the agent. That is, the agent could choose to learn more, gain more skills and thus be able to raise p. As we have seen, to the extent that a higher value of p may force the principal to switch from a soft to a hard authority mode, the principal could be made worse off by such an increase in p. In such a situation, the principal could benefit from contractually limiting the agent's discretion, either by reducing the number of actions in her action set, or by committing to a hard authority mode. By only granting partial control to the agent, the principal could discourage the agent from learning too much and thus make the agent more 'obedient'.

As an illustration, suppose that a contract can be written giving the agent only *partial* authority over projects $j = \xi, ..., N$, where $\xi \geq 1$. Suppose in addition that if the agent puts in effort e she becomes informed with probability $p(e) = e$, and that her effort cost is $\psi_A(e) = \frac{1}{2}e^2$. Then, given that projects are ex-ante identical, under a mode of soft authority, the agent

chooses e to maximize her expected payoff:

$$\frac{N-\xi+1}{N}eb + (1-e)\beta b - \frac{1}{2}e^2.$$

The probability $p(e^*) = e^*$ that the agent becomes informed is then given by $e^* = ((N-\xi+1)/N-\beta)b$, so that the agent can be induced to become more and more *obedient* by limiting her authority by letting ξ rise.

6.1.2 Authority and financial incentives

Except for the very top managers in a firm, it is rare to see other employees' compensation be based mainly on firm performance.[13] As Van den Steen (2007) argues, one reason may be that, although high-powered output-based incentives boost employees' effort incentives, they may also exacerbate differences of opinion and conflicts over the optimal choice of projects. He argues that output-based incentive schemes may therefore be dominated by input-based incentive schemes.

To illustrate this point in our simple model suppose that the monetary payoff of any project q_j can take two possible values $q_j \in \{0, X\}$, where $X > 0$. The principal and agent's "private benefits", B and b, now stand for the principal's and agent's respective *beliefs of success*: $B = \Pr_P(q_j = X) \in [0,1]$ and $b = \Pr_A(q_j = X) \in [0,1]$. Each project costs $\kappa > 0$ to undertake and as always we assume that only two actions generate positive net payoffs given the principal's and agent's respective beliefs: $\alpha BX - \kappa > 0$ and $\beta bX - \kappa > 0$ (where as before $\alpha \in (0,1]$ and $\beta \in (0,1]$). In addition, when uninformed all projects look alike to each agent and have a negative net expected value: $E_0[q_j] - \kappa < 0$.

[13] Two notable exceptions, in which output-based incentive schemes are more common, are the high-technology start-ups and investment banking.

Suppose that the principal learns the payoffs associated with each project with fixed probability $E \in (0,1)$. The principal could offer the agent an output-based incentive scheme to induce her to invest in knowledge and discover the payoffs associated with each project. The benefit to the principal is that a positive net present value project is then more likely to be selected. Let $\theta \in (0,1)$ denote the share of profits the agent gets under an output-based incentive scheme and suppose that θ maps into a probability $e = \upsilon\theta$ of the agent learning the payoffs of each project, where $\upsilon \in (0,1)$ is a parameter measuring the incentive efficiency of the performance-based compensation contract.

The agent's knowledge about projects is only valuable in states of nature where the principal is uninformed. In those states the net benefit to the principal is $\alpha B(1-\theta)X - \kappa$, since the agent chooses her preferred project and thus generates an expected payoff αB given the principal's beliefs, and since the principal only gets a share $(1-\theta)$ of the financial returns, and pays the set-up cost κ. From an ex-ante perspective, the expected value to the principal of giving the agent a performance-based incentive contract θ under a *soft authority* mode is then

$$E([(1-\upsilon\theta)B + \upsilon\theta\alpha B](1-\theta)X - \kappa - k) + (1-E)\upsilon\theta(\alpha B(1-\theta)X - \kappa - k).$$

