
Selected Paper No. 56

The Disciplining of Corporate Managers

Eugene F. Fama

**Graduate School
of Business
The University
of Chicago**

Eugene F. Fama is the Theodore O. Yntema Professor of Finance at the University of Chicago Graduate School of Business. His research interests encompass the broad areas of macro-economics, finance, and statistics. Recently, he has been particularly concerned with the determinants of inflation and the behavior of security returns in an inflationary environment. A leader in developing the so-called "efficient markets" hypothesis, his influential writings have stimulated a large volume of related research at Chicago and elsewhere. Professor Fama received the B.A. degree (magna cum laude) from Tufts University in 1960 and the Ph.D. degree from the University of Chicago in 1964.

The Disciplining of Corporate Managers

I. Introduction

Economists have always been troubled by the incentive problems that arise when decision making in a firm is the province of managers who are not the firm's securityholders. Modern interest in these problems began with the treatise of Berle and Means (1932),* but Adam Smith, in 1776, saw fit to state his concern:

The directors of such [joint-stock] companies, however, 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. . . . Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such a company.

Adam Smith (1776), p. 700.

One outcome of academic study on the separation of corporate security ownership and control has been the development of "behavioral" and "managerial" theories of the firm. This approach rejects the classical model of an entrepreneur, or owner-manager, who singlemindedly operates the firm to maximize profits, in favor of theories that focus more on the motivations of a manager who controls but does not own and who has little resemblance to the classical "economic man." The works of Simon (1959), Baumol (1959), Cyert and March (1963), and Williamson (1964) are examples.

More recently the economics literature has moved toward theories that also reject

*References are at the end of the paper. This is a shortened and simplified version of Fama (1980).

the classical model of the firm but assume classical forms of behavior on the part of participants in the firm. The firm is viewed as a set of contracts among factors of production, for example, labor, management, and securityholders, with each factor motivated by and acting rationally in its self-interest. Because of its emphasis on the importance of rights in the organization established by contracts, this approach is characterized under the rubric "property rights." The papers of Alchian and Demsetz (1972) and Jensen and Meckling (1976) are the best examples. The antecedents of their work are in the seminal articles of Coase (1937, 1960).

The property rights literature, and in particular the work of Jensen-Meckling and Alchian-Demsetz, provides a striking insight in viewing the firm as a set of contracts among production factors. In effect, the corporation is viewed as a team whose members act out of self-interest but realize that their destinies depend to some extent on the survival of the team in its competition with other teams. This insight, however, is not carried far enough. In the classical theory, the agent who personifies the firm is the entrepreneur who is taken to be both manager and residual riskbearer. Although his title sometimes changes (for example, Alchian and Demsetz call him "the employer") the entrepreneur continues to play a central role in the firm of the property rights literature. As a consequence, this literature does not really analyze the large modern corporation in which control is in the hands of managers who are more or less separate from the firm's securityholders. In short, the problems typically attributed to separation of security ownership and control are left unresolved.

My main thesis is that the separation of security ownership and control can be ex-

plained as an efficient form of economic organization within the "set of contracts" perspective. I shall first set aside the common presumption that a corporation is owned by its securityholders. The attractive concept of the entrepreneur is also laid to rest, at least for the purposes of the large modern corporation. Instead I shall treat the two functions usually attributed to the entrepreneur—management and riskbearing—as naturally separate ingredients within the set of contracts called a firm. The firm is disciplined by competition from other firms which forces the evolution of devices for efficiently monitoring the performance of the entire team and its individual members. In addition, individual participants in the firm, and in particular its managers, face both the discipline and opportunities provided by the markets for their services, both within and outside of the firm.

II. The Irrelevance of the Concept of Ownership of the Firm

To set a framework for the analysis, let us first describe roles for management and riskbearers in the set of contracts called a firm. Management is a type of labor but with a special role—coordinating the activities of inputs and carrying out the agreed contracts among inputs, all of which can be characterized as "decision making." To explain the role of the riskbearers, suppose for the moment that the firm rents all other factors of production, for example, labor, capital, etc., and that the rental contracts are negotiated at the beginning of each production period with payoffs at the end of the period. The riskbearers then contract to receive or pay the uncertain difference between total revenues and costs at the end of each production period.

When other factors of production are paid at the end of each period, it is not necessary for the riskbearers to invest anything in the firm at the beginning of the period. Most commonly, of course, the riskbearers guarantee performance of their contracts by putting up wealth *ex ante*, and this front money is used to purchase capital and perhaps also the technology that the firm uses in its production activities. In this way the riskbearing function is combined with ownership of capital and technology.

We also commonly observe that the joint functions of riskbearing and ownership of capital are repackaged and sold in different proportions to different groups of investors. For example, when front money is raised by issuing both bonds and common stock, the bonds involve a combination of riskbearing and ownership of capital with a low amount of riskbearing relative to the combination of riskbearing and ownership of capital inherent in the common stock. Unless the bonds are riskfree, the riskbearing function is in part borne by the bondholders. In any case, ownership of capital or any other factor purchased by the firm is shared by bondholders and stockholders.

However, ownership of capital should not be confused with ownership of the firm. Each factor in a firm is owned by somebody. The firm is just the set of contracts covering the way inputs are joined to create outputs and the way receipts from outputs are shared among inputs. In this "nexus of contracts" perspective, ownership of the firm is an irrelevant concept. The point is more than semantic. Dispelling the tenacious notion that a firm is owned by its securityholders is important because it is a first step toward understanding that control over a firm's decisions is not necessarily the province of securityholders. The second step is setting

aside the equally tenacious role in the firm usually attributed to the entrepreneur.

III. Management and Riskbearing: A Closer Look

The entrepreneur, or combined manager-riskbearer, is ubiquitous in models of the firm, recent as well as ancient. However, to understand the large modern corporation, it is better to separate managers and riskbearers and analyze their functions separately. The rