ON THE DESIGN OF HIERARCHIES: COORDINATION VERSUS SPECIALIZATION*

by

Oliver Hart
Harvard University and London School of Economics

and

John Moore
London School of Economics and University of St Andrews

Contents:

The Suntory Centre
Suntory and Toyota International Centres for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 0171-405 7686

Discussion Paper
No. TE/99/375
October 1999

---

* We thank Philippe Aghion, Antoine Faure-Grimaud, Ulrike Malmendier, Fausto Panunzi, Rohan Pitchford, Andrei Shleifer, Jean Tirole, Chenggang Xu, and seminar audiences at the Western Economic Association and European Association for Research in Industrial Economics (EARIE) meetings for helpful comments. We acknowledge financial support from the U.S. National Science Foundation through the National Bureau of Economic Research, the Leverhulme Trust, the LSE Financial Markets Group, and the Suntory and Toyota International Centres for Economics and Related Disciplines at the LSE.

Abstract

We develop a model of hierarchies based on the allocation of authority. A firm's owners have ultimate authority over a firm's decisions, but they have limited time or capacity to exercise this authority. Hence owners must delegate authority to subordinates. However, these subordinates also have limited time or capacity and so further delegation must occur. We analyze the optimal chain of command given that different agents have different tasks: some agents are engaged in coordination and others in specialization. Our theory throws light on the nature of hierarchy, the optimal degree of decentralization, and the boundaries of the firm.

Keywords: organisations, hierarchies, coordination, specialisation.
JEL Nos.: D23, L22.

1. Introduction

The purpose of this paper is to develop a model of hierarchies based on the allocation of authority. We take the view that a firm's owners have ultimate authority over a firm's decisions, but that they have limited time or capacity to exercise this authority. Hence the owners must delegate some authority to subordinates, i.e., they must grant the subordinates the right to make decisions that they themselves are unable or unwilling to make. However, these subordinates also have limited time or capacity to exercise authority and so further delegation must occur to other subordinates. Thus we view a firm as a chain of command over decisions. We use our model to analyze the optimal chain of command given that different agents have different tasks; in particular, some agents are engaged in coordination and others in specialization. Our theory throws light on the nature of hierarchy and the optimal degree of decentralization inside a firm, as well as on the boundaries of the firm.

There is a vast literature on many of the issues we consider, and this is not the place to provide a review. Economists have studied hierarchical structure from the point of view of supervision and task assignment (see, e.g., Williamson (1967) and Rosen (1982)); from the point of view of incentive theory (see, e.g., Calvo and Wellisz (1979)); and from the point of view of information processing and team theory (see, e.g., Keren and Levhari (1979), Radner (1992), Bolton and Dewatripont (1994), and Segal (1998)). By and large, however, the existing literature does not analyze hierarchy in terms of authority; that is, in contrast to our model, it is not the case that if i is above j in the hierarchy, then i necessarily has authority over j. Rather, in much of the literature, if i is above j in the hierarchy, then j provides information to i. Also the literature does not distinguish between what happens inside a firm and what happens between firms. In other

words, the optimal hierarchies derived could apply just as well to the organization of production in the U.S.A. as to the organization of production in Microsoft. In contrast, our approach does distinguish between the firm and the economy. In our model, one firm has one person or group with ultimate authority over all decisions (one owner or group of owners), whereas the economy has many people with ultimate authority over different subsets of decisions (many owners or groups of owners).[1]

Although our approach differs from much of the literature, it has parallels with the paper by Aghion and Tirole (1997). In our model, a boss (e.g., an owner) has formal authority in the Aghion-Tirole sense, while a subordinate has real authority if his boss cannot exercise authority but he can. We discuss the relationship further in Section 5.

The basic elements of our model are as follows. We consider an economy consisting of a set of assets and a set of identical individuals. Each asset represents a (residual) decision; that is, a decision must be taken with respect to that asset. We assume that these decisions are noncontractible, both ex ante and ex post. In addition to these basic decisions, there are also "higher-level" decisions, which correspond to the coordination of assets or to synergies among assets.[2] To be precise, we assume that, for each subset of assets A, there is a task t, which consists of trying to come up with an idea about what to do with the assets in A. A task does not

---

[1]In future work, it would obviously be desirable to combine the informational approach in the literature and the authority approach in this paper.

[2]Decisions rather than assets are the key feature of the model. We introduce assets because they are a convenient way to think about decisions, particularly higher-level decisions (synergies).

necessarily reach fruition, that is, become an idea ex post. If an individual's task becomes an idea, we say that the individual is active. In this case he will carry out his idea if he can. If the individual does not have an idea, he is inactive; in this case not only can the individual not implement his own idea, but also he cannot implement anyone else's.

Because any particular individual may not have an idea, it is important for efficiency that each asset has a hierarchy of bosses, that is, a chain of command. The way a hierarchy works is as follows. If the first person in the hierarchy (the ultimate boss) has an idea, she implements it. If she does not have an idea, control passes to her subordinate, who implements his idea if he has one. If the subordinate does not have an idea, the subordinate's subordinate has a chance to implement his idea, and so on.

A key assumption that we make is that ideas are mutually exclusive in the following sense. If one individual implements an idea involving an asset, then someone else cannot implement an idea involving that asset (whether or not the idea involves other assets too). This assumption has a significant implication. The benefit of putting someone high up in a hierarchy is that, if the person has a good idea, he is likely to be able to implement it. The cost is that the person may block others from implementing better ideas.

A (stylized) example may help. Imagine that there are two assets, a hotel and an airplane. Then there are three tasks: coming up with an idea about the hotel; coming up with an idea about the plane; and coming up with an idea about how to coordinate the hotel and the plane (synergy). The model supposes that these ideas conflict. For example, the "synergy" idea to offer hotel discounts to airplane passengers is inconsistent with the "hotel" idea to refurbish the hotel in the next three months.

Assume that there are three individuals, one carrying out each task.

One hierarchical structure would make the coordinator--that is, the person working on the hotel-plane synergy--senior on both assets. In this case, if the coordinator is active, i.e., has an idea, she can implement her idea, whether or not the specialists have ideas, since she has authority. In contrast, each specialist can implement his idea only if (a) he has one; (b) the coordinator does not.

Another hierarchical structure would reverse the roles: the coordinator would be junior to the specialists on both assets. In this case, each specialist can implement an idea whenever he has one, while the coordinator can implement her idea only if (a) she has one; (b) neither specialist does. (If either specialist implements his idea, this preempts the use of one of the assets, which means that the coordinator cannot implement her idea.)

A priori it is not clear which of these two hierarchical structures is better, although the second one seems unconventional. It is in fact an implication of one of our main results that the second hierarchical structure is suboptimal (given some additional assumptions).

Returning to the general case, we assume that the organizational form--characterized by a chain of command over each asset and an assignment of tasks to each individual--is chosen ex ante to maximize expected total surplus. From the design point of view, the key questions are, what tasks should each individual be assigned to carry out and what is the optimal chain of command for each asset? One of our principal results is that, given the assumption that the probability of an idea is decreasing in the set of assets being looked after, individuals with a broad remit, i.e., whose tasks cover a large subset of assets, should appear higher in the chain of command than those with a narrow remit. In other words, big thinkers or coordinators should be senior to small thinkers or specialists. We also establish that "criss-cross" hierarchies are never optimal; that is, if individual i appears

above individual j on one asset, j will not appear above i on another asset. Finally under an additional assumption, we show that the optimal hierarchy is a pyramid, in the sense that each individual has at most one boss.

We use these results to analyze the trade-off between centralization and decentralization. We define an organization to be centralized if most individuals in it are coordinators, and an organization to be decentralized if most individuals in it except for the top people are specialists. We show that if the gains to coordination are large enough, it is optimal for the organization to be centralized; if the gains to coordination are significant but not too large, it is optimal for the organization to be decentralized; and finally if the gains to coordination are small, then it is optimal for the organization to split up into several independent firms.

The paper is organized as follows. We set out the model in Section 2. In Section 3 we establish our main result, that individuals with a broad remit should be senior to those with a narrow remit. In Section 4 we provide a detailed analysis of the symmetric two-asset case. Section 5 is devoted to some foundational issues. Finally, Section 6 contains extensions and further discussion.

2. The Model

We consider an economy consisting of m assets, $a_1,\ldots,a_m$, and n (risk neutral) individuals $1,\ldots,n$. The economy begins at date 0, and at this point organizational form is chosen. Each asset represents a decision that has to be made in the future at date 1. These decisions are noncontractible both ex ante and ex post. However, authority over decisions can be allocated at date 0, as can the tasks that people are engaged in.

For simplicity we assume that all individuals are identical. At date 0 each individual is assigned a task. (More than one person can be assigned the same task.) A task consists of trying to come up with an idea about what to do with a subset of the m assets, i.e., what decisions to make with respect to these assets at date 1. For each set $A \subset \bar{A} = \{a_1,\ldots,a_m\}$, there is a corresponding task t(A). Not all tasks reach fruition, that is, become ideas. We write the probability that task t(A) becomes an idea as p(A), where $0 < p(A) < 1$.

An individual who has an idea about the subset of assets A, and is able to implement it, generates value $v(A) \geq 0$ (measured in money). We put few restrictions on the function v, other than to suppose that $v(\{a_k\}) > 0$ for all k. In particular, v may depend on the identity of the assets in A as well as on their number. Also v may not be superadditive or even nondecreasing in A.

This last point deserves discussion. We have in mind a situation where thinking about how to use two assets is a very different activity from thinking about how to use one of them. The first activity involves coordination while the second does not. If coordination possibilities are limited, then the value of having an idea about how to coordinate two assets may be very low. Thus $v(\{a_1, a_2\})$ could be smaller than $v(\{a_1\}) + v(\{a_2\})$, or even than $\text{Min}\left[v(\{a_1\}), v(\{a_2\})\right]$ if synergies between the assets are

sufficiently small.[3]

We will make the following (quite strong) assumptions about the generation of value:

(A1) To realize v(A), an individual carrying out task t(A) needs access to all the assets in A. If he has an idea but has access only to a non-empty, strict subset of A, he obtains a positive, but insignificant, value.

(A2) (No externalities.) All the value from an idea accrues to the individual whose idea it is (think of a pet project). In particular, ideas cannot be transferred: i cannot carry out j's idea.

(A3) Having an idea is an independent event across individuals.

(A4) There is no ex post renegotiation (e.g., because of shortage of time). That is, authority cannot be bought and sold at date 1.

We suspect that not all of these assumptions (except possibly for (A4)) are essential, but they greatly simplify the analysis. We discuss (A4) further in Section 5. Note that, as will be seen below, it is the absence of costless ex post renegotiation that provides a role for hierarchical

---

[3]Take the hotel-plane example of the introduction. Consider the comparison between the profit from offering hotel discounts to airplane passengers and the profit from refurbishing the hotel and charging higher hotel prices. The first may be bigger than the second if the plane flies to an airport near the hotel; but smaller if it does not.

structure in our model. In fact, we view hierarchical structure as a substitute for ex post renegotiation.

(A1) - (A4) have a simple but useful implication:

(*) An individual who has an idea and can implement it (even if only partially, i.e., even if he obtains only an insignificant value) will always do so; he will never defer to someone junior, however productive the junior person is. Also a senior individual who is inactive will never wish to veto the idea of a subordinate.

The first part of (*) follows from the fact that, if a senior person with an idea defers to a junior person, he loses his private value (for which he cannot be compensated--given (A4)). The second part follows from the fact that, given (A2), a senior person without an idea neither gains nor loses from his subordinate's idea.

We now turn to the allocation of authority at date 0. We associate with each asset a hierarchy of bosses, that is, a chain of command. Formally, a chain of command is a list, i.e., a sequence of a subset of the numbers 1,...,n (the list may contain all the numbers 1,...,n, none of the numbers, or a strict subset of the numbers; no number is repeated). The first number in the list refers to the ultimate boss, the second number to his subordinate, the third number to the subordinate's subordinate, and so on. Given a chain of command, the most senior person on the list with an idea implements it. If no one in the chain has an idea, the asset yields zero value.

We define an organizational form at date 0 to be a delineation of a chain of command for each asset and an assignment of tasks to each individual. We assume that both the chain of command and the tasks can be

specified in an enforceable contract.[4]

We make a final assumption:

(A5) There is costless (Coasian) bargaining at date 0, and individuals are not wealth-constrained.

(A5) is in stark contrast to (A4). We have in mind that there is plenty of time for the parties to negotiate at date 0, but very limited time (no time) to negotiate at date 1. (A5) implies that organizational form will be chosen at date 0 to maximize expected total surplus, with the surplus being divided up using lump sum transfers.

Before we write down the formula for expected surplus, we can simplify matters a little. Suppose an individual's task consists of looking after assets in the subset A. Then it makes no sense to put the individual in the list (chain of command) involving an asset $a_k \notin A$, since the individual will never have an idea about a_k.

Similarly, suppose an individual is assigned the task t(A). Then it makes no sense <u>not</u> to put him in the list (chain of command) involving each asset $a_k \in A$, since he generates no significant value unless he has control over each of the assets in A.

Putting these two observations together, we can conclude the following. Once the lists for all the assets have been determined, we can figure out which task each person is doing by seeing which list he appears in: if the union of the lists he appears in corresponds to the set of assets A, then he will be doing task t(A).

---

[4]In Section 6 we briefly discuss what happens if tasks are noncontractible.

An example might be useful at this point. Suppose there are two assets and four people (m = 2, n = 4). Figure 1 illustrates three possible organizational forms.

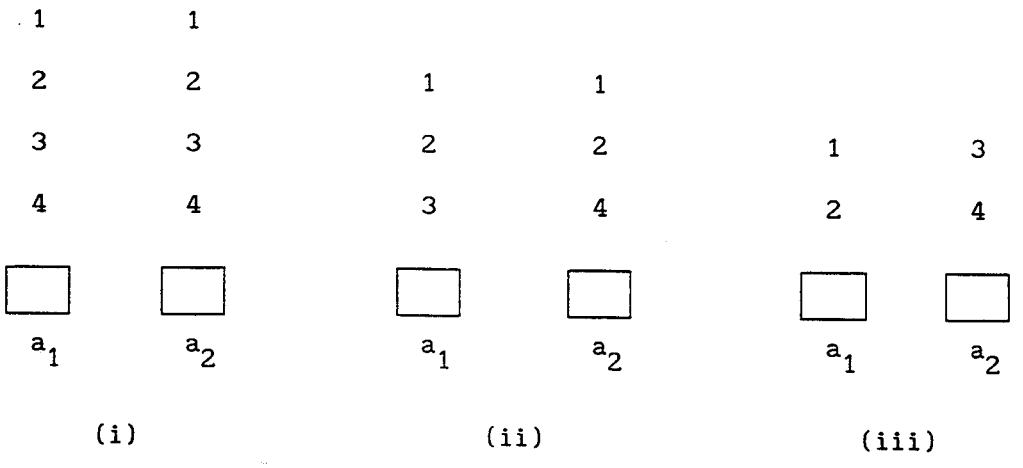

Figure 1

In the first form, 1 is the boss of a_1 and a_2, and has 2, 3, and 4 as subordinates on both assets. The tasks correspond to this assignment of authority: all four individuals are engaged in looking after both assets. In the second form, 1 is the boss of both assets, 2 is 1's subordinate on a_1 and a_2, 3 is 2's subordinate on a_1, and 4 is 2's subordinate on a_2. Again, the tasks correspond: 1 and 2 look after a_1 and a_2, 3 looks after a_1 and 4 looks after a_2. In the third form, 1 is the boss of asset a_1, 2 is his subordinate on a_1, 3 is the boss of a_2, and 4 is his subordinate on a_2. 1 and 2 look after a_1 and 3 and 4 look after a_2.

The forms have a natural economic interpretation. The first two represent a single firm since both assets have the same ultimate boss, individual 1 (who can be interpreted as the owner of the assets). The second form can be thought of as corresponding to a more decentralized firm than the

first because authority is more likely to be exercised by someone with a narrow remit--a specialist--and less likely to be exercised by a coordinator (in the second form, 3 or 4 gets to exercise authority if 1 and 2 don't have an idea). Finally, the third form represents two firms since assets a_1 and a_2 have different ultimate bosses: 1 is the boss of a_1, 3 is the boss of a_2.

We now write down the general formula for expected total surplus in the m asset, n individual case. Let L_k be the list associated with asset a_k. For individual i, define

$$A_i = \{\text{assets } a_k \mid i \text{ appears on list } L_k\}.$$

A_i is the set of assets over which i can exercise authority. From the above we know that individual i will be engaged in task $t(A_i)$. Also for individual i define

$$S_i = \{\text{individuals } j \mid \text{for some asset } a_k, \text{ i and j both appear on list } L_k \text{ and j appears above i}\}.$$

S_i is the set of individuals who are senior to i on some asset. Now we know that individual i receives value $v(A_i)$ if and only if i has an idea and nobody senior to i on any of the assets i looks after has an idea. Given (A3), we can therefore write the formula for total expected surplus as

$$V = \sum_{i=1}^{n} p(A_i) \left(\prod_{j \in S_i} (1 - p(A_j)) \right) v(A_i). \tag{2.1}$$

According to (A5), organizational form will be chosen to maximize (2.1).

3. An Example and the Main Theorem

In this section we establish some general results about optimal organizations. Part (a) of Theorem 1 provides a surprisingly powerful characterization of an optimal hierarchy. It says that an optimal organizational form has the property that an individual's place in the hierarchy is determined (entirely) by his probability of having an idea: individuals with the lowest probability of having an idea are placed at the top of the hierarchy, individuals with the next lowest probability of an idea are placed next in the hierarchy, and so on. Part (b) (which pretty much follows from part (a)) says that criss-cross arrangements are never optimal. That is, if j is above i on one asset, i will never be above j on another asset.

Before we state Theorem 1, and two corollaries, it is useful to get some intuition from a special case. Suppose there are two assets, a_1 and a_2, and two individuals, 1 and 2 ($m = n = 2$). Given our assumption that individuals are identical, but assets may not be, there are nine distinct organizational forms. (In what follows, everything is unique up to the permutation of the individuals' names.) To see this, note that there are two organizational forms where both individuals look after both assets (2 can be senior to 1 on both, or senior on one and junior on the other); four forms where 2 looks after two assets and 1 looks after one (2 can be senior or junior on the asset 1 looks after; and 1 can look after a_1 or a_2); and three forms where 1 and 2 both look after one asset (they can look after different assets or the same asset, which may be a_1 or a_2).