Indeed, whenever the agent is informed she ignores the principal's recommendation and just picks her favorite action. Therefore with probability $\upsilon\theta$ the principal only gets $\alpha B(1-\theta)X$. The marginal benefit of raising θ is then

$$(1-E)\upsilon(\alpha BX(1-2\theta) - \kappa - k)$$

and the marginal cost is

$$EBX(1 + \upsilon(1-\alpha)(1-2\theta)).$$

Therefore, whenever

$$EBX(1+\upsilon(1-\alpha)) > (1-E)\upsilon(BX\alpha-\kappa-k)$$

it is optimal for the principal not to put the agent under a high-powered output-based incentive scheme. Note in particular that this condition is more likely to hold the lower is α. In other words, the more the principal and agent's beliefs differ the less the principal wants to reward the agent based on output performance:[14] he would rather protect his real authority over the agent.

6.1.3 Abuse of authority

So far we have only considered situations where both the principal and agent benefit from an action choice. Moreover, we have assumed that the congruence of the principal's and agent's objectives is commonly known to both parties. As a result, an uninformed agent is happy to follow the principal's recommendation, as she knows that she would also benefit from taking the recommended course of action.

But suppose now that the principal's recommended actions do not always benefit the agent and that the principal may have better information than the agent on the two parties' objectives. Specifically, suppose that the degree of congruence can be either (α_H, β_H) or (α_L, β_L), where $1 > \alpha_H > \alpha_L > 0$ and $1 > \beta_H > 0 > \beta_L$. And assume that the principal knows the true value of α and β but the agent's prior beliefs put equal probability weight on high and low congruence. Under these parameter values, the principal may want to take an action under low congruence even though the agent is made worse off by that action choice. Thus, the agent now worries that with a 50%

[14] Similarly, the less effective is the incentive scheme (the lower υ) the more likely is this condition to hold.

probability the principal's favorite project is worse than doing no project at all.

Under these circumstances the preferred mode of authority may vary depending on the realization of (α, β) so that the principal's chosen mode of authority may itself be a *signal* of congruence. For example, the principal may choose soft authority when congruence is high and hard authority when it is low. This is the case when the following conditions hold:

$$p\alpha_L B + (1-p)B - k \leq B - K - \beta_L b$$

and

$$p\alpha_H B + (1-p)B - k \geq B - K$$

This simple example illustrates the effect first highlighted in Dessein (2005), and mentioned at the beginning of Section 4.2 above that, when congruence is high but unknown to the agent, the principal may prefer to give her more discretion. The principal's recommendation then carries more weight, as the agent always chooses to follow the principal's recommendation when she is uninformed herself.

The example above illustrates a situation of benign potential abuse of authority, which can be overcome through an appropriate choice of mode of authority. In the next section, however, we explore a situation with more severe abuse of authority, which cannot be overcome in a simple one-shot interaction between the principal and agent. In this situation the principal always has an informational advantage over the agent ($p = 0$) and may prefer actions which make the agent strictly worse off. For example the principal may have an action in his choice set with payoffs $(\theta B, \lambda b)$, where $\theta > 1$ and $\lambda < 0$, but where $\theta B + \lambda b < B + \beta b$. Clearly, in a one-shot contracting relation the agent will not be able to stop the principal from recommending

the *abusive* or *exploitative* action. But, as we illustrate in the next section in a repeated–long-term employment–relation similar to relational contracting situations studied in Bull (1987), MacLeod and Malcomson (1989), Levin (2003), and Baker et al. (1999) among others, it is possible for the principal to *gain authority* over the agent by developing a reputation for treating employees fairly.

6.2 Authority as an Optimal Relational contract

This final section also discusses a form of authority which has to do with issuing orders that are expected to be followed by subordinates, because of an informational advantage of the principal, and which is sustained by an ongoing relation built on *trust*, as Bolton and Rajan (2000) have argued (see Bolton and Dewatripont, 2005 chapter 12). The central questions concerning this form of authority are why it exists in the first place, and why this mode of transaction is preferred over a spot-contracting mode.

