Corporate Finance, the Theory of the Firm, and Organizations

Patrick Bolton and David S. Scharfstein

In his classic 1937 article, “The Nature of the Firm,” Ronald Coase asked one of the most fundamental questions in economics: Why are there firms? Or, in Coase’s words, “If production could be carried out without any organisation at all, well might we ask, why is there any organisation?” Initially, few economists seemed interested in answering Coase’s question, but it is now one of the central questions in economic theory, industrial economics and management strategy.

Five years before the publication of Coase’s article, Adolf Berle and Gardiner Means (1932) published their now famous book, The Modern Corporation and Private Property. Berle and Means documented that a large portion of corporate assets were controlled by managers with minimal ownership stakes in their firms. Like Coase, they raised a fundamental question about the functioning of firms: Will corporate managers continue to act in the interest of investors despite their small ownership stakes? Understanding the agency costs stemming from the “divorce of ownership from control” is now the central issue in corporate finance, and has been for some time.

Over half a century after these famous works, we have some of the answers to these questions. Williamson (1975), Klein, Crawford, and Alchian (1978), Grossman and Hart (1986), and Hart and Moore (1990) have made substantial progress answering Coase’s question. Alchian and Demsetz (1972), Mirrlees (1976), Jensen and Meckling (1976), and Holmström (1979) have gone a long way towards un-
derstanding the implications of Berle and Means's observations about the separation of ownership and control.

What we lack and what we need is a more unified theory of the firm based on the insights of Coase and Berle and Means. We believe that one cannot fully answer Coase's question without understanding the issues raised by Berle and Means. That is, a fully satisfactory theory of the firm must recognize that the managers of the firm are not its owners, but are the agents of the firm's shareholders. Yet, the Coasian literature—notably Grossman and Hart (1986) and Hart and Moore (1990) which provided the first complete, formal model of integration—has modeled the integration decision between enterprises that are owned and managed by the same person. Despite the fundamental importance of this work, it is not so clear how one would use this model to understand, for example, the acquisition by a large multidivisional firm of one of its suppliers. Managers don't own their companies' assets, though they may control their use. How then might we think about the boundaries of the firm when managers control assets but don't own them, when they are subject to the sort of agency problems identified by Berle and Means?

At the same time, a complete understanding of the corporate ownership and control issues discussed by Berle and Means requires some notion of what a firm is; in other words, an answer to Coase. This literature, unlike the theory of the firm, recognizes that at least one actor in the firm, the chief executive officer (CEO), is the agent of shareholders. But it ignores the other agents in the firm and, as a result, delivers little in the way of a theory of organizations. Can we better understand the conflict between shareholders and management, if we recognize that there is more to management than the CEO?

In this article, we argue that the time has come to begin to integrate the Coasian view of the firm—which is concerned with the interactions between owner-managers—and the Berle and Means perspective—which emphasizes the separation of ownership and control in most corporations. To illustrate the importance of integrating both perspectives, we will revisit the Coasian literature's favorite example of a vertical merger, the acquisition of Fisher Body by General Motors in 1926.¹ This example has been interpreted and reinterpreted by numerous authors in the Coasian literature, but all implicitly assume that each firm was run by its sole owner—that there was no separation of ownership and control. A more realistic account of the merger must recognize that General Motors was not owner-managed and that it was a large organization run by professional (non-owner) managers. The real result of the merger was that Fisher Body became part of the complex multidivisional organization called General Motors, overseen by GM's chief executive officer, Alfred P. Sloan, and his professional staff at corporate headquarters. Thus, it seems to us that to understand the effect of the merger one must understand

¹ Many papers cite this example, but interestingly Coase himself takes issue with the interpretation. In fact, Coase (1988) takes issue with much of the work that followed from his initial paper.
both the complex interplay between divisional (non-owner) managers and corporate headquarters, and the relationship between corporate headquarters and outside investors. Our main point is that the firm should be understood as being comprised of (at least) two tiers of agency relationships: one between investors and corporate headquarters and the other between corporate headquarters and the divisions. This is a more subtle and complex view of the firm than is found in the literatures of Coase and Berle and Means.

Unfortunately, we know relatively little about how two-tiered (or multi-tiered) organizations like these work. To be sure, this article does more to highlight the problem than to solve it. However, we do review some recent theoretical and empirical work that gets at one of the most important activities of two-tiered organizations, the allocation of capital across divisions. This literature suggests that when there are many agents competing for capital within an organization, and when they are agents, not owners, the allocation of capital need not be efficient. In fact, headquarters may practice a kind of internal socialism (egalitarianism?) in which investment is equalized across divisions regardless of their investment opportunities. These models and empirical results suggest that to understand the boundaries of the firm one must understand how agents interact in complex organizations. Clearly, we are at only the earliest stages of such an understanding.

**Coase and Integration**

Coase (1937) begins his discussion of the costs and benefits of integration by arguing that some transactions between firms are potentially inefficient because it is difficult to write contracts that fully specify what should happen in future situations that are hard to foresee or describe. Firms emerge in response to the inefficiencies arising from these incomplete contracts because a firm’s owner can simply direct employees to do the right (that is, the efficient) thing. Against these benefits of integration, there are costs, chief among them “diminishing returns to management” (p. 394) and “waste of resources” (p. 395).