Some of these forms are illustrated in Figure 2. (We leave out the symmetric version of (iii), (iv) and (v), where assets a_1 and a_2 are reversed.)

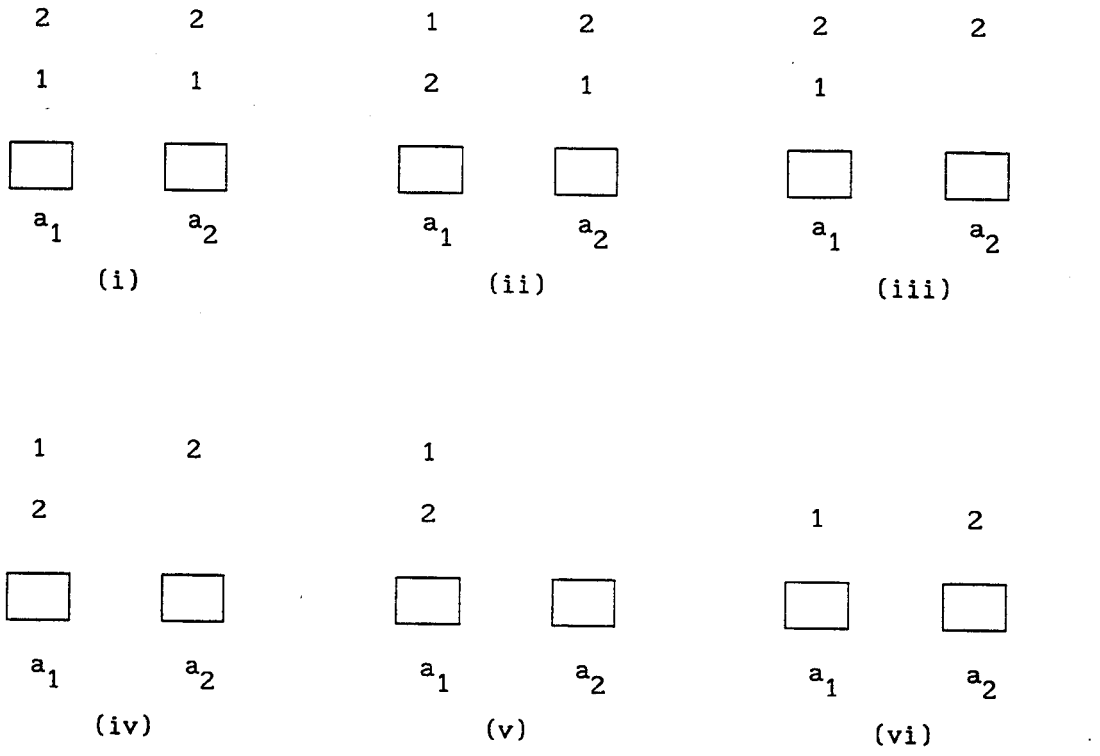

Figure 2

Forms (i), (iii), (v) and (vi) make good economic sense. We would expect form (i) to be desirable if coordination is very important; form (iii) to be desirable if some coordination is important but not too much; form (v) to be desirable if coordination is not valuable and a_1 is very productive; and form (vi) to be desirable if coordination is not valuable and both assets are comparably productive. (Each of these forms can be shown to be optimal for a suitable choice of the parameters.)

However, forms (ii) and (iv) seem strange. Does it make sense to have someone coordinate and yet be junior on some asset, given that this implies that he will rarely be able to implement his coordination idea?

Fortunately, (ii) is never optimal, and neither is (iv) (under an

additional assumption). To see this, note that the expected surpluses (values) from forms (i) - (v) (represented by $V_1,\ldots,V_5$, respectively) are given by

$$V_1 = [1 - (1 - p_2)^2]v_2 = (2p_2 - p_2^2)v_2,$$
$$V_2 = 2p_2(1-p_2)v_2,$$
$$V_3 = p_2 v_2 + (1 - p_2)p_1 v_1,$$
$$V_4 = p_1 v_1 + (1 - p_1)p_2 v_2,$$
$$V_5 = [1 - (1 - p_1)^2] v_1 = (2p_1 - p_1^2)v_1,$$

where $v_1 = v(\{a_1\})$, $v_2 = v(\{a_1,a_2\})$, $p_1 = p(\{a_1\})$, $p_2 = p(\{a_1, a_2\})$. (To understand these formulae, note that in (i) coordination occurs if either 1 or 2 (or both) is active; in (ii) coordination occurs if exactly 1 or 2 is active (but not both); in (iii) coordination occurs if 2 is active whether or not 1 is active; and in (iv) coordination occurs if 2 is active but 1 is not.)

It is immediate that $V_2 < V_1$ and so (ii) is not optimal. To see whether (iv) can be optimal, note that, if it is, we must have $V_4 \geq V_1$ and $V_4 \geq V_5$. The first implies $p_1v_1 \geq p_2 v_2 (1 + p_1 - p_2)$, while the second implies $p_2v_2 \geq p_1v_1$. These cannot both be true, as long as we are prepared to assume $p_2 < p_1$.

This example illustrates the theorem (and corollaries) stated below: (a) If individual 2's task is such that he has a lower probability of an idea than individual 1, then it is not optimal to put 1 above 2 on any asset ((iv) is not optimal). (b) Criss-cross arrangements are not optimal, i.e., it cannot be the case that 2 is above 1 on one asset and 1 is above 2 on another ((ii) is not optimal).

Without the monotonicity assumption on probabilities, $p_2 < p_1$, we cannot

rule out organizational forms like (iv). In the above example, let $v_1 = 10$, $v_2 = 8$, $p_1 = 1/4$, $p_2 \simeq 1$. Then direct calculation shows that (iv) is optimal.[5]

Let us return to the m asset, n individual case. Recall from Section 2 that A_i is the set of assets over which i can exercise authority and L_k is the list (chain of command) associated with asset a_k. We now state the main theorem.

Theorem 1. Consider an optimal organizational form.

(a) Suppose $p(A_i) > p(A_j)$. Then, for all assets $a_k \in A_i \cap A_j$, j appears above i on list L_k.

(b) Criss-cross arrangements are not optimal, i.e., if j appears above i on list L_k for asset a_k, then there does not exist k' such that i appears above j on list $L_{k'}$ for asset $a_{k'}$.

Proof: See appendix.

Notice that, when $p(A_i) \neq p(A_j)$, part (b) of the theorem follows immediately from part (a). The heart of part (b), therefore, lies in showing that criss-cross arrangements are also not optimal when $p(A_i) = p(A_j)$.

Of course, the leading case where $p(A_i) = p(A_j)$ is when $A_i = A_j$: part

---

[5]Although (iv) may be optimal when $p_2 > p_1$, we can rule out (iii) in this case. It follows from direct calculation that $p_2 > p_1 \Rightarrow$ either $V_3 < V_1$ or $V_3 < V_5$. This finding provides another illustration of part (a) of Theorem 1: If a specialist has a lower probability of an idea than a coordinator ($p_2 > p_1$), then it is not optimal to put a coordinator above a specialist on any asset ((iii) is not optimal).

(b) of the theorem then says that, given two people with the same remit, one should be senior to the other on all the assets they work on. Below we state a corollary that deals with the more interesting case where $A_i \subset A_j$ $A_i \neq A_j$, i.e., j's remit is broader than i's.

First, we make an observation about Theorem 1. Part (a) at first looks a little suspicious. It would appear that the decision about who to put on top of a hierarchy is determined solely by the probability of the success of an idea and not at all by the value of an idea. For example, suppose $p(A_j)$ and $v(A_j)$ are very low. Then putting j at the top of the hierarchy is very inefficient, and yet the theorem suggests that this is optimal. The reason there is no contradiction is that the theorem says nothing about which tasks should be carried out. Given that j has a low probability of success and is unproductive even when he has an idea, j is clearly doing the wrong task. That is, in an optimal organizational form no-one will be doing task $t(A_j)$.

This observation about unproductive individuals with low probabilities having the wrong tasks gives the clue to how part (a) of the Theorem is proved. Take an asset a_k, and suppose that agent i is senior to agent j on list L_k, but $p(A_i) > p(A_j)$. In broad terms, we show that expected surplus can be increased by making one of two changes to the organizational form. Either i is relatively unproductive ($v(A_i)$ is relatively low), in which case expected surplus can be increased by switching i to task $t(A_j)$ and placing him just under j in seniority on all assets in A_j--akin to changing from hierarchy (iv) to hierarchy (i) in Figure 2 (with $i \equiv 1$, $A_i \equiv \{a_1\}$, $j \equiv 2$, $A_j \equiv \{a_1,a_2\}$, and $a_k \equiv a_1$). Or j is relatively unproductive ($v(A_j)$ is relative low), in which case expected surplus can be increased by switching j to task $t(A_i)$ and placing her just under i in seniority on all assets in A_i--akin to changing from hierarchy (iv) to hierarchy (v). The merit of these two kinds of maneuver is that one can keep track of how overall expected surplus

changes. By contrast, if we consider a third kind of change to the organizational form, the ostensibly more straightforward maneuver of simply switching the seniorities of i and j, then the people who lie between i and j (on list L_k) are affected in subtle ways and so the overall change to expected surplus is complicated. (In terms of our earlier two-person analysis, this third maneuver amounts to switching the seniorities of 1 and 2 on asset a_1; i.e., changing from hierarchy (iv) to hierarchy (iii). This method happens to work when there are just two people, but not if there are others in between them.)

