The Ownership of ENTERPRISE

Henry Hansmann

THE BELKNAP PRESS OF
HARVARD UNIVERSITY PRESS
Cambridge, Massachusetts
London, England
1996

1

An Analytic Framework

A firm's "owners," as the term is conventionally used and as it will be used here, are those persons who share two formal rights: the right to control the firm and the right to appropriate the firm's profits, or residual earnings (that is, the net earnings that remain with the firm after it has made all payments to which it is contractually committed, such as wages, interest payments, and prices for supplies). The reference to "formal" rights in this definition is important. Formal control, for instance, does not necessarily mean effective control. In firms that are incorporated—which comprise most of the institutions of interest to us here, including business corporations, cooperatives, nonprofits, and mutual companies—formal control generally involves only the right to elect the firm's board of directors and to vote directly on a small set of fundamental issues, such as merger or dissolution of the firm. Moreover, in large business corporations the shareholders, who hold formal control, are often too numerous and too dispersed to exercise even these limited voting rights very meaningfully, with the result that corporate managers have substantial autonomy. Hence it has long been common to speak of "the separation of ownership from control," reflecting the substantial autonomy of corporate managers.[1]

Nevertheless, I shall principally be concerned with exploring assignment of the *formal* legal or contractual rights to control and residual earnings. As we shall see in the chapters that follow, there are often strong reasons for giving the formal right of control to a particular class of persons even when those persons are not in a position to

exercise that right very effectively. For this reason, among others, the assignment of these formal rights—which is to say, the assignment of ownership—tends to follow strong and clear patterns.

In theory, the rights to control and to residual earnings could be separated and held by different classes of persons. In practice, however, they are generally held jointly. The obvious reason for this is that, if those with control had no claim on the firm's residual earnings, they would have little incentive to use their control to maximize those earnings, or perhaps even to pay out the earnings received. To be sure, this problem would not arise if all important decisions to be made by those with control could be appropriately constrained in advance by contractual arrangements between them and the holders of the rights to residual earnings. But the essence of what we term "control" is precisely the authority to determine those aspects of firm policy that, because of high transactions costs or imperfect foresight, cannot be specified ex ante in a contract but rather must be left to the discretion of those to whom the authority is granted.[2]

Not all firms have owners. In nonprofit firms, in particular, the persons who have control are barred from receiving residual earnings. As we shall see, however, the same factors that determine the most efficient assignment of ownership also determine when it is appropriate for a firm to have no owners at all.

The Structure of Ownership

In the discussion that follows, it will be helpful to have a term to comprise all persons who transact with a firm either as purchasers of the firm's products or as sellers to the firm of supplies, labor, or other factors of production. I shall refer to such persons—whether they are individuals or other firms—as the firm's "patrons."

Nearly all large firms that have owners are owned by persons who are also patrons. This is obvious in the case of consumer and producer cooperatives, which by definition are firms that are owned, respectively, by their customers and by their suppliers. It is also true of the standard business corporation, which is owned by persons who lend capital to the firm. In fact, the conventional investor-owned firm is nothing more than a special type of producer cooperative—a lenders' cooperative, or capital cooperative. Because we so commonly associate ownership with investment of capital, and because the comparison of

investor-owned firms with cooperatives of other types will be at the core of the analysis that follows, it is worthwhile to elaborate briefly on this point.