Coase’s argument was an extremely useful starting point, but it left many questions unanswered. What are the costs of writing incomplete contracts? Why might diminishing returns to management exist? Do large firms indeed waste more resources? The literature that followed—notably Williamson (1975, 1985) and Klein, Crawford and Alchian (1978)—began to offer some answers. The difficulty with incomplete contracts is that they enable parties to take advantage of ambiguities in the contract to their own advantage, particularly when buyer and supplier have tailored their products to each others’ needs, and thus are somewhat locked into the existing relationship. In the language of Klein, Crawford and Alchian, incomplete contracts raise the possibility of *hold-up problems* when there are *relationship-
specific investments.\textsuperscript{2} (Or, as Williamson puts it, incomplete contracts allow for opportunism after investment decisions are made in situations of asset specificity.) The possibility of hold-up discourages contracting parties from making otherwise efficient relationship-specific investments, and can lead to inefficiencies when contracts are renegotiated. This line of argument implies that integration into a single corporation should occur when renegotiation costs are high and when important relationship-specific investments exist.

While these theories are clear on the costs of transacting in the marketplace with incomplete contracts, they are somewhat less clear on the benefits and costs of integration. Here, Grossman and Hart (1986) and Hart and Moore (1990) have provided some important answers and a more formal theoretical framework—sometimes referred to as the property rights approach—in which to think about the issues. They view ownership of a firm as giving the owner residual control rights over the use of the firm’s non-human assets; that is, the right to use assets in whatever way the owner likes unless otherwise prohibited in a contract. In particular, and importantly, the owner of an asset has the right to exclude others from its use. In this paradigm, ownership is synonymous with control.\textsuperscript{3}

For example, consider two firms, A and B. To be concrete, suppose that A is a textbook publisher and B is a printer. Suppose they have a long-term supply contract, and that something happens that is unanticipated in the initial contract; for example, it becomes crucial to use four colors instead of the standard two colors. The Grossman-Hart-Moore paradigm assumes that the contract is always successfully and efficiently renegotiated so that four colors are used. Because each party owns its assets, each can threaten not to renegotiate. As a result, it is likely that they will share the benefits of moving from two to four colors. Contrast this to a situation in which the publishing house owns the printing press. There is still a manager in charge of running the printing press, but now the manager of the press is no longer the owner and does not control the use of the press. Instead, the manager of the press is an employee of the publishing house, taking instructions from its manager/owner on the number of colors to be used. As an employee, the manager has no bargaining power and need not be paid off to implement the change (provided that the manager’s human capital is also replaceable).

\textsuperscript{2}Anyone who has ever renovated a home knows firsthand what this means. After the project has begun, situations arise that are unforeseen in the original contract. The builder is in a great position to extract a tidy sum from the homeowner to make even the smallest changes (the hold-up problem) because the homeowner would have a hard time using someone else (because of relationship-specific investments).

\textsuperscript{3}Notice the contrast with the emphasis of Berle and Means (1932) on the separation of ownership and control. This is because Berle and Means implicitly define ownership as claims on residual cash flows—what is available after paying other stakeholders. By contrast, Grossman and Hart define ownership as residual control rights—the right to make decisions when not specified in a contract. Thus, the Grossman-Hart-Moore line of thought would call shareholders “owners” because they have the voting power to determine how assets are deployed, whereas Berle and Means think of shareholders as “owners” because they get residual cash flows.
Note that the ownership allocation in the Grossman-Hart-Moore paradigm has no effect on the level of efficiency achieved; in both cases, the model assumes that the efficient outcome (four colors) is achieved through renegotiation. However, because ownership affects the division of the benefits from renegotiation, it can affect the incentives of the two parties to make relationship-specific investments. More concretely, the publishing house may decide not to tailor its publishing software to the specifications of the printing press when it does not own the printing press because it thereby weakens its position in future anticipated renegotiations with the printer. By contrast, if the publisher owns the press, the publisher does not have to worry about sharing the benefits of its software investments with the printer because the publisher doesn't give away anything in renegotiation. Ownership encourages parties to make more relationship-specific investments.

The model therefore implies that A should own B when A's relationship-specific investments are considerably more important than those of B; that B should own A when the reverse is true; and that there should be no integration when both make important relationship-specific investments. In sum, the important contribution of Grossman, Hart and Moore is to explain how ownership affects economic decisions and to provide a framework in which the costs and benefits of integration are clearly identified.

**Berle and Means and the Separation of Ownership and Control**

There are two main contributions of *The Modern Corporation and Private Property*, by Berle and Means (1932). The first is to document that, by 1930, large corporations accounted for a sizable share of corporate activity, and that in many of these companies control and ownership were separated. According to Berle and Means, the 200 largest non-financial corporations accounted for roughly half of all corporate assets, and 65 percent of these companies were controlled by management with small ownership stakes. The second contribution was to argue that the separation of ownership and control may lead managers to pursue their own objectives at the expense of owners. They quote Adam Smith, who raised a similar concern about companies with dispersed ownership in *The Wealth of Nations* (p. 304), "The directors of such companies... being the managers rather of other people's money than of their own, it cannot well be expected that they should watch over it with the same anxious vigilance with which the partners in a private copartnery frequently watch over their own."

---

*An alternative to no integration which seems to provide equivalent protections to both managers is integration with joint ownership by both managers. In this situation, which resembles a partnership, both managers appear to have equal bargaining power from their respective ownership stakes. Depending on the specific set-up, joint ownership may actually be better or worse than non-integration (Hart, 1995; Hausmann, 1996).*
A large and ever-growing literature has tried to flesh out these ideas. Jensen and Meckling (1976) is, perhaps, the starting point for the modern literature. They show that if managers own only a portion of the company’s equity, they will over-indulge in perquisites since they get all of the benefits, but bear only a portion of the costs. Jensen and Meckling then analyze ownership and capital structures that mitigate these “agency costs.” For example, to the extent that managers need external financing to fund capital investment, they will rely more on debt financing, because it allows managers to retain a greater portion of the company’s equity and gives them the incentive not to consume excess perks. Much of the subsequent theoretical literature builds on Jensen and Meckling’s insight by spelling out different kinds of agency costs and other mechanisms by which such agency costs can be mitigated. We cannot hope to provide a full picture of the many strands of this literature; Harris and Raviv (1991) and Shleifer and Vishny (1997) are excellent surveys.

One weakness of this perspective on agency costs is that it fails to explain why a firm’s capital structure must be used to provide incentives when managerial compensation packages offer a more direct and possibly cheaper way of doing so—say, by linking compensation to the firm’s stock price. It also cannot explain the allocation of control implicit in capital structure decisions; for example, why creditors only obtain control when the firm is in financial distress, or why shares have both dividend and voting rights. Such aspects of corporate finance are the main focus of a more recent literature which puts the spotlight on the effect of capital structure on control allocations.

Interestingly, the Grossman-Hart-Moore property rights approach has been very useful for understanding the control issues related to the choice of capital structure. Indeed, this perspective sees debt financing as a way of allocating control in a “state-contingent” fashion, with equity holders or managers retaining control in non-default states and creditors taking control in default states. Aghion and Bolton (1992) have shown that the control shifts that come with debt instruments can be an optimal mechanism in mitigating agency conflicts. In addition, Bolton and Scharfstein (1990) and Hart and Moore (1998) have shown that if the debtor is able to divert cash flow to itself, it is optimal for the creditors to take control of the firm’s assets in default, to limit the debtor’s incentive to divert cash flow.

Another application of the Grossman-Hart-Moore paradigm is to explain why the separation of ownership and control documented by Berle and Means may be optimal in some cases (Burkhart, Gromb and Panunzi, 1997). Diffuse equity ownership, as Berle and Means argued, gives managers effective control of the firm. As already noted, this raises concerns because it allows managers to run the firm to their own benefit, possibly at the expense of investors. However, the ability of man-

---

5 Ironically, while the Grossman-Hart-Moore paradigm has been used to explain the separation of ownership and control, there has been no real attempt to examine the implications of this separation on the integration decision initially addressed by the model.
agers to derive such benefits has a silver lining; it gives them the incentive to invest in firm-specific human capital.\textsuperscript{6}

In addition to this theoretical work, there is a sizeable empirical literature characterizing the nature of agency conflicts in corporations and analyzing the effect of various corporate governance arrangements on those conflicts, reviewed in some detail in Shleifer and Vishny (1997). Much of the evidence comes from analyzing managerial behavior in corporate control transactions. CEOs tend to resist takeovers, even when acquirers offer large premiums. Moreover, CEOs acquire companies even if doing so reduces shareholder value. However, both value-destroying takeover resistance and acquisitions seem to be less likely when top managers have larger ownership stakes, suggesting that agency conflicts really are at the heart of this behavior (Lewellen, Loderer and Rosenfeld, 1985; Walkling and Long, 1984).

One can interpret certain corporate governance arrangements and capital structures as attempts to mitigate agency problems. For example, because large creditors and large shareholders have a lot at risk, they will have incentives to monitor and control management. Similarly, management buyouts, in which managers use debt to finance an acquisition and become large shareholders, can also be seen as responses to agency problems. The high leverage used to finance these deals limits the ability of managers to overinvest, and the larger percentage ownership of the company gives managers greater financial incentives to perform.\textsuperscript{7}

The agency literature has considerably improved our understanding of the kinds of problems that can arise when ownership and control are separated, and how they might be mitigated by designing the firm’s capital structure or the manager’s compensation. However, its scope is limited by an excessively narrow view of the firm. The Berle and Means (1932) perspective is to identify the firm with a single individual, the CEO, and a single productive asset. But firms are actually organizations with many different kinds of employees; with internal capital and labor markets; with formal management structures and decision-making procedures; and with a wide variety of rules governing the allocation of inputs, capital, and the hiring and promotion of employees. To develop a better understanding of the conflict between the firm’s management and its investors, this perspective must be broadened to introduce the Coasian themes of internal organization and interaction among managers inside the firm.

\textsuperscript{6}Shleifer and Vishny (1989) take the opposite view. They argue that because of the benefits that managers receive from being in control, managers will make inefficient investment decisions with the aim of making themselves indispensable—and thus further entrenching themselves.

\textsuperscript{7}The effect of management buyouts on incentives is less clear than this explanation may suggest. While managers have larger percentage stakes in management buyouts, in many cases the dollar amount they have at risk is less because of the smaller equity base in these highly leveraged transactions (Kaplan and Stein, 1993).
A Reexamination of the Merger between General Motors and Fisher Body

One of the best-known examples of integration in response to hold-up and asset-specificity problems is the case of General Motors’ 1926 acquisition of its car body supplier, Fisher Body, which was first discussed in the Klein, Crawford and Alchian (1978) paper that has been mentioned several times earlier. It is worth taking another look at this example viewed from the perspective of Coase and his descendants, as well as from the perspective of Berle and Means and their descendants. Revisiting this case will make clear some of the limitations of both theories, and will highlight the need for a theory that integrates the vertical integration questions posed by Coase with the corporate finance questions posed by Berle and Means.

As told by the Coase descendants, the General Motors/Fisher Body story goes something like this. In 1919, General Motors signed a long-term exclusive contract with Fisher Body to supply General Motors with car bodies. Fisher Body had to tailor its production to GM’s needs, forcing it to make GM-specific investment in stamping machines and dies. To limit the ability of General Motors to threaten to use another supplier after Fisher made these specific investments, General Motors agreed to a ten-year exclusive-dealing clause requiring General Motors to buy only from Fisher at presupposed prices. Over the next few years, demand for closed metal car bodies grew rapidly and unexpectedly. Fisher Body became even more critical to General Motors, and was in a position to hold up General Motors. For example, General Motors wanted Fisher Body to locate its plant near General Motors’ plants, but Fisher allegedly refused to do so. To prevent future hold-ups, General Motors bought Fisher Body in 1926.

In the Grossman-Hart-Moore paradigm, by acquiring Fisher Body, the owner/manager of General Motors gains the residual control rights over the Fisher Body assets. This means that the owner/manager of General Motors can get much of the surplus from renegotiation with the Fisher manager should contingencies arise that are not covered by the terms of the contract. As a result, General Motors has an incentive to invest in relationship-specific investments with Fisher, such as tailoring its chassis production to Fisher’s frame production.

If this model is taken literally, there would be someone at General Motors who makes all the decisions and receives all of the General Motors’ profits, and someone at Fisher Body who makes all of the decisions and receives all of Fisher Body’s profits. In 1926, the president of General Motors was Alfred P. Sloan, and the president of Fisher Body was William A. Fisher. So in this framework, integration means that Sloan gets to decide how Fisher’s assets are used, which induces Sloan to tailor his assets to Fisher’s, but reduces Fisher’s incentives to do so.

However, viewing the General Motors/Fisher Body case from the perspective of Berle and Means (1932) raises some different issues. Fisher Body was owned and
controlled by the Fisher family; as a result, agency costs were likely to be small. By contrast, in 1926, ownership and control were separated at General Motors. General Motors was run by Sloan, who had joined the firm in 1916 when his Hyatt Roller Bearing Co. was purchased by the United Motors division of General Motors. Sloan ran the United Motors division and then became the chief operating officer under Pierre du Pont until 1924, when he was appointed chief executive officer. We do not know exactly how much equity Sloan held in 1926, but he is listed as owning 0.82 percent of General Motors stock in 1939 (Temporary National Economic Committee, 1940). He is unlikely to have held much more in 1926. All officers and directors together held 6.26 percent of General Motors stock. From the Berle and Means perspective, since ownership was separated from control at General Motors, agency costs were likely to have been substantial.

The main implication of the General Motors/Fisher Body merger in the Berle and Means paradigm is that now Fisher Body is run by a manager and agent of the owner (Sloan) instead of by the owners themselves (the Fisher brothers). Thus, there is scope for increased agency costs and, as a result, greater inefficiencies. One might even speculate that the merger itself was the manifestation of agency costs, a desire to build an empire by a manager who does not bear the full costs of the decision.

Both the Coase and the Berle- Means perspectives may highlight potentially important aspects of the merger, but both also fail to capture other potentially important elements. The Berle and Means perspective lacks a real theory of the effects of the merger on the combined organization. A firm is nothing more than a manager—in this case Sloan—who makes decisions for General Motors and Fisher Body, perhaps at the expense of investors. But General Motors was a much more complex organization than would be suggested by CEO-centric agency theories. After all, the merger would likely affect more than just the CEO's incentives, but also those of managers deeper down in the organization.

The Coasian perspective, by contrast, is a starting point for thinking about the effect of the merger on the combined organization; it focuses on bargaining problems among managers, and makes arguments about whether those problems are better resolved inside a firm or as a market interaction between firms. In this view, a firm is a collection of managers bargaining for their piece of the pie. But at least in its simple form, the theory envisions managers who are also owners, and who are driven purely by the profit motive. This obviously leaves out important agency considerations. For a truly compelling theory of the firm one would need to meld the Coasian perspective on integration and organizations with the Berle and Means perspective on agency costs.