Corollary 1 follows directly from Theorem 1, given a further assumption.

Monotonicity (M). $p(A) > p(B)$ if $A \subset B$, $A \neq B$.

Corollary 1. Assume (M). Consider an optimal organizational form. Suppose $A_i \subset A_j$, $A_i \neq A_j$, where $i, j \in \{1,\ldots,n\}$. Then, for each asset a_k such that $a_k \in A_i$, j will appear above i on list L_k.

Corollary 1 says that, under the assumption that the probability of an idea is decreasing in an individual's span of control, it is optimal for someone with a broad remit to be senior to someone with a narrow remit.

Assumption (M) warrants further discussion. We would argue that this assumption is plausible. Consider the two functions p(A) and v(A) as the set A increases. It would be surprising if p and v moved in the same direction. If p and v both increase, this would say that coordinators are supermen, while if p and v both decrease, it would say that specialists are supermen. It is more likely that p and v move in opposite directions, i.e., p falls and v rises, or p rises and v falls. The first of these seems more reasonable

than the second: that is, it seems to accord with common sense that a coordinator can (on average) achieve a sizable efficiency gain if he has an idea, but that he is not that likely to have an idea. (Note, however, that (M) requires only that p is decreasing in A; it does not require that v is increasing in A.)

Corollary 1 covers only the case where the remits of individuals can be ranked. To this extent, the corollary leaves open the possibility that someone can have two bosses (a non-pyramidal hierarchical structure). That is, person i may be senior to person j on one asset; while person i′ is senior to person j on another asset, whereas i is not. This can happen if the remits of persons i, j and i′ cannot be ranked.

For example, consider the situation illustrated in Figure 3. There are three individuals and six assets, and synergies exist only between assets a_1, a_2, a_3, assets a_4, a_5, a_6, and assets a_3, a_4.

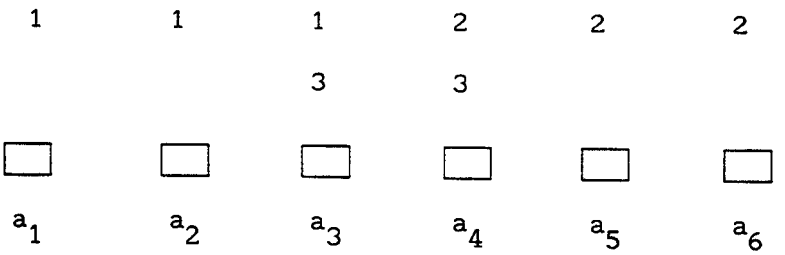

Figure 3

Assume $v(\{a_1,a_2,a_3\}) = v(\{a_4,a_5,a_6\}) = 14$, $v(\{a_3,a_4\}) = 8$, $p(\{a_1,a_2,a_3\}) = p(\{a_4,a_5,a_6\}) = 1/4$, $p(\{a_3,a_4\}) \simeq 1$. Then it is straightforward to show that it is optimal to put individual 1 in charge of a_1,a_2,a_3, individual 2 in charge of a_4,a_5,a_6, and to make individual 3 a subordinate on a_3,a_4. Individual 3 then has different bosses on a_3 and a_4.

In order to rule out this kind of situation, we need a further assumption.

Let us define a set of assets A to be synergistic if $v(A) > 0$.

Nestedness (N). Synergies are nested if given two synergistic sets A, B, either $A \subset B$, or $B \subset A$, or $A \cap B = \emptyset$.

In other words, (N) says that if there is a synergy between a set of assets, then any synergy involving one of the set and a new asset requires the presence of the other assets in the set too. If synergies are nested, the situation in Figure 3 cannot arise, since if $\{a_3, a_4\}$ is a synergistic set (which is why 3 is working on these assets), then $\{a_1, a_2, a_3\}$ is not a synergistic set (and so 1 will not work on these assets).

(N) is quite strong. Note, however, that it is trivially satisfied in the two asset case.

Corollary 2. Assume (M) and (N). Consider an optimal organizational form. If j appears above i on list L_k for some asset a_k, then j appears above i on every list on which i appears.

The proof of Corollary 2 is direct. Suppose i looks after the set of assets A_i and j looks after the set of assets A_j. A_i and A_j must be synergistic since otherwise one of the individuals creates zero value and expected surplus could be increased by assigning this individual to a single asset (any one) and making him the most junior person on this asset. It follows from (N) that $A_i \subset A_j$, $A_j \subset A_i$ or $A_i \cap A_j = \emptyset$. The last is impossible since j appears above i on some asset. If $A_i \subset A_j$, $A_i \neq A_j$, the

conclusion of Corollary 2 follows from Corollary 1. If $A_i = A_j$, the conclusion follows from Theorem 1, part (a). Finally, $A_j \subset A_i$, $A_j \neq A_i$, is inconsistent with Corollary 1 since we know that i does not appear above j on list L_k.

At this point it is worth returning to the two asset-two individual example. Theorem 1 and Corollary 1 imply--and we have also observed this from direct calculation--that the organizational forms (ii) and (iv) in Figure 2 cannot be optimal (assuming $p_2 < p_1$). In fact, if we are prepared to make the additional assumption of symmetry ($v(\{a_1\}) = v(\{a_2\})$, $p(\{a_1\}) = p(\{a_2\})$), (iii) and (v) can also be ruled out. To see this, note that the expected surplus from organizational form (vi) is given by

$$V_6 = 2p_1v_1,$$

which is obviously greater than V_5. (This is just an implication of diminishing returns.) In addition, it is easy to show that either $V_6 > V_3$ or $V_1 > V_3$ (so (iii) is not optimal).

So in the two asset-two individual example, the optimal arrangement is symmetric: either there should be two coordinators (as in (i)) or no coordinators (as in (vi)).

In the next section, we will show that symmetry always holds in the two asset case when the number of individuals (n) is even. This has an interesting implication for Figure 1. In Figure 1 we illustrated three possible organizational forms for the two asset-four individual case. It turns out that, under symmetry, these are the only candidates for optimality (as the next section will show). Moreover, the trade-off between them is as one would expect (at least if $p_1 > p_2$, p_1 not too close to p_2). Form (i) is optimal if the gains to coordination are large enough; (ii) is optimal if the

gains to coordination are moderate; and (iii) is optimal if the gains to coordination are small.

4. The Symmetric Two Asset Case

In this section we analyze in detail the case of n individuals and two symmetric assets (m = 2). We refer to an individual who looks after both assets as a cordinator and an individual who looks after one asset as a specialist. The values they generate if they have an idea and exercise authority over the appropriate assets are given by v_2, v_1, respectively (i.e., $v_2 = v(\{a_1,a_2\})$, $v_1 = v(\{a_1\}) = v(\{a_2\})$). The probabilities of having an idea are given by p_2, p_1, respectively (i.e., $p_2 = p(\{a_1, a_2\})$, $p_1 = p(\{a_1\}) = p(\{a_2\})$).

We assume $p_1 > p_2$. Hence, from the Theorem, coordinators will be senior to specialists and, given any two coordinators, one will be senior to the other on each asset. Thus we can represent an optimal organization as follows:

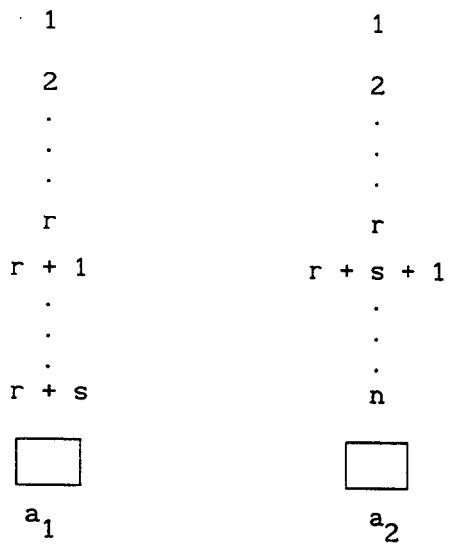

Figure 4

In Figure 4 there are r coordinators and (n - r) specialists. Note that the optimal organization may not be symmetric, i.e., the number of specialists, s, on asset a_1 need not be the same as the number of specialists, n - r - s, on asset a_2.

In fact it turns out that the optimal organization <u>is</u> symmetric, i.e., s = n - r - s, except in the case where n is odd and there are no coordinators at all.

<u>Result 1</u>. Assume m = 2; the assets are symmetric, i.e., $v(\{a_1\}) = v(\{a_2\}) = v_1$, $p(\{a_1\}) = p(\{a_2\}) = p_1$; and $p_1 > p_2$. Consider an optimal organizational form in which individuals 1,...,r look after both assets ($A_i = \{a_1,a_2\}$ for i = 1,...,r), individuals r + 1,...,r + s look after asset a_1 ($A_i = \{a_1\}$ for i = r + 1,...,r + s) and individuals r + s + 1,...,n look after asset a_2 ($A_i = \{a_2\}$ for i = r + s + 1,...,n). Then, unless n is odd and r = 0, s = n - r - s = (n-r)/2.

<u>Proof</u>: See appendix.

The key step in proving Result 1 is to show that if there is some coordination (r > 0) and, say, one more specialist working on asset a_1 than on asset a_2 (s = n - r - s + 1), then it would be better either for the most junior coordinator to switch to specializing on a_2, or for the "extra" specialist on asset a_1 to switch to coordination. Either way, the number of specialists on a_1 and a_2 should be equalized: an asymmetric compromise is never optimal. And if there are no coordinators (r = 0), the number of specialists on each asset should be as equal as possible--i.e., the same when n is even.

From now on, to simplify matters, we will assume that n is even, in

which case the optimal organization is symmetric even if there are no coordinators. As above, write the number of specialists on each asset as s. Since n is even, the number of coordinators must also be even: $r = n - 2s$.

We now consider the optimal choice of s. Before we get into the details, let us note a simple way of thinking about this. Imagine that everyone is a specialist, i.e., $2s = n$. Then the value of the organization is $2(1 - (1 - p_1)^{n/2})v_1$, which is a strictly concave function of n. In other words, not surprisingly, there are diminishing returns to having more specialists. As we will see, this has the following implication: in a class of cases (more precisely, when the surplus maximization problem is convex, and the solution to the problem is interior), then, after s has reached a certain value, it is better not to have further specialists, but rather to make any additional people in the organization (i.e., those at the top) coordinators. That is, for large enough n, the optimal value of s is independent of n.

However, this is not the only possibility. There is another class of cases (when the surplus maximization problem is nonconvex) where the optimum is a corner solution: $s = 0$ or $s = n$.

Now to the details. Denote by V_s the expected surplus (value) of an organizational form in which there are s specialists on each asset and $n - 2s$ coordinators. Suppose $s \leq n/2 - 1$. Let θ be the probability that at least one of the first $(n - 2s - 2)$ coordinators is active (i.e., has an idea). Then we can write

$$V_s = \theta v_2 + (1 - \theta)\hat{V}_s, \tag{4.1}$$

where $\hat{V}_s$ is the value of the organizational form (call it "hat") consisting of everyone <u>but</u> the first $(n - 2s - 2)$ coordinators, i.e., the organizational

form consisting of 2 coordinators and s specialists below them on each asset. The justification for (4.1) is that if one of the first (n - 2s - 2) coordinators is active, which happens with probability θ, v_2 is realized; while, if not, $\hat{V}_s$ is realized.

In turn we can write

$$\hat{V}_s = (2p_2 - p_2^2)v_2 + (1 - p_2)^2(2W_s), \tag{4.2}$$

where W_s is the value of the organizational form consisting of s specialists working on asset a_1 (or a_2). The justification for (4.2) is that in the organizational form "hat," the probability that at least one of the coordinators is active is $(1 - (1 - p_2)^2)$, in which case v_2 is realized; otherwise $2W_s$ is realized.

Now increase s by 1. Since there are now 2 more specialists altogether, the number of coordinators falls to n - 2s - 2. Using the same logic as above, we can write

$$V_{s+1} = \theta v_2 + (1 - \theta)\hat{V}_{s+1}, \tag{4.3}$$

where $\hat{V}_{s+1}$ is the value of an organizational form consisting of (s + 1) specialists on each asset and no coordinators. In turn,

$$\hat{V}_{s+1} = 2[p_1 v_1 + (1 - p_1)W_s]. \tag{4.4}$$

Define $q_1 = 1 - p_1$, $q_2 = (1 - p_2)^2$. Combining (4.1) - (4.4), carrying out some manipulation, and using the fact that $W_s = (1 - q_1^s)v_1$, we obtain

$$V_{s+1} \geq V_s \Leftrightarrow \hat{V}_{s+1} \geq \hat{V}_s \tag{4.5}$$

$$\Leftrightarrow (q_2 - q_1)q_1^{\,s} \geq (1 - q_2)\left(\frac{v_2}{2v_1} - 1\right).$$

(4.5) provides us with important (marginal) information; it tells us when total value can be raised by increasing the number of specialists on each asset from s to s + 1 (and correspondingly reducing the number of coordinators by 2).

In what follows, it is helpful to consider separately the cases $q_1 \geq q_2$ and $q_1 < q_2$. (The first is likely to occur when $p_1 \simeq p_2$ and the second when $p_1 >> p_2$.)

Case 1: $q_1 \geq q_2$

From (4.5), the crucial inequality is:

$$(4.6) \qquad (q_2 - q_1)q_1^{\,s} \geq (1 - q_2)\left(\frac{v_2}{2v_1} - 1\right).$$

Now if $q_1 \geq q_2$ the left-hand side (LHS) is nonpositive and increasing in s. The right-hand side (RHS) may be positive or negative, but it is constant. It follows that if (4.6) holds at a particular value of s, it also holds at s + 1, i.e.,

$$(4.7) \qquad V_{s+1} \geq V_s \Rightarrow V_{s+2} \geq V_{s+1}.$$

(4.7) tells us that the problem of maximizing V_s in Case 1 is nonconvex. An implication is that an interior value of s is never (uniquely) optimal (if $0 < \hat{s} < n/2$ maximizes V_s, then $V_{\hat{s}} \geq V_{\hat{s}-1}$ and so, from (4.6), $V_{n/2} \geq V_{\hat{s}}$, which is a contradiction). Hence, in Case 1, we have a corner solution: s = 0 or s = n/2.

To see which corner is better, we compare

$$V_0 = [1 - (1 - p_2)^n]v_2 = (1 - q_2^{n/2})v_2$$

and

$$V_{n/2} = 2[1 - (1 - p_1)^{n/2}]\, v_1 = (1 - q_1^{n/2})2v_1.$$

It follows that, if $v_2 \geq 2v_1$, $s = 0$ is optimal (since $q_1 \geq q_2$); while, if $v_2 < 2v_1$, $s = n/2$ is optimal for large enough n (since $q_2^{n/2}$, $q_1^{n/2} \to 0$ as $n \to \infty$).

The results for Case 1 are summarized in Result 2.

Result 2. Assume $m = 2$; the assets are symmetric; n is even; $p_1 > p_2$; and $q_1 \geq q_2$, where $q_1 = 1 - p_1$ and $q_2 = (1 - p_2)^2$. Then if $v_2 \geq 2v_1$, it is optimal to have n coordinators on the two assets (a completely centralized firm). On the other hand, if $v_2 < 2v_1$, then, for large enough n, it is optimal to have $n/2$ specialists on each asset (two independent firms).

The intuition behind Result 2 is straightforward. If $v_2 \geq 2v_1$, coordination adds value. Since $q_1 \geq q_2$ (i.e., p_2 is not much smaller than p_1), the expected return from coordination is also quite large. It is therefore not surprising that all individuals will be assigned to the task of coordination. On the other hand, if $v_2 < 2v_1$, two specialists create more value than one coordinator, conditional on all of them having ideas.. Moreover, when n is large, the probability that at least one of n coordinators is active, or that at least one of $n/2$ specialists is active, is close to 1. So value, conditional on having an idea, is the only thing that matters. It follows that specialization is better when $v_2 < 2v_1$.

We turn next to the case $q_1 < q_2$.

Case 2: $q_1 < q_2$

In Case 2 the LHS of (4.6) is positive and decreasing in s, while the RHS is constant. It follows that, if (4.6) fails to hold for s = 0, it fails to hold for all s. Hence $(q_2 - q_1) \leq (1 - q_2)\left(\frac{v_2}{2v_1} - 1\right)$ is a sufficient condition for $V_{s+1} \leq V_s$ everywhere. In other words, if

$$v_2 \geq 2v_1\left(\frac{1 - q_1}{1 - q_2}\right), \tag{4.8}$$

then V_s is maximized at s = 0.

At the other extreme, if $v_2 \leq 2v_1$, the RHS of (4.6) is nonpositive and so the LHS > RHS for all s. In this situation $V_{s+1} \leq V_s$ for all s, i.e., s = n/2 is optimal.

The interesting parameter range in Case 2 is where

$$2 < \frac{v_2}{v_1} < 2\left(\frac{1 - q_1}{1 - q_2}\right). \tag{4.9}$$

In this range, $V_{s+1} > V_s$ for small s, and $V_{s+1} < V_s$ for large s. To be more precise, when (4.9) holds, the problem of maximizing V is convex and there is an interior solution (for large enough n) characterized by the "first order condition": $V_{s-1} \leq V_s > V_{s+1}$. Write the optimal value as $\bar{s}$. From (4.5), $\bar{s}$ is given by the smallest integer greater than or equal to the solution x of

$$q_1^x = \frac{(1 - q_2)}{(q_2 - q_1)}\left(\frac{v_2}{2v_1} - 1\right). \tag{4.10}$$

The results for Case 2 are summarized in Result 3.

Result 3. Assume m = 2, the assets are symmetric, n is even, and $q_1 < q_2$, where $q_1 = 1 - p_1$, $q_2 = (1 - p_2)^2$. Then:

(1) If $v_2 \geq 2v_1 \left(\frac{1 - q_1}{1 - q_2} \right)$, it is optimal to have n coordinators on the two assets (a completely centralized firm).

(2) If $v_2 \leq 2v_1$, it is optimal to have n/2 specialists on each asset (two independent firms).

(3) If $2 < \frac{v_2}{v_1} < 2 \left(\frac{1 - q_1}{1 - q_2} \right)$, it is optimal to have $(n-2\bar{s})$ coordinators on the assets, with $\bar{s}$ specialists below them on each asset, where $\bar{s}$ is the smallest integer greater than or equal to the solution of (4.10).

The thrust of Result 3 is similar to that of Result 2. The higher v_2 is relative to $2v_1$, the greater are the gains from coordination and the more coordinators one would expect to have. Result 3 tells us that, if $(v_2/2v_1)$ is high enough, everyone will be a coordinator (complete centralization). If $(v_2/2v_1)$ is low enough, no one will be a coordinator (two independent firms). In between there will be a mix of senior coordinators and junior specialists (a decentralized firm); moreover, the number of specialists is increasing in $(v_2/2v_1)$.

One aspect of the interior solution (part (3) of Result 3) deserves

particular mention. Provided n is large enough (specifically, provided $n > 2\bar{s}$), the optimal number of specialists on each asset is independent of n! This is an instance of a more general idea: the shape of an optimal sub-hierarchy--that is, everyone below a group of senior coordinators--is independent of how many senior coordinators there are in that group. The reason is that the sub-hierarchy makes a contribution to the organization if and only if no one in the group of senior coordinators is active. Let us call this event E. By assumption (A3), E is statistically independent of which agents in the sub-hierarchy are active. Hence, for a given number of agents in the sub-hierarchy, the optimal design of the sub-hierarchy is independent of E: conditioning on an independent event does not change the nature of the design problem. In particular, for the parameter values in part (3) of Result 3, consider the sub-hierarchy of (say) m people, comprising the $\bar{s}$ specialists on each asset together with the $m - 2\bar{s}$ most junior coordinators (take $m > 2\bar{s}$). Then the optimality of this sub-hierarchy is unaffected by the number, n - m, of senior cooordinators. As the total number, n, of agents grows, provided the optimal number, $\bar{s}$, of specialists on each asset has been reached, all additional agents should be assigned to be coordinators. Of course, for large n the event E (the event that none of the n - m senior coordinators is active) becomes extremely unlikely; but that does not affect the optimality of having just $\bar{s}$ specialists on each asset.

Given this logic, it is not surprising that in all cases in Results 2 and 3, the optimal number of coordinators, $n - 2\bar{s}$, grows as n grows. In effect, an increase in n leads to larger, more centralized firms. To put it another way, a fall in the opportunity cost of workers would make it optimal to have bigger, more centralized firms. We shall return to this point in Section 6 .

Finally, there are some other comparative statics properties that

follow from Results 2 and 3. First, if q_2 falls (i.e., p_2 rises), centralization increases (i.e., the number of specialists, $\bar{s}$, does not rise). This is intuitive since the probability of a coordinator having an idea has increased. Second, if q_1 falls (i.e., p_1 rises), inspection of (4.10) reveals that $\bar{s}$ can rise or fall, i.e., centralization may decrease or increase. The intuition for this is that, although the returns to specialization have increased, which suggests that the number of specialists should rise, diminishing returns to specialization will set in earlier, i.e., it may not be necessary to have as many specialists.

5. Foundational Issues

In this section we discuss further some of the assumptions that we have made.

The model is one of symmetric information and there are no ex ante investments. What matters are ex post decisions. The decisions are noncontractible both ex ante and ex post, but the right to make them can be allocated ex ante (think of allocating the right to press a button). An obvious question to ask is, why doesn't the Coase theorem apply here, i.e., why don't the n individuals simply sit down at date 1 and negotiate over who can press which button? (The tasks would presumably still be agreed on contractually at date 0.) A related question is, even if the Coase theorem doesn't apply, aren't there perhaps mechanisms other than the assignment of authority through hierarchy that could do a better job of maximizing total surplus?

The first point to notice is that in some cases a hierarchical structure can achieve a remarkably efficient outcome--even the first-best. Take the two asset, n individual case analyzed in Section 4 (where n is even). Suppose $v_2 > 2\ v_1$. According to Results 2 and 3, the optimal hierarchy has $(n - 2\bar{s})$ coordinators on the two assets, with $\bar{s}$ specialists junior to them on each individual asset (where $\bar{s}$ may be zero). It is easy to see that this arrangement yields an ex post efficient outcome since (a) if one or more coordinators has an idea, then one of these ideas will be implemented (since coordinators are senior); (b) a coordinator creates at least as much value as two specialists combined (and so if a coordination idea exists, it is efficient to implement it).

In fact, this hierarchical structure is also ex ante efficient; that is, surplus cannot be increased by reassigning tasks. The reason is that, when $v_2 > 2v_1$, the above argument shows that, given <u>any</u> assignment of tasks,

ex post value can be maximized through a hierarchical structure (in which coordinators are senior to specialists). Thus it is impossible to find an assignment of tasks, and an ex post rule for implementing ideas, that creates more value than the optimal hierarchy. Hence, when $v_2 > 2v_1$, the optimal hierarchy identified in Results 2 and 3 achieves the first-best.[6]

In general, of course, a hierarchy will not achieve the first-best. We now try to provide two (very tentative) justifications for being interested in hierarchies in this case too.

First, suppose that having an idea corresponds to being active or "awake," and not having an idea corresponds to being passive or "asleep." Moreover, being "asleep" means being totally incommunicado (e.g., the person is away from the office). Then renegotiation or revelation-type mechanisms that require the participation of all the parties are infeasible, since those who do not have ideas cannot be included.

It is still possible to have renegotiation or mechanisms involving only those who are active at date 1. In fact, this issue is key. Determining the ex post outcome when more than one person is active is what the analysis is

---

[6]The case $v_2 < 2v_1$ is a bit more complicated. Results 2 and 3 tell us that (when n is large) the optimal hierarchy has n/2 specialists on each asset. Since there are no coordinators, this arrangement is trivially ex post efficient. However, we suspect that the outcome may not be ex ante efficient. That is, it may be possible to increase surplus by having some coordinators and some specialists, in combination with a nonhierarchical structure: this nonhierarchical structure would have the feature that a coordinator gives up authority to a specialist if and only if there is at least one specialist on each asset with an idea. (See footnote 7 for a related point.)

all about: given that decisions do no harm, if only one person is active, his decision should be implemented.

We take the view that time pressure at date 1 may make it very costly to have an elaborate exchange of messages or renegotiation even among those who are active. Instead the parties may want to rely on an ex ante mechanism that is "simple." A hierarchical structure is a leading example of such a mechanism. Under a hierarchy, individuals are ranked and the individual with the highest rank among those who are active has the right to decide: complex messages and side-payments at date 1 are thereby avoided.[7]

---

[7]We are not suggesting that a hierarchy is necessarily the only mechanism that is "simple." A hierarchy has the feature that an individual's ranking does not depend on who the other active individuals are. However, one could depart from this. For instance, take the example in Figure 3, but suppose that $v(\{a_3, a_4\}) = 15$. Then the following mechanism achieves first-best (and is more efficient than the hierarchy illustrated in Figure 3): If individuals 1 and 2 both have ideas, then individual 1 has authority over a_3 and individual 2 has authority over a_4; in all other cases individual 3 has authority over a_3 and a_4. This mechanism ensures that, in the event that individual 3 has an idea and only one of individuals 1 and 2 has an idea, a value of 15 is generated instead of 14. Enforcing a mechanism like this may be difficult, however, to the extent that the number of people with ideas may not be verifiable.

In future work it would be interesting to analyze the role of more general mechanisms. Note that nondeterministic mechanisms are also a possibility. For example, one could have a first-come/first-served rule, where the person who first has an idea gets to implement it (see Lando (1998)). We doubt that nondeterministic mechanisms have a useful role to

The assumption that no communication or exchange of messages at all is possible ex post is extreme. There is a second interpretation of our model that is also of interest. Imagine that all ideas can in principle be passed up to the top person in the organization, i.e., even if the boss does not have an idea, he can still consider the ideas of others. The top person is overloaded, however, and so cannot vet every idea carefully. Moreover, he is (approximately) indifferent about which of his employees' ideas are implemented since he does not benefit directly from them. Then the boss may employ a simple "satisficing" rule (which is set up ex ante): he implements his own idea if he has one; otherwise he ratifies or rubber-stamps whatever his subordinate proposes. The subordinate behaves in the same way: he proposes to the boss his own idea if he has one; otherwise he proposes (ratifies / rubber-stamps) his subordinate's idea. And so on. This second interpretation also corresponds to our model, although the foundations for it are less clear.

This second interpretation serves to distinguish the model from Aghion-Tirole (1997). Aghion and Tirole (1997) differentiate between formal and real authority. An owner has formal or legal authority but may not have the information to exercise his authority. Instead he defers to an informed subordinate whose preferences are not too different from his; the subordinate has real authority. In our model, a senior individual has formal authority, while a junior individual has real authority if he has an idea and his boss does not. Aghion and Tirole's interpretation is that the senior individual rubber-stamps the junior individual's proposal in this case since he doesn't

---

play in the present model.

have a better proposal of his own.[8]

Aghion and Tirole focus on the case of two individuals. The question is, what happens if there are more than two, e.g., there is one senior individual and two junior individuals? One natural extension of Aghion-Tirole is to suppose that the senior individual, if he does not have an idea of his own, considers the proposals of both the juniors and chooses the one he likes better. According to this interpretation, no chain of command needs to be set up in advance. All that matters is that one individual is designated the boss: this individual then decides which proposal to favor on an ex post basis.