Consider, first, the basic structure of a typical producer cooperative. For concreteness, we can take as a simple, stylized example a dairy farmers' cheese cooperative, in which a cheese factory is owned by the farmers who supply the factory with raw milk. (The example is not fanciful; farmer-owned cooperatives account for 45 percent of all natural cheese produced in the United States.)[3] The firm pays its owners—or "members," as they are usually termed in a cooperative—a predetermined price for their milk. This price is set low enough so that the cooperative is almost certain to have positive net earnings from the manufacture and sale of its cheese. Then, at the end of the year, the firm's net earnings are divided pro rata among the members according to the amount of milk they have sold to the cooperative during the year, and distributed as patronage dividends. All voting rights in the firm are also apportioned among its farmer-members, either according to the amount of milk each member sells to the firm or, more simply, on a one-member-one-vote basis. Some or all of the members may have capital invested in the firm. In principle, however, this is unnecessary; the firm might borrow all of the capital it needs. In any case, even where members invest in the firm, those investments generally take the form of debt or preferred stock that carries no voting rights and is limited to a stated maximum rate of dividends. Upon liquidation of the firm, any net asset value—which may derive from retained earnings or from increases in the value of assets held by the firm—is divided pro rata among the members, according to some measure of the relative value of their cumulative patronage.

In short, ownership rights are held by virtue of, and proportional to, one's sale of milk to the firm. Not all farmers who sell milk to the firm need be owners, however; the firm may purchase some portion of its milk from nonmembers, who are simply paid a fixed price (which may be different from the price paid members) and do not participate in net earnings or control.

The structure of a consumer cooperative is similar, except that net earnings and votes are apportioned according to the amounts that members purchase from the firm rather than the amounts they sell to it.

Now imagine a hypothetical "capital cooperative" with a form pre-

cisely analogous to that of the dairy cooperative. The members of the capital cooperative each lend the firm a given sum of money, which the firm uses to purchase the equipment and other assets it needs to operate (say, to manufacture widgets—or cheese). The firm pays the members a fixed interest rate on their loans, set low enough so that there is a reasonable likelihood that the firm will have net earnings after paying this interest and all other expenses. The firm's net earnings are then distributed pro rata among its members according to the amount they have lent, with the distributions taking place currently, as dividends, or upon liquidation. Similarly, voting rights are apportioned among members in proportion to the amount they have lent the firm. To supplement the capital that it obtains from its members, the firm may borrow money from lenders who are not members, but who simply receive a fixed rate of interest (which may be different from the fixed rate paid to members) without sharing in profits or control.

This hypothetical capital cooperative is, transparently, a producers' cooperative just as is the dairy cooperative. Yet this capital cooperative in fact has precisely the structure that underlies the typical business corporation. If this is not immediately obvious, it is perhaps just because, in a business corporation, the fixed interest rate paid on loans from the firm's lender-members—whom we conventionally term "shareholders" or "stockholders"—is typically set at zero for the sake of convenience, thus obscuring the fact that the members' contributions of capital are, in effect, loans.

To be sure, there are also various other ways in which capital cooperatives (that is, business corporations) are often structured a bit differently from other types of cooperatives. For example, in a business corporation the loans from members are usually not arranged annually or for other fixed periods, but rather are perpetual; members can withdraw their capital only upon dissolution of the firm, although an individual member may be free to sell his or her interest in the firm to another person before then. In other types of cooperatives, in contrast, members often remain free to vary the volume of their transactions with the firm over time, and even to terminate their patronage altogether. This distinction is not, however, fundamental. Investor-owned business corporations sometimes permit members to redeem their invested capital at specified intervals or even (as in the standard partnership) at will; open-ended mutual funds are a familiar example. Conversely, cooperatives often require that members make a long-term

commitment to remain patrons. For example, electricity generation and transmission cooperatives commonly insist that their members, which are local electricity distribution cooperatives, enter into thirty-five-year requirements contracts.[4] Agricultural marketing and processing cooperatives, such as the cheese cooperative just described, often require that their members commit themselves to sell to the cooperative a given amount of their production each year for a period of several years.[5] And mutual life insurance companies, which are essentially consumer cooperatives owned by their policyholders, originally issued only nonredeemable policies that committed policyholders to make premium payments—that is, to continue to purchase a specified amount of insurance from the firm—for the rest of their lives.[6]

The allocation of voting rights is another area where business corporations often differ somewhat from other types of cooperatives. In business corporations, the general rule is one-share-one-vote; that is, votes are apportioned according to the amount of capital contributed to the firm. In many cooperatives, in contrast, the rule is one-member-one-vote, with no adjustment for the volume of patronage of the individual members. Again, however, the difference is neither universal nor fundamental. The charters of many eighteenth- and nineteenth-century American business corporations limited the number of votes an individual shareholder could exercise regardless of the number of shares he owned; only in the twentieth century did the practice of one-share-one-vote become nearly universal.[7] And, while the statutes governing cooperatives sometimes still impose a rule of one-member-one-vote, this is not universal and many cooperatives allocate votes proportionally to their members' volume of patronage. (We shall consider later why these different voting rules arose and survived.)

In sum, a business corporation is just a particular type of cooperative: a cooperative is a firm in which ownership is assigned to a group of the firm's patrons, and the persons who lend capital to a firm are just one among various classes of patrons with whom the firm deals.

Conversely, supplying capital to the firm is simply one of many transactional relationships to which ownership can be tied, and there is nothing very special about it. Ownership of a firm need not, and frequently does not, attach to investment of capital. Indeed, contrary to some popular perceptions and even to some more sophisticated organizational theory, ownership of the firm need have nothing to do with ownership of capital, whether physical or financial.[8]

To be sure, it might be argued that ownership is necessarily connected to capital in the sense that the owners of a firm, whether they are suppliers or customers or workers or whatever, are the persons who effectively own the firm's capital, such as its plant and equipment. For example, in our cheese cooperative, one might argue that the farmer-members own the firm's capital in the sense that they collectively have title to, and will profit or lose from fluctuations in the value of, the cheese factory's plant and equipment.

But this is not necessarily true. The firm could rent rather than own the land, buildings, and equipment it uses. It could in fact have title to no physical assets whatever yet still be a large and prosperous firm.[9] It could even have no net financial assets, distributing all profits to members as they are earned and maintaining a line of credit at a bank sufficient to ensure that it can pay bills in periods when expenses temporarily exceed receipts.[10] The members of the cooperative might *choose* to invest some of their personal funds in the firm, or to have the firm retain some of its profits for internal investment. Indeed, as subsequent chapters will discuss at greater length, there are good reasons why the owners of most types of firms, including producer and consumer cooperatives, choose to invest some financial capital in the firm they own. But it is not *necessary* that owners of a firm also be investors in the firm.

Even though the ordinary business corporation is, as I have just argued, essentially a lenders' cooperative, I shall continue to follow the usual convention here and generally use the term "cooperative" to refer only to patron-owned firms *other than* investor-owned firms.

The Structure of Organizational Law

From these observations we can gain a helpful perspective on the general structure of corporation law.

In the United States, basic corporation law is state law rather than federal law. The typical state has three general corporation statutes: a business corporation statute, a cooperative corporation statute, and a nonprofit corporation statute. Most of the organizations we shall be concerned with in this book are formed under one or another of these three types of statutes. (There are, however, a number of exceptions. For example, mutual banks, mutual insurance companies, and housing condominiums are often formed under special corporation statutes specifically designed for them. And the employee-owned firms that are common in the service professions are often formed as partnerships or professional corporations, which are also governed by separate statutes.)

A cooperative corporation statute typically accommodates all types of producer and consumer cooperatives, from retail grocery cooperatives on the one hand to farm processing and marketing cooperatives, such as our cheese factory, on the other. Once it is understood that investor-owned firms are in essence capital cooperatives, it follows that in principle investor-owned corporations could also be formed under cooperative corporation statutes rather than, as is customary, under the separate business corporation statutes.[11] There is no fundamental reason to have business corporation statutes at all; they are just specialized versions of the theoretically more general cooperative corporation statutes. It is appropriate to have separate business corporation statutes simply because it is convenient to have a form that is customized for the most common type of cooperative—the lenders' cooperative—and to signal to patrons more clearly the type of cooperative with which they are dealing.[12] For similar reasons, some agricultural states have separate corporation statutes for another particularly common type of producer cooperative, the agricultural marketing cooperative; some states have special statutes for worker cooperatives; and some states have separate statutes for consumer, as opposed to producer, cooperatives.

The partnership statutes, in contrast, are not as specialized as the corporation statutes. Each state has only one general partnership statute, and under that statute partnership shares can be given in return for any type of patronage—whether it involves the provision of inputs such as labor or capital or the purchase of the firm's products—or to persons who are not patrons at all.

Although cooperatives are sometimes loosely said to be "nonprofit," nonprofit corporations are conceptually quite distinct from cooperatives. The defining characteristic of a nonprofit organization is that the persons who control the organization—including its members, directors, and officers—are forbidden from receiving the organization's net earnings. This does not mean that a nonprofit organization is barred from earning profits; rather, it is the *distribution* of the profits to controlling persons that is forbidden. Thus by definition, a nonprofit organization cannot have owners. A well-drafted nonprofit corporation

statute imposes this "nondistribution constraint" on any organization formed under the statute, and hence prohibits the formation, as a nonprofit corporation, of any form of cooperative and of any other form of owned enterprise.

What Must a Theory of Ownership Explain?

In principle, a firm could be owned by someone who is not a patron. Such a firm's capital needs would be met entirely by borrowing. Its other factors of production would likewise be purchased on the market, and its products would be sold on the market. The owner would then be a pure entrepreneur, of roughly the character described in Frank Knight's classic work,[13] simply controlling the firm and receiving its (positive or negative) residual earnings after all output was sold and inputs were paid for. Such firms are rare, however. Rather, ownership is commonly in the hands of one or another group of the firm's patrons—that is, in the hands of persons who have some other transactional relationship with the firm, either as suppliers or as customers.[14]

It follows that a general theory of enterprise ownership must explain at least two things: First, why is ownership generally given to the firm's patrons? Second, what factors determine the particular group of patrons—whether lenders of capital, suppliers of labor or other inputs, or purchasers of the firm's products or services—to whom ownership is given in any particular firm?

The remainder of this chapter sketches such a theory, and the two following chapters flesh out its details. Parts II–IV then offer illustration and further refinement of the theory through detailed application to particular industries and particular organizational types.

The Firm as a Nexus of Contracts

In developing a theory of ownership, it helps to view the firm—as economists increasingly do these days—as a nexus of contracts.[15] More precisely, a firm is in essence the common signatory of a group of contracts. Some of these contracts are with vendors of supplies or services that the firm uses as inputs, some are employment contracts with individuals who provide labor services to the firm, some are loan agreements with bondholders, banks, and other suppliers of capital, and some are contracts of sale entered into with purchasers of the

firm's products. In small firms organized as sole proprietorships, the individual proprietor signs these contracts. In a corporation or a partnership, the party that signs the contracts is a legal entity. Indeed, one of the most important functions of organizational law is to permit the creation of a juridical person—a single legal entity—that can serve as the signatory to contracts.

A firm's contracts generally commit it to certain actions, such as making payments to vendors or delivering goods or services to customers. But contracts typically also leave the firm with some discretion. An employment contract, for example, generally gives the firm some freedom to choose the particular tasks to which the employee will be assigned; a loan contract commonly gives the firm some choice concerning the uses of the borrowed funds; and a contract of sale often affords the firm some latitude in the methods to be used to produce the goods or services promised to a given customer. The right to exercise this discretion is a vital component of control over the firm, and is by definition the prerogative of the firm's owners. The firm may itself also own assets outright, of course, in which case the exercise of discretion over the use of those assets is included among the control rights belonging to the owners of the firm. Again, however, outright ownership of assets is not an essential aspect of what we call a firm.[16]

Broadly speaking, each transaction that a firm enters into is embedded in one or the other of two relationships between the firm and the patron who is the other party to the transaction. In the first of these relationships, which I shall call "market contracting," the patron deals with the firm only through contract and is not an owner. In the second, which I shall simply call "ownership," the patron is also an owner of the firm.

By terming the first of these two relationships "market contracting" I do not mean to imply that there is necessarily a competitive market for the goods or services in question. The relationship between the firm and its patron may, for example, be one of bilateral monopoly, with only one potential trading partner on each side of the transaction. Rather, I use the expression "market contracting" here simply to emphasize that the patron in question can control the firm's behavior only by seeking enforcement of his contract with the firm, or by threatening to cease transacting with the firm in favor of whatever other alternatives the market offers him. Where the relationship is one of ownership, in contrast, the patron has the additional option of seeking to

control the firm's behavior directly through the firm's mechanisms for internal governance. Moreover, by using the term "market contracting" I do not mean to suggest that the relationships in question are necessarily short-term, as on a spot market; rather, I shall use the term to encompass also long-term, highly interdependent contracting of the type sometimes referred to as "relational" contracting.[17]

Using this terminology, we would then say that, in an investor-owned firm, the transactions between the firm and the patrons who supply the firm with capital occur in the context of ownership, while transactions with workers, other suppliers, and customers all take the form of market contracting. An employee-owned firm, in contrast, obtains labor inputs from workers whose relationship is one of ownership, but obtains its capital and other supplies, and sells its products, through market contracting. And a consumer cooperative, in turn, obtains capital, labor, and all other inputs through market contracting while selling the goods or services it produces in transactions embedded in ownership.

To be sure, patrons occasionally have some but not all of the prerogatives of ownership, putting their relationship with the firm somewhere ambiguously between ownership and market contracting. The relationship between a firm and its employees under German codetermination, which will be examined in Chapters 5 and 6, is a conspicuous example. In general, however, the simple dichotomy between market contracting and ownership that I have described here will be adequate for our purposes.

An Overview of the Theory

If a firm were entirely owned by persons who were not among the firm's patrons, then all the firm's transactions involving inputs and outputs would take the form of market contracting. Although feasible in principle, in practice this is likely to be quite inefficient. Market contracting can be costly, especially in the presence of one or more of those conditions loosely termed "market failure"—for example, where there is an absence of effective competition, or where one of the parties is at a substantial informational disadvantage. We shall examine the costs of market contracting more closely in Chapter 2. For the present we need simply note that, where these costs are high, they can often be reduced by having the purchaser own the seller or vice versa. When

both the purchaser and the seller are under common ownership, the incentive for one party to exploit the other by taking advantage of market imperfections is reduced or eliminated. Assigning ownership of a firm to one or another class of the firm's patrons can thus often reduce the costs of transacting with those patrons—costs that would otherwise be borne by the firm or its patrons. To assign ownership to someone who is not among the firm's patrons would waste the opportunity to use ownership to reduce these costs.

Pursuing this logic we can then ask, for any given firm: what is the lowest-cost assignment of ownership? By "lowest-cost assignment of ownership" I mean the assignment of ownership that minimizes the total costs of transactions between the firm and all of its patrons. (Alternatively, I mean the assignment of ownership that maximizes the total net benefits—benefits minus costs—of transactions between the firm and its patrons. Since a forgone benefit can be considered a cost, these definitions are equivalent.) The analysis just offered suggests that, all other things equal, costs will be minimized if ownership is assigned to the class of patrons for whom the problems of market contracting—that is, the costs of market imperfections—are most severe. For example, if the firm is a natural monopoly vis-à-vis its customers, but obtains its capital, labor, and other factors of production in reasonably competitive markets, then total costs are likely to be minimized by assigning ownership to the firm's customers. This presumably helps explain why, as discussed in Chapter 9, so many rural electric utilities are organized as consumer cooperatives.

If ownership were always perfectly effective, in the sense that it eliminated all costs of market contracting without imposing any new costs of its own, then there would be no more to a theory of ownership than this. In fact, however, ownership itself involves costs. Some of these costs are what might be called "governance" costs; they include the costs of making collective decisions among the owners, the costs of monitoring managers, and the costs of the poor decisions and excessive managerial discretion that result when collective decision making or managerial monitoring are imperfect. Another cost is the risk bearing associated with receipt of residual earnings. We shall explore these and other costs of ownership in detail in Chapter 3. For the moment we need simply note that, like the costs of market contracting, these costs can vary greatly from one class of patrons to another. Some patrons, for example, are in a much better position than others to govern the

firm effectively. Similarly, some are better able than others to bear the risk associated with the right to residual earnings. Consequently, when deciding which class of patrons is to own the firm, the costs of ownership must be considered in addition to the costs of market contracting. For example, Chapter 9 offers evidence that the costs of consumer ownership in an electric utility are significantly higher in urban areas than in rural areas, and that this is an important reason why utility cooperatives are much less common in urban areas than in rural areas.

The least-cost assignment of ownership is therefore that which minimizes the sum of all of the costs of a firm's transactions. That is, it minimizes the sum of (1) the costs of market contracting for those classes of patrons that are not owners and (2) the costs of ownership for the class of patrons who own the firm.

Although this theory is simple in basic concept, it is important when applying the theory to realize that the costs of market contracting for any given class of patrons may depend on which of the other classes of patrons owns the firm.[18] This will become clearer in Chapter 3.

Survivorship

It is reasonable to expect that, over the long run, cost-minimizing forms of organization will come to dominate most industries. Two mechanisms press in this direction. The first is conscious design and imitation on the part of the entrepreneurs who organize firms: a firm's entrepreneurs, together with those persons who expect to be among the firm's patrons, have an incentive to adopt a cost-saving organizational form and share the resulting savings among themselves. The second is market selection: higher-cost forms of organization tend to be driven out of business by their lower-cost competitors. If we observe that a particular form of ownership is dominant in a given industry, this is a strong indication that the form is less costly than other forms of ownership would be in that industry.

In Parts II–IV we shall use this "survivorship test" as important evidence of the relative cost of different forms of ownership. There are, however, a number of reasons why this test might not be an entirely accurate measure of comparative organizational costs. Most obviously, public subsidies or regulation might give a special advantage to one form over another. Moreover, the diffusion of new forms through conscious imitation does not always happen quickly,[19] and for

various reasons market selection can operate quite slowly as well.[20] In interpreting the pattern of ownership that appears in any given industry, we must be attentive to these considerations. In fact, we shall gain important insight into the processes of organizational evolution when we consider the temporal pattern of change in ownership forms in some of the industries examined in later chapters.

What Kinds of Costs?

Some might object that there are other values served by assignment of ownership besides cost minimization and that therefore the cost-minimizing form of ownership might not be the one that is most desirable from a social point of view, or even the one that is chosen by the parties involved. I use the term "cost" here, however, to include all interests and values that might be affected by transactions between a firm and its patrons. For example, among the costs of contracting for labor on the market might be a subjective sense of alienation or disempowerment that could be alleviated if the workers instead owned the firm. In fact, one of the fruits of this inquiry is a better understanding of the range of values, both subjective and objective, that are served by ownership, and of the relative significance of those values to persons who deal with the firm.

Thus I use the expression "cost-minimizing" here to mean "efficient" in the economist's very broad sense of that word—that is, to refer to a situation in which there is no alternative arrangement that could make any class of patrons better off, by their own subjective valuation, without making some other class worse off to a greater degree.[21]

In general, the only persons whose interests are importantly affected by the assignment of ownership in a firm are the firm's patrons. In the long run, moreover, all costs that patrons bear under any particular assignment of ownership—whether those costs are pecuniary or nonpecuniary—should be reflected in the contractual terms under which they will agree to transact with the firm. As a consequence the firms that survive in the market should not be those that simply minimize pecuniary costs, but those that are efficient in the broader sense.

To give the theory sketched here more substance, the next two chapters examine in greater detail the most important costs inherent in market contracting and ownership, respectively.