Bolton and Rajan (2000) assume that giving orders requires in the first place superior information by the principal (the issuer of orders). The principal must have superior information on which action by the agent is best for her. Second, an ongoing authority relationship based on trust is fundamentally more flexible than a spot-market negotiation-based transaction. However, such a relation is not always sustainable if the agent cannot have enough trust that the principal will not exploit the agent in the authority relation.

As in Simon (1951), Bolton and Rajan compare two modes of transacting, a *negotiation/contracting mode* in anonymous markets and an *authority mode* built on long-term personal relations. In the contracting mode, the services or goods to be provided by an agent (the *seller*) as well as the terms of

trade are spelled out in detail in a spot contract. In the authority mode, the principal (the *buyer*) writes a long-term employment contract with the agent (seller), specifying only the terms of employment, leaving the details of which service to provide in any given period unspecified. In this mode, the buyer gives orders or *directs* the seller to perform a specific service in each period. The seller only has the choice of executing the order, or quitting. There are no ongoing negotiations about which service to provide, or at what terms.

As the authority mode is based on a long-term contract and an ongoing relationship, the timing of the seller's payments can be made more *flexible*. The principal can now compensate the agent with a *bonus* after the latter has carried out a particularly costly service. The principal's incentive to voluntarily pay such a bonus, which is purely discretionary, is supported by the agent's threat to dissolve the relationship should the principal not compensate him adequately. When the principal is expected to always fully cover the agent's costs ex-post, the agent is also willing to execute the order. Moreover, since the principal always ends up paying the true cost of services ordered under the authority relation, she has an incentive to always demand the value maximizing services. In other words, the principal chooses the first-best action in the authority mode, and thus generates an efficiency gain, which will be lost should the agent (employee) decide to quit. It is the prospect of losing this efficiency rent which preserves the principal's incentives to adequately compensate the employee's costs and induce him to stay. Thus, as long as the principal and agent transact sufficiently frequently and do not discount the future too much they will be better off in a long-term employment relationship than by trading anonymously in a spot market.

A related analysis of authority as an optimal relational contract has been proposed by Wernerfelt (1997). He considers an employment relationship as

an efficient relational contract to economize on explicit contract negotiation costs. The main benefit of the employment relation in his model is that fewer contractual terms have to be bargained over before trade can take place.

7 Conclusion

As the literature discussed in this chapter highlights, approaching the analysis of the *internal organization of firms* from the perspective of *allocation of authority* to managers is a fruitful avenue of research. This approach emphasizes the dilemmas managers face in exercising authority and thereby sheds new light on the importance of (formal and informal) internal communication protocols, as well as institutional commitments to preserve managerial discretion. The rational contracting perspective of most articles reviewed in this chapter, however, paints an excessively hopeful picture of the efficiency of firms' internal organizations. In reality, the sheer complexity of the problem of dynamic organizational design is likely to overwhelm even the most persistent efforts at rationalizing the firm's internal hierarchy. What is more, power struggles and office politics, forces which this chapter has entirely ignored, substantially complicate the implementation of a rational internal order.

References

Aghion, P. and P. Bolton (1992). "An Incomplete Contracts Approach to Financial Contracting," *Review of Economics Studies*, 59: 473-494.

Aghion, P., M. Dewatripont and P. Rey (2002). "On Partial Contracting," *European Economic Review*, 46, 745-53.

Aghion, P., M. Dewatripont and P. Rey (2004). "Transferable Control." *Journal of the European Economic Association*, 2: 115-138.

Aghion, P., M. Dewatripont and J. Stein (2008). "Academic Freedom, Private-Sector Focus and the Process of Innovation", *Rand Journal of Economics*, 39: 617-635.

Aghion, P. and J. Tirole (1997). "Formal and Real Authority in Organizations", *Journal of Political Economy*, 105: 1-29.

Alchian, A., and H. Demsetz (1972). "Production, Information Costs, and Economic Organization." *American Economic Review, 62: 777-95.*

Auriol, E., G. Friebel and L. Pechlivanos (2002). "Career Concerns in Teams," *Journal of Labor Economics* 20: 289-307.

Baker, G., R. Gibbons, and K. Murphy (1999). "Informal Authority in Organizations." *Journal of Law, Economics, and Organization*, 15: 56-73.

Bolton, P. and M. Dewatripont (2005). *Contract Theory*, Cambridge, MA: MIT Press.

Bolton, P. and J. Farrell (1990)."Decentralization, Duplication, and Delay". *Journal of Political Economy*, 98: 803-826.

Bolton, P. and A. Rajan (2000). "The Employment Relation and the Theory of the Firm: Arm's Length Contracting Versus Authority", mimeo, Princeton.

Boot, A. W., S. I. Greenbaum and A. V. Thakor (1993). "Reputation and Discretion in Financial Contracting," *American Economic Review*, 83:

1165-83.

Bresnahan,T., E. Brynjolffson, and L.M. Hitt (2002). "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics*, 117: 339-376.

Bull, C. (1987). "The Existence of Self-Enforcing Implicit Contracts", *Quarterly Journal of Economics*, 102: 147-160.

Burkart, M., D. Gromb, and F. Panunzi (1997). "Large Shareholders, Monitoring, and the Value of the Firm". *Quarterly Journal of Economics*, 112: 693-728.

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

Caroli, E. (2001). "New Technologies, Organizational Change, and the Skill Bias: What Do We Know?" in *Technology and the Future Employment of Europe*, edited by P. Petit and L. Soete. Edward Elgar.

Caroli, E., and J. Van Reenen (2001). "Skill Biased Organizational Change? Evidence from a Panel of British and French Establishments." *Quarterly Journal of Economics*, 116: 1449-1492.

Coase, R.H. (1937). "The Nature of the Firm." *Economica, 4: 386-405.*

Crawford, V. and J. Sobel (1982), "Strategic Information Transmission." *Econometrica,* 50: 1431-1452.

Dessein, W. (2002). "Authority and Communication in Organizations." *Review of Economic Studies*, 69: 811-838.

Dessein, W. (2005). "Information and Control in Ventures and Alliances." *Journal of Finance, 60: 2513-2550.*

Dessein, W. and T. Santos (2006). "Adaptive Organizations". *Journal of Political Economy*, 114: 956-995.

Dewatripont, M. (2006). "Costly Communication and Incentives", *Jour-*

nal of the European Economic Association, Papers and Proceedings, 4, 253-68.

Dewatripont, M., I. Jewitt and J. Tirole (1999). "The Economics of Career Concerns, Part I: Comparing Information Structures; and Part II: Application to Missions and Accountability of Government Agencies, *Review of Economic Studies,* 66: 183-217.

Dewatripont, M., I. Jewitt and J. Tirole (2000)."Multitask Agency Problems: Task Separation and Clustering", *European Economic Review*, 44: 869-877.

Dewatripont, M. and J. Tirole (1994a). *The Prudential Regulation of Banks*, Cambridge, MA: MIT Press.

Dewatripont, M. and J. Tirole (1999). "Advocates", *Journal of Political Economy*, 107: 1-39.

Dewatripont, M. and J. Tirole (2005). "Modes of Communication", *Journal of Political Economy*, 113: 1217-1238.

Fama, E. F. and M. C. Jensen (1983). "Separation of Ownership and Control," *Journal of Law and Economics*, 26: 327-49.

Friebel, G. and M. Raith (2004). "Abuse of Authority and Hierarchical Communication," *Rand Journal of Economics*, 35: 224-244.

Garicano, L. (2000). "Hierarchies and the Organization of Knowledge in Production". *Journal of Political Economy*, 108: 874-904.

Garicano, L. and T. Santos (2004). "Referrals". *American Economic Review*, 94: 499-525.

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: 1119-58.

Hart, O., and J. Moore (2000). "On the Design of Hierarchies: Coordination versus Specialization", mimeo, Harvard University.

Holmstrom, B. (1979). "Moral Hazard and Observability," *Bell Journal of Economics*, 10: 74-91.