In contrast we take the Aghion-Tirole model in a different direction. We assume that the boss cannot decide between all the ideas of his employees either because he is not there (he's asleep) or because he is overloaded. Instead he uses a shortcut: he decides in advance who will have authority if he can't exercise it, who will have authority if that person can't exercise it, and so on. In other words, he decides on a chain of command.

Needless to say, the above discussion is quite preliminary. Providing a more formal foundation for the model would be extremely desirable in future research.

We close this section with some remarks about the other assumptions we have made. (A1) - (A3) are all strong, but greatly simplify the analysis. It would be interesting to examine cases where (a) the value of an idea does not accrue entirely to the person having it, i.e., values are not purely private (part of value may be contractible); (b) an individual can achieve value even if he does not have access to all the assets corresponding to his

---

[8]For an alternative interpretation of the Aghion-Tirole model, based on the idea of an implicit contract, see Baker et al. (1999).

task. Note that an implication of allowing for non-private values is that a boss may sometimes defer to a subordinate if the subordinate has a better idea than the boss; or may veto the idea of a subordinate if the idea creates negative value for the boss.

It is also desirable to relax the assumption that the probability $p(A_i)$ of having an idea is exogenous. In a more general model, $p(A_i)$ would depend on individual i's effort decision, which in turn might depend on where i is in the hierarchy. This would bring incentives into the picture and would greatly enrich the analysis.

6. Further Discussion

We have established several results. First, we have shown that, given the assumption that the probability of an idea is decreasing in the set of assets being looked after, individuals with a broad remit, i.e., whose tasks cover a large subset of assets, should appear higher in the chain of command than those with a narrow remit. In other words, big thinkers or coordinators should be senior to small thinkers or specialists. Second, we have shown that "criss-cross" hierarchies are never optimal; that is, if individual i appears above individual j on one asset, j will not appear above i on another asset. Third, if synergies are nested, the optimal hierarchy is a pyramid, in the sense that each individual has at most one boss.

There are several ways in which our work could be extended. In Section 4 we analyzed the two-asset case and considered what happens to the optimal hierarchy as the number of agents increases. We found that if coordination occurs at all it increases (Results 2 and 3). However, this is not the only margin of interest. One could also ask what happens as the number of assets increases along with the number of agents (perhaps assets can be purchased at a fixed price; perhaps the ratio of assets to individuals remains constant). Depending on the nature of synergies, expansion might then take the form of further specialists being hired rather than increased coordination.

In a similar fashion, one could fix the number of assets at more than two and consider what happens as the number of agents increases. For instance, suppose that there are four assets, a_1, a_2, a_3, a_4, and the synergistic sets are $\{a_1, a_2\}$, $\{a_3, a_4\}$ and $\{a_1, a_2, a_3, a_4\}$, in addition to the singleton sets. We know from our two main results that the optimal hierarchy consists of a combination of bosses (working on all four assets), middle-level managers (working on a_1, a_2, and a_3, a_4, respectively) and low-level employees (working on a_1, a_2, a_3, a_4, respectively). But the mix

of these groups can be almost anything, depending on the parameters. For example, it might be optimal to have very few or no middle-level managers; or it might be optimal to have very few or no bosses. We suspect, however, that Result 3 generalizes: that is, if the surplus maximization problem is convex and the solution to the problem is interior (which means that low-level employees, middle-level managers, and bosses are all hired), then after the total number of agents has reached a critical level, all further people in the organization will be bosses.

The case just described helps to cast light on the subtleties of the term decentralization. Consider an organization with several bosses and several low-level employees, but no middle-level managers. This organization might be said to be centralized, given that a decision is likely to be made by a boss (a high-level coordinator). However, the organization might be said to be decentralized given that, if decisions are not made by a boss, they are made by specialists.

On the other hand, consider an organization with no bosses, but many middle-level managers and some low-level employees. This organization might be said to be decentralized, given that a decision is never made by a high-level coordinator (a boss). However, the organization might be said to be centralized, given that decisions are usually made by middle-level managers and rarely by specialists.[9]

---

[9]We should acknowledge that our concept of decentralization is far from standard in the literature. It is not uncommon for economists to define decentralization in terms of span of control. For example, an organization where four workers report to one boss is said to be centralized, whereas an organization where two pairs of workers report to two middle-level managers, who in return report to a boss is said to be decentralized (see, e.g.,

We turn now to some more general remarks about the model. First, it is useful to relate it to the literature on the theory of the firm. The property rights theory (see Grossman and Hart (1986), Hart and Moore (1990), and Hart (1995)) can be thought of as a special case of the current model in which an owner of an asset is always able to exercise control. In effect, $p(A) = 1$ for all A (there is no uncertainty). There is still a nontrivial allocation decision even under these conditions. For example, suppose $m = n = 2$ (two assets and two individuals). Then the choice is between putting one person in charge of a_1 and a_2 and getting $v(\{a_1, a_2\})$; or putting one person in charge of a_1 and the other in charge of a_2 and getting $v(\{a_1\}) + v(\{a_2\})$.

Apart from uncertainty, another important difference between the current model and the standard property rights model is that the current model ignores effort (or investment) incentives. As a consequence, the current model is probably biased toward finding that large firms are optimal. For example, consider Result 3 of Section 4, which says that, given an interior solution, the optimal hierarchy becomes more centralized as the number of workers increases. This result is much less likely to hold once we take into account the fact that lower-level employees' effort incentives will be dulled by the presence of many senior coordinators.

At the same time the property rights model, which is based on entrepreneurial incentives, has a hard time explaining the existence of large firms (except in the presence of very strong complementarities between assets). Thus to have a theory that is biased in the direction of large

---

Wernerfelt (1992)). There does not seem to be a simple connection between this notion of decentralization and ours. The reason is probably that the one in the literature is based on information flowing up the system, whereas ours is based on authority being imposed from above.

firms may not be such a bad thing.

The current model exhibits another, more subtle bias toward large firms. Given our assumption that tasks are contractible, there is in a formal sense no limit to the size of a firm. Take a situation in which there are two assets and two individuals, one looking after each asset. We have interpreted this to represent two independent firms. However, one could also interpret the situation to represent one firm with a single owner, who has committed himself to look after only one asset. Of course, if tasks are noncontractible, this equivalence breaks down: in general, the only way for someone to commit himself to look after one asset is for him to give up the legal right to intervene in a second asset. In future work it would be desirable to drop the assumption that tasks are contractible. We conjecture that this is unlikely to change our results about the nature of an optimal hierarchy very substantially.

We end by mentioning two potential applications of the model. The first is to understand how economies differ during periods of national emergency ("war") as opposed to periods of normality ("peace"). There is a sizeable informal literature arguing that centralization becomes more desirable when urgent decisions are required. However, it has been hard to formalize this idea.[10] Our model seems well placed to deal with this issue since it delivers the prediction that centralization is optimal when coordination benefits are sufficiently great.

Second, ever since Chandler's work (Chandler (1962)), economists have been interested in the choice between the U form and the M form.[11] To

---

[10]For an interesting attempt in this direction, see Bolton and Farrell (1990).

[11]For a recent formalization, see Maskin, Xu and Qian (1999). For a

understand this choice in our model, consider a stylized automobile company. Suppose that there are four assets, a_1, a_2, a_3, a_4, where a_1 represents an asset for producing luxury cars, a_2 represents an asset for producing family cars, a_3 represents an asset for marketing luxury cars and a_4 represents an asset for marketing family cars. One way to structure the automobile company is to combine a_1 and a_2 in one division and a_3 and a_4 in another division. This corresponds to the U form (organization by function). A second way to structure the company is to combine a_1 and a_3 in one division and a_2 and a_4 in another. This corresponds to the M form (organization according to self-contained units). It would be interesting to use the model to analyze the trade-off between these two arrangements. Note that to do this it would seem necessary to drop the assumption that synergies are nested (since presumably there are synergies both between a_1 and a_2 and between a_1 and a_3).

---

discussion of how the U form and M form compare in their ability to solve coordination problems, see Argyres (1995).

APPENDIX (Proofs of Theorem and Result 1)

Proof of Theorem

Consider an optimal organizational form, and two individuals i and j with tasks $t(A_i)$ and $t(A_j)$ respectively. To ease notation, let $p_i = p(A_i)$ and $p_j = p(A_j)$. Suppose $0 < p_j \leq p_i < 1$. Define

X = expected surplus conditional on neither i nor j being active;

X_i = expected surplus conditional on i being active but j not;

X_j = expected surplus conditional on j being active but i not;

X_{ij} = expected surplus conditional on both i and j being active.

That is, the (unconditional) expected surplus equals

$$V = (1 - p_i)(1 - p_j)X + p_i(1 - p_j)X_i + (1 - p_i)p_jX_j + p_ip_jX_{ij}. \quad (A.1)$$

Since we are at an optimum, we know that no organizational form can generate an expected surplus strictly in excess of V. That is, V is maximal.

Now consider a change to this organizational form. Give j the same task as i (task $t(A_i)$) and move her immediately below i in seniority on the lists pertaining to assets A_i, including those assets not in A_j (j no longer appears on the lists pertaining to those assets in A_j that aren't in A_i). The new expected surplus cannot exceed the maximum:

$$(1 - p_i)^2X + [1 - (1 - p_i)^2]X_i \leq V. \quad (A.2)$$

The symmetric argument holds, reversing the roles of i and j. So

$$(1 - p_j)^2 X + [1 - (1 - p_j)^2] X_j \leq V. \qquad (A.3)$$

Inequalities (A.2) and (A.3) will be of use in proving both parts (a) and (b) of the Theorem.