---

8 A little-noted fact is that at the time of the merger, General Motors already owned 60 percent of Fisher Body, a stake that had been acquired in 1919.
A more detailed look at General Motors organization at the time of the merger makes the need for this sort of framework apparent. We know a lot about General Motors’ organization circa 1926 from *Strategy and Structure*, Alfred Chandler’s (1962) landmark study of the evolution of the modern form of corporate organization. Chandler credits Sloan with being the principal architect of the multi-divisional form of organization that most companies now use. In the M-form, as Williamson (1985) calls it, the role of corporate headquarters is to design broad strategies for the corporation, and the role of the operating divisions is to implement their parts of those strategies with little interference from headquarters. As Chandler (1962, p. 161) put it, Sloan’s structure involved “decentralized operating responsibilities and centralized policy formation . . . Sloan firmly believed that divisional independence encouraged initiative and innovation. At the same time, the activities of these divisions had to be coordinated and controlled in the interest of the corporation as a whole.”

This structure replaced one in which there was little coordination across divisions, and no clear lines of authority. Others had tried to resolve these problems, notably Pierre du Pont who stepped in to take control of the company during General Motors’ financial difficulties in 1920. But they ran into severe organizational obstacles and “failed to make General Motors into more than an expanding agglomeration of different companies making automobiles, parts, accessories, trucks, tractors, and even refrigerators” (Chandler, 1962, p. 127).

This suggests that the effect of the General Motors/Fisher Body merger depended critically on the organizational structure of General Motors. A merger of Fisher into the old, disorganized General Motors might have had little value. But perhaps the new structure created by Sloan enabled the company to get the benefits of increased coordination without dampening Fisher’s initiative and innovation—which was one of the stated goals of the new organizational form.

Sloan’s reorganization of General Motors streamlined the company into four operating groups: cars and trucks, accessories, parts, and miscellaneous. Fisher Body became part of the cars and trucks group, joining the Chevrolet, Cadillac, Buick, Olds, and Oakland car lines. Each of these units reported to the group vice-president in charge of cars and trucks. In turn, the four group vice-presidents served on the company’s Executive Committee along with Sloan and other top management in headquarters.

The placement of Fisher into the cars and trucks division makes clear some of the limitations of the Grossman-Hart-Moore paradigm in this context. Presumably, Fisher would not be negotiating directly with Sloan about the supply of car bodies to Cadillac and the other car lines. The negotiations would likely occur between the car lines and Fisher Body and would be overseen by the group vice president in charge of cars and trucks.⁹ This differs from the Grossman-Hart-Moore frame-

---

⁹ Incidentally, the Cadillac division was run by one of the Fisher brothers at the time of the merger.
work in three ways. First, the parties to the bargaining are not owners, but rather division managers—agents of the true owners. In addition, they are not bargaining over the profits they will pocket themselves, but rather for their division’s share of corporate resources to be used mainly for investment. Finally, and importantly, there is a third party in the bargaining, the group vice president or corporate headquarters, who has ultimate control over the distribution and use of these resources.

The Grossman-Hart-Moore framework has a hard time explaining the presence of third parties such as headquarters and group vice-presidents. It predicts that control should be allocated to parties whose relationship-specific investments are most important to the relationship. Yet headquarters is given control, even though it does not really make such investments. But once one integrates the Berle and Means perspective into the Grossman-Hart-Moore framework, it is easier to see why an institution like corporate headquarters emerges. Again, the history and evolution of General Motors prior to the merger with Fisher Body helps clarify the point.

Before Alfred Sloan’s regime at General Motors, there was no corporate headquarters acting as an intermediary between and coordinator of the various divisions of General Motors. The Executive Committee was composed of division managers. According to J.L. Pratt, Chairman of the General Motors Appropriations Committee (quoted in Chandler, 1962, p. 127), “When one of them had a project, why he would get the vote of his fellow members; if they would vote for his project, he would vote for theirs. It was a sort of horse trading.”

As a result of this type of horse trading, spending and costs grew out of control and Sloan’s creation of corporate headquarters was largely an attempt to bring tighter control on the allocation of resources inside General Motors. Indeed, Sloan saw headquarters as mainly a coordinator of the activities and investments of General Motors’ divisions. In his mind, a merger with Fisher Body was called for essentially to bring Fisher body under the umbrella of the newly created administrative structure and thus take full advantage of the coordination benefits of General Motors’ general office.

What this discussion suggests is that we need a framework that builds on Berle and Means and Coase, but that recognizes: 1) a role for corporate headquarters in the bargaining process among managers; 2) that managers both at corporate headquarters and divisions are not owners but rather agents of shareholders; and 3) that bargaining takes place over the allocation of corporate resources (broadly defined) not manager’s compensation (narrowly defined).

The fundamental consequence of integration is to bring all divisions under the umbrella of a single administrative structure. Integration results in greater centralization of decision-making. This includes setting corporate strategy, allocating corporate resources and raising capital. This more centralized structure could be more or less efficient than the more decentralized structures of two stand-alone entities. In truth, we know little about the relative efficiency of centralized and decentralized organizational structures. However, we are at the early stages of understanding the differences in how organizations function in one crucial activity,
the allocation of capital. We review this research in an attempt to illustrate the importance for understanding organizations of the three elements listed above.

**Corporate Headquarters and the Allocation of Capital**

As we have discussed, one of the key roles of corporate headquarters is the allocation of capital.\(^\text{10}\) There are several competing views about the effectiveness of corporate headquarters in allocating capital, or what we will refer to as “internal capital markets.” The first, expounded by Williamson (1970) and Alchian (1969) among others, is that internal capital markets are more effective than external capital markets because headquarters is a more informed provider of capital and can more effectively ensure that its funds are used appropriately. A second view, which is a variant of the Modigliani-Miller theorem, holds that internal and external capital markets will result in the same resource allocation. A third view—now the favorite of just about everyone—holds that internal capital markets are less efficient than external capital markets because they replace the profit-based decision-making of investors with the bureaucratic decision-making of corporate executives.

It is safe to say that these competing assertions are not based on clearly articulated models of the resource allocation process. The arguments used to support these theories are typically not much more developed than the initial Coasian arguments on product-market integration. Take, for example, the Williamson-Alchian argument about the value of internal capital markets. They have in mind a comparison of an internal capital market in which corporate headquarters makes the capital allocation decision, and an external capital market in which there are many small debt and equity investors who provide capital to the firm. In this situation, it might be reasonable to suppose that small debt and equity investors lack incentives to monitor management and will be ill-informed about investment opportunities. Perhaps then internal capital markets organized by corporate headquarters really are better. But suppose that the external capital comes from a single bank, rather than many small investors. Here, the assertion that a bank is less informed and monitors less is not obviously true.\(^\text{11}\) In fact, a wealth of evidence exists that close banking relationships mitigate information and agency problems typical of more arms-length capital markets.\(^\text{12}\) What then is the crucial difference between an in-

---

\(^{10}\) Gertner (1996) models corporate headquarters as an arbitrator of disputes across managers when there are potential gains from competition. Holmström and Tirole (1991) model the role of corporate headquarters in transfer pricing.

\(^{11}\) One might ask who monitors the bank; that is, why should the bank be informed? For an answer, see Diamond (1984).

\(^{12}\) One of the symptoms of information and agency problems is that firms tend to cut investment when they have cash flow shortfalls because external financing from “arms-length” investors is more costly than internal financing (Fazzari, Hubbard and Petersen, 1988). However, Hoshi, Kashyap and Scharfstein (1989) have shown that Japanese firms with close bank ties tend to cut investment less in response
ternal capital market, run through corporate headquarters, and an external capital market like a bank?

One important difference seems to be that the relationship between a bank and a firm is governed by a debt contract, whereas there are less formal or different kinds of contracts between headquarters and a division. However, this need not be the case. Why can’t headquarters put its division managers on a debt-like incentive contract, in which division managers are supposed to make a fixed payment to headquarters and are able to keep whatever is left over? Indeed, if debt is an optimal contract between a bank and a firm, shouldn’t it be the optimal contract between headquarters and the division?

These questions are reminiscent of Coase’s fundamental question: What is the difference between a contract between firms and a contract within firms? Here we are asking: What is the difference between a contract to supply capital between a bank and an independent firm and a contract to supply capital between headquarters and a division within the firm? Just as the Grossman-Hart-Moore paradigm provided a useful answer to Coase’s question, it can provide a useful answer to this question, as well. In the spirit of this perspective, Gertner, Scharfstein and Stein (1994) have pointed out that the difference between internal and external capital markets is that headquarters owns and controls the assets of the division. In contrast, the bank only owns and controls the assets of the firm should it default—and even then may not control the assets due to bankruptcy law protection. As the Grossman-Hart-Moore paradigm emphasizes, ownership comes with the right to do with the assets as the owner pleases. This has two consequences for an internal capital market, one which makes it more attractive than an external capital market and one which makes it less attractive.

On one hand, an internal capital market provides greater monitoring incentives than an external capital market. This is essentially what Williamson (1970) and Alchian (1969) claimed, but the reasons are more grounded in the theory of the firm. To see why, suppose that headquarters comes up with an idea that would improve the value of its division’s assets. As the owner of those assets, it can implement the idea without the manager’s approval. Since it doesn’t need approval, headquarters will get all (or most) of the rents from this improvement. This gives headquarters strong incentives to do this sort of monitoring. By contrast, a bank might come up with an equally good idea, but cannot force the firm to implement the idea. Instead, it would have to coax the firm into doing so by sharing with it some of the value of the improvement. The result is that a bank has less incentive to monitor than corporate headquarters.

There is, however, a cost of an internal capital market. The very same control that gives headquarters the incentive to monitor has an adverse effect on the in-

to cash flow shortfalls than firms that do not have close bank ties. This finding has been confirmed in a variety of other settings including the United States at the turn of the century, when J. P. Morgan held large financial stakes in many companies (Ramirez, 1995).
centives of division managers to act in an innovative or entrepreneurial fashion. Since headquarters has control, it has the ability to extract rents from division managers should they want to do something that is not covered by their employment contract. Control by headquarters discourages managerial entrepreneurial activity relative to bank lending.

This perspective gets at some of the differences between internal and external capital markets, yet it remains an overly simplistic account of an internal capital market in two ways. First, there is no capital allocation decision across divisions. How does corporate headquarters allocate capital when there is competition among divisions for potentially scarce resources? Does it play “Robin Hood,” in the language of Stein (1997), taking from cash-rich divisions with poor investment opportunities and giving to cash-poor divisions with good investment opportunities? Or does it practice a kind of (dysfunctional) “socialism” by reallocating resources regardless of the ability of divisions to use those resources productively? A second and potentially more important shortcoming is that our discussion so far does not account for potential agency problems at headquarters. Our analysis of the difference between corporate headquarters and large banks could apply equally well to a large shareholder contracting with a manager of the firm it owns. To describe corporate headquarters adequately we need to allow for the fact that the chief executive officer is an agent of shareholders, and the division managers are agents of headquarters. We need a two-tiered agency model.

The Grossman-Hart-Moore paradigm remains useful in analyzing a two-tiered agency model of this sort. The first point to notice is that, while headquarters is not the owner, it has effective control of the firm, in the sense that it can control the investment of its divisions, giving more funds to some and less to others. In principle, it can even take cash from one division and give it to another. By contrast, there is no way that a bank can prevent a firm from going to another bank to fund an investment. Nor can a bank take cash from one firm and give it to another. Thus, the exclusive funding relationship divisions have with headquarters gives headquarters the ability to allocate capital as it sees fit, whereas a bank is constrained by the ability of firms to seek funding elsewhere. The question then is whether headquarters will allocate capital efficiently, in the so-called “Robin Hood” approach, or inefficiently, as in (dysfunctional) socialism.

When one recognizes that headquarters is itself an agent of investors, the possibility that it will make inefficient allocation decisions becomes a real worry. However, there may be reasons to believe that headquarters plays the role of Robin Hood, as Stein (1997) has pointed out. Suppose the main agency problem is that managers—whether at headquarters or divisions—like to invest; the larger the empire they run, the more “rents” or “private benefits” they get. It is also reasonable to suppose that headquarters is more informed than external investors, like a bank, for the reasons described above. As long as headquarters gets some benefit from funding profitable projects, it will invest in the right projects. After all, headquarters may like running a large empire, but it doesn’t particularly care in which division
the capital goes. If it can put the capital in the division with more favorable investment opportunities, all the better. Thus, it is not obvious that agency problems at the top induce inefficient capital allocation decisions at lower levels.

However, things may not work out this well. The divisions in Stein’s (1997) model are rather passive; they make no attempt to influence the capital allocation decisions of headquarters. But of course, the real question is how well will internal capital markets work when there is internal politicking for resources? We believe that corporate politics may be the ultimate reason that internal capital markets do not work so well. The problem is that because the marginal product of the managers of weak divisions are relatively low, they are willing to spend more time trying to get organizational rents (Meyer, Milgrom and Roberts, 1992), or increase their own compensation, perhaps by trying to improve their value outside the firm (Scharfstein and Stein, 1998). Moreover, headquarters might prefer to compensate managers of these divisions with an excessively large capital budget rather than with cash—since headquarters is an agent of shareholders it does not bear the full costs of capital misallocation (Scharfstein and Stein, 1998).

Ultimately, the question of whether internal capital markets work well is an empirical one. Not long ago, the common view was that they worked well. For example, the Boston Consulting Group long propounded the view that firms ought to assemble a portfolio of businesses in which the cash flows of the highly profitable, slow growth businesses (the “cash cows”) would be used to finance the growth of the more cash needy, but promising businesses. Indeed, the conglomerate merger wave of the 1960s and early 1970s was partly fueled by the belief that internal capital markets were more efficient than external capital markets. There is some evidence that during this period, announcements of diversifying acquisitions tended to raise the acquirer’s stock price (Matsusaka, 1993). It also seems that stock prices of acquirers went up even more when cash rich acquirers purchased cash poor firms, suggesting that the acquisitions were viewed as mitigating capital market inefficiencies (Hubbard and Palia, 1998).15

However, the more recent empirical literature seems to suggest that internal capital markets don’t work very well, and management consultants now routinely recommend focusing on one core business rather than many unrelated ones. Much of the negative evidence comes from examining the diversified conglomerates that were assembled during the 1960s and 1970s. Since allocating capital is probably the most important activity of headquarters in a diversified conglomerate, it’s a good place to look to understand the workings of an internal capital market—and many of these conglomerates were dismantled in takeovers and corporate restructurings of the 1980s (Comment and Jarrell, 1995). Moreover, acquisitions of companies unrelated to the acquirer’s core business were more likely to be divested (Kaplan

15 However, Servaes (1997) still finds that conglomerates were worth less than a portfolio of stand-alone firms in the same industries.
and Weisbach, 1992). Finally, when non-core businesses were divested following hostile takeover attempts they were often sold to competitors in the same line of business (Bhagat, Shleifer and Vishny, 1991). Consistent with these findings, the stock market now appears to be skeptical of unrelated acquisitions and is more receptive to related acquisitions. In the 1980s, stock prices of acquiring firms fell on the announcement of diversifying acquisitions (Servaes, 1996) and rose on the announcement of sales of non-core divisions (Comment and Jarrell, 1995).

This evidence seems to suggest that diversified firms might to be worth less than the sum of their parts. Indeed, if one values business segments of multidivisional firms based on industry-wide multiples of assets, sales, or operating income, one finds that the combined, imputed value of these segments is 10–15 percent greater than the value of the whole company (Lang and Stulz, 1994; Berger and Ofek, 1995; Lin and Servaes, 1997). Lamont (1997) provided evidence that investment behavior in conglomerates isn’t quite like that of stand-alone businesses. In 1986, oil prices halved, dramatically reducing the cash flows of oil companies. Not surprisingly, they cut investment in their oil business. But many of the oil companies had non-oil businesses—such as Mobil’s Montogomery Ward department store chain—and they cut their investment in these businesses as well. They even cut investment in petrochemical businesses which benefit from an oil price decline, because oil is an input for those businesses. Thus, Lamont showed that there is a meaningful internal capital market in which the resources of one division are used to finance others. However, he was not able to address whether the internal capital market worked well or not.14

Berger and Ofek (1995) were perhaps the first to pin the underperformance of conglomerates on poor investment decisions. They showed that conglomerates that invest more in divisions with poor investment opportunities tend to trade at bigger discounts to their break-up value. They also found that conglomerates with larger discounts relative to their stand-alone values were more likely to be busted up and to increase their focus. Scharfstein (1997) looks in more detail at the investment behavior of 165 conglomerates operating in unrelated lines of business in 1979, just before the breakup of many conglomerates in the 1980s. Divisions in industries with relatively poor investment opportunities tended to invest more than their stand-alone industry peers, while divisions in industries with relatively good investment opportunities tended to invest less than their industry peers. Capital misallocation is more pronounced in conglomerates where management has small ownership stakes; in contrast, where headquarters has strong financial incentives, conglomerate divisions behave more like stand-alone firms.15 Thus, capital misallocation appears to be tied to agency problems at the headquarters level. Finally, Rajan, Servaes and Zingales

14 Shin and Stulz (1996) extended Lamont’s work by looking at all multidivisional firms. They too found that cash flow generated by one division affects the investment of other divisions, principally small ones.

15 In addition, firms with low ownership are more likely to be diversified in the first place (Denis, Denis and Sarin, 1997).
(1998) show that capital misallocation of this type seems to be more pronounced in firms with very different investment opportunities across divisions.

While much of the recent empirical work is skeptical of the value of internal capital markets, it is important to keep in mind that not all conglomerates function poorly. Economists have long been singing the praises of General Electric (Alchian, 1969), which produces products as diverse as refrigerators, light bulbs, aircraft engines, and television news. Baker and Montgomery (1994) have identified a number of diversified conglomerates that appear to manage their businesses quite well by delegating substantial authority to their business-unit managers. The conglomerates act in some ways like leveraged buyout associations—firms such as Kohlberg, Kravis and Roberts—which own large equity positions in a broad array of unrelated businesses and have been very successful. Conglomerates in some form or another have been an enduring feature of most non-Anglo-American economies for a long time. In fact, Faver, Houston and Naranjo (1998) show that in countries with poorly developed capital markets, the conglomerate discount is not very large. It is fair to say that we do not really know why some conglomerates work well and others do not.

The more general point that this discussion raises is that integration fundamentally changes the resource allocation process by increasing centralized decision making under corporate headquarters. By contrast, the Grossman-Hart-Moore paradigm, as discussed earlier, does not see integration as leading to greater centralization; only as reallocating bargaining power. This discussion also brings out the point that integration can lead to inefficient outcomes from decision-making processes (in this case in the allocation of capital), in contrast to the efficient outcomes from bargaining that always occur in the Grossman-Hart-Moore paradigm. In our view, corporate headquarters, agency problems, and the resource allocation process, must play a key role in any realistic theory of the firm.

**Pulling Themes Together**

We have ranged over a fair amount of ground in this article, including Coase, Berle and Means, Chandler’s study of the development of the multi-divisional corporate structure, capital allocation in diversified conglomerates, and more. It is worth pulling together the main points of our analysis.

1. While there are limitations to the Coasian perspective and the descendent property rights view of the firm developed by Grossman, Hart and Moore, these views emphasize the critical observation that integration reallocates control to different parties and that control is an important element of any business relationship when contracts are incomplete.

2. The simple form of the property rights view of the firm sees integration as a shift in control to the owner/manager of the acquiring firm. However, the Berle and Means perspective emphasizes that managers are not typically owners. Thus,
the effect of integration is to bring bargaining among managers, not owners, under the corporate umbrella.

3. According to Chandler (1962), corporate headquarters evolved as a relatively efficient way of overseeing multi-agent bargaining under a single corporate umbrella. One can fruitfully think of the integrated firm as being comprised of two tiers of agency relationships—at the top between corporate headquarters and investors, and below that, between corporate headquarters and division managers.

4. To understand the effects of integration, one must study the bargaining process among divisional managers and headquarters, all agents of shareholders, with headquarters ultimately determining the outcome of bargaining.

5. As recent work on internal capital markets makes clear, not all decision-making processes lead to efficient resource allocation. One of the keys to understanding the boundaries of firms is understanding the efficiency of decision-making within firms.

- We are grateful to Brad De Long, Robert Gertner, Denis Gromb, Oliver Hart, Bengt Holmström, Alan Krueger, Julio Rotemberg, Jeremy Stein, and Timothy Taylor for many helpful comments.

References


Shleifer, Andrei, and Robert Vishny, “Management Entrenchment: The Case of Manager-