Holmstrom, B. (1982a). "Moral Hazard in Teams," *Bell Journal of Economics*, 13: 324-340.

Holmstrom, B. (1982b). "Managerial Incentive Problems: A Dynamic Perspective", in *Essays in Economics*

and Management in Honor of Lars Wahlbeck (Helsinki: Swedish School of Economics); reprinted in 1999 in the *Review of Economic Studies*.

Holmstrom, B. and P. Milgrom (1991). "Multi-Task Principal Agent Analyses," *Journal of Law, Economics and Organization*, 7, Special Issue.

Itoh, H. (1991). "Incentives to Help in Multi-Agent Situations." *Econometrica* 59: 611-636.

Kartik, N. (2009). "Strategic Communication with Lying Costs", *Review of Economic Studies*, 76: 1359-1395.

Kartik, N., M. Ottaviani and F. Squintani (2007)."Credulity, Lies, and Costly Talk". *Journal of Economic Theory*, 134: 93-116.

Levin, J. (2003). "Relational Incentive Contracts", *American Economic Review*, 93: 835-857.

MacLeod, B., and J. Malcomson (1989). "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica*, 57: 447-480.

Martimort, D. (1996). "Exclusive Dealing, Common Agency and Multiprincipals Incentive Theory." *Rand Journal of Economics,* 27: 1-31.

Maskin, E. (1999). "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 66, 23-38.

Maskin, E., and J. Tirole (1989). "The Principal-Agent Relationship with an Informed Principal: The Case of Private Values", *Econometrica 58:*. 379-409.

Maskin, E. and J. Tirole (1999), "Unforeseen Contingencies and Incomplete Contracts," *Review of Economic Studies*, 66: 83-114.

Mirrlees, J. A. (1999), "The Theory of Moral Hazard and Unobservable Behaviour: Part I," *Review of Economic Studies*, 66: 3-21

Myerson, R., (1983) "Mechanism Design by an Informed Principal", *Econometrica* 51: 1767-98.

Radner, R. (1993). "The Organization of Decentralized Information Processing." *Econometrica* 61: 1109-1146.

Riordan, M. and D. Sappington (1987) "Information, Incentives and Organizational Mode", *Quarterly Journal of Economics*, 102(2), 243-63.

Roberts, J. (2004) *The Modern Firm: Organizational Design for Performance and Growth*, Oxford University Press

Rotemberg, J., and G. Saloner (1993) "Leadership Styles and Incentives," *Management Science*, 39, 1299-1318.

Segal, I. (1999), "Complexity and Renegotiation: A Foundation for Incomplete Contracts," *Review of Economic Studies*.

Simon, H. (1951). "A Formal Theory of the Employment Relationship." *Econometrica*, 19: 293-305.

Stern, S. (2004). "Do Scientists Pay to Be Scientists?" *Management Science*, 50: 835-853.

Townsend, R. (1979). "Optimal Contracts and Competitive Markets with Costly State Verification". *Journal of Economic Theory* 21, 265-293.

Van den Steen, E. (2007), "The Cost of Incentives under Disagreement (Can an Employee be Too Motivated?)", mimeo, MIT.

Van den Steen, E. (2010), "Interpersonal Authority in a Theory of the Firm", *American Economic Review* 100: 466-490.

Waldman, M. (1990), "Up-or-Out Contracts: A Signaling Perspective". *Journal of Labor Economics* 8: 230-250.

Wernerfelt, B. (1997) "On the Nature and Scope of the Firm: An Adjustment-Cost Theory", *Journal of Business, Vol. 70, No. 4: 489-514.*

Williamson, O. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.

Williamson, O. (1979). "Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics*, 22: 3-61.

Williamson, O.(1985). *The Economic Institutions of Capitalism*. New York: Free Press.

Williamson, O., M. Wachter and J. Harris, (1975). "Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange." *Bell Journal of Economics*, 6: 250-80.

Wilson, J.Q. (1989). *Bureaucracy: What Government Agencies Do And Why They Do It*. New York: Basic Books.