Part (a): $p_j < p_i$.

Suppose part (a) of the Theorem is false; i.e., there exists some $a_k \in A_i \cap A_j$ such that i appears above j on list L_k, and yet

$$p_j < p_i. \qquad (A.4)$$

Conditional on i being active, j cannot implement any idea of her own because i is senior to j on one of her assets, $a_k \in A_j$. In effect, i has the seniority (at least, on asset a_k) to block j: when i is active, j cannot generate any value. Worse, when j too is active, not only is she blocked (by i) from generating value herself, she in turn may block someone else who is active but junior to her from implementing their idea. In short, conditional on i being active, j's presence in the organization cannot raise expected surplus:

$$X_{ij} \leq X_i. \qquad (A.5)$$

Taken together, (A.4) and (A.5) are incompatible with the maximality of V.

To see why, first perform some manipulations on (A.1)-(A.3). Take a convex combination of $1/(2 - p_i)$ times (A.2) plus $(1 - p_i)/(2 - p_i)$ times (A.3). Then substitute for V from (A.1), to yield

$$\frac{p_j(1 - p_i)(p_i - p_j)}{2 - p_i}(X_j - X) \quad + \quad p_i p_j(X_i - X_{ij}) \quad \leq \quad 0. \qquad (A.6)$$

If (A.4) and (A.5) hold, then (A.6) implies

$$X_j \leq X. \qquad (A.7)$$

Notice that (A.5) and (A.7) together say that j is making no contribution to the organization, irrespective of whether i is active or inactive. Indeed, combining (A.1), (A.5) and (A.7), we obtain

$$V \leq (1 - p_i)X + p_i X_i; \qquad (A.8)$$

and the RHS of (A.8) is the expected surplus of the organization with j taken out. Since V is maximal, (A.8) cannot hold as a strict inequality.

But j can make a positive contribution. Suppose her seniority is lowered so that, for each asset $a_{k'} \in A_j$, she is placed at the bottom of list $L_{k'}$. Now she is so junior that she can no longer block anyone. Moreover, despite her junior status, she sometimes contributes $v(A_j)$ (when she is active and no one senior to her is active). The modified organization thus yields an expected surplus strictly in excess of the RHS of (A.8), which contradicts the maximality of V. Part (a) of the Theorem is proved.

Part (b): $p_j = p_i$.

As noted in the text, part (b) of the Theorem follows directly from part (a) when $p_j \neq p_i$. Therefore the only case we need consider is when $p_j = p_i$.

For $p_i = p_j$, suppose part (b) of the Theorem is false. That is, suppose there exists some asset $a_k \in A_i \cap A_j$ such that i appears above j on list L_k, and there exists another asset $a_{k'} \in A_i \cap A_j$ such that j appears above i on list $L_{k'}$. Then i and j block each other.

In particular, consider the event that i is active. j cannot implement any idea of her own (because i is senior on asset a_k). Moreover, when j too is active she blocks i from implementing his idea (because j is senior on asset $a_{k'}$). Hence, conditional on i being active, j's presence in the organization strictly reduces expected surplus:

$$X_{ij} < X_i. \tag{A.9}$$

The symmetric argument holds, reversing the roles of i and j. So

$$X_{ij} < X_j. \tag{A.10}$$

Given $p_i = p_j = p$, say, we can substitute (A.9) and (A.10) into (A.1) to obtain

$$V < (1 - p)^2 X + [1 - (1 - p)^2]\max\{X_i, X_j\}. \tag{A.11}$$

But (A.11) contradicts one of (A.2) or (A.3). Part (b) of the Theorem is proved.

Q.E.D.

Proof of Result 1

Suppose Result 1 is not true. Then there are two cases to consider. Either (a) there exists an optimal organizational form for which $r \geq 1$ and (without loss of generality) $s \geq n - r - s + 1$. Or (b) there exists an optimal organizational form for which $r = 0$ and (again, without loss of generality) $s \geq n - s + 1$, where n is even.

Case (a): $r \geq 1$

Let θ denote the probability that at least one of the most senior $r - 1$ coordinators has an idea (if $r = 1$ then define θ to equal 0):

$$\theta = 1 - (1 - p_2)^{r-1}.$$

Let Y_1 denote the expected surplus generated by the bottom $s - 1$ specialists on asset a_1, conditional on none of the $r + 1$ people senior to them (r coordinators and 1 specialist) having an idea (if $s - 1 = 0$ then define Y_1 to equal zero):

$$Y_1 = [1 - (1 - p_1)^{s-1}]v_1. \qquad (A.12)$$

Let Y_2 denote the expected surplus generated by all the $n - r - s$ specialists on asset a_2, conditional on none of the r coordinators senior to them having an idea (if $n - r - s = 0$ then define Y_2 to equal zero):

$$Y_2 = [1 - (1 - p_1)^{n-r-s}]v_1. \qquad (A.13)$$

Note that given $s \geq n - r - s + 1$, it follows from (A.12) and (A.13) that

$$Y_2 \leq Y_1 < v_1. \tag{A.14}$$

The organizational form yields expected surplus

$$V = \theta v_2 + (1 - \theta)p_2 v_2 + (1 - \theta)(1 - p_2)[p_1 v_1 + (1 - p_1)Y_1 + Y_2]. \tag{A.15}$$

Consider an alternative organizational form. Change the task of the most junior coordinator so that he is a specialist on asset a_2, and make him the most senior of such specialists. Leave everyone else's tasks and seniorities unchanged. This alternative organizational form yields expected surplus

$$V' = \theta v_2 + 2(1 - \theta)p_1 v_1 + (1 - \theta)(1 - p_1)[Y_1 + Y_2]. \tag{A.16}$$

Consider a second alternative. Starting from the original organizational form, change the task of the most senior specialist on asset a_1 so that she is a coordinator, and make her the most junior coordinator. Leave everyone else's tasks and seniorities unchanged. This second alternative organizational form yields expected surplus

$$V'' = \theta v_2 + (1 - \theta)[1 - (1 - p_2)^2]v_2$$

$$+ (1 - \theta)(1 - p_2)^2[Y_1 + Y_2]. \qquad (A.17)$$

Given that the original organizational form is optimal, $V' \leq V$ and $V'' \leq V$. From (A.15), (A.16) and (A.17), these inequalities can be simplified and then combined to eliminate v_2, yielding

$$p_2 v_1 + (1 - p_2)Y_1 \leq Y_2. \qquad (A.18)$$

But (A.18) contradicts (A.14).

Case (b): r = 0 and n even

When r = 0, the organizational form yields expected surplus

$$V = [1 - (1 - p_1)^s]v_1 + [1 - (1 - p_1)^{n-s}]v_1.$$

Now move one of the specialists from asset a_1 to asset a_2. The new expected surplus equals

$$V' = [1 - (1 - p_1)^{s-1}]v_1 + [1 - (1 - p_1)^{n-s+1}]v_1.$$

Given that n is even and $s \geq n - s + 1$, V' is greater than V, which contradicts the optimality of the original organizational form. Q.E.D.

REFERENCES

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

Argyres, N., "Technology Strategy, Governance Structure and Interdivisional Coordination," Journal of Economic Behavior and Organization, 28 (1995), 337-58.

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

Bolton,P., and M. Dewatripont, "The Firm as a Communication Network," Quarterly Journal of Economics, 109 (1994), 809-39.

Bolton, P., and J. Farrell, "Decentralization, Duplication, and Delay," Journal of Political Economy, 98 (1990), 803-26.

Calvo, G. A., and S. Wellisz, "Hierarchy, Ability, and Income Distribution," Journal of Political Economy, 87 (1979), 991-1010.

Chandler, A. D., Jr., Strategy and Structure: Chapters in the History of the American Industrial Enterprise (Cambridge, MA: MIT Press, 1962).

Grossman, S., and O. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," Journal of Political Economy, 94 (1986), 691-719.

Hart, O., and J. Moore, "Property Rights and the Nature of the Firm," Journal of Political Economy, 98 (1990), 1119-58.

Hart, O., Firms, Contracts, and Financial Structure (Oxford: Oxford University Press, 1995).

Keren, M., and D. Levhari, "The Optimum Span of Control in a Pure Hierarchy," Management Science, 25 (1979), 1162-72.

Lando, H., "Communication Costs, Individual and Collective Ownership," mimeo, Copenhagen Business School, January 1998.

Maskin, E., Y. Qian, and C. Xu, "Incentives, Information and Organizational Form," mimeo, Harvard University, Department of Economics, February 1999.

Radner, R., "Hierarchy: The Economics of Managing," Journal of Economic Literature, 30 (1992), 1382-1415.

Rosen, S., "Authority, Control, and the Distribution of Earnings," Bell Journal of Economics, 13 (1982), 311-23.

Segal, I., "Communication Complexity and Coordination by Authority," mimeo, Department of Economics, University of California at Berkeley, 1998.

Wernerfelt, B., "The Structure and Scope of Firms and Economies of Scale in Contracting," mimeo, MIT Sloan School of Management, November 1992.

Williamson, O. E., "Hierarchical Control and Optimum Firm Size," Journal of Political Economy, 75 (1967), 123-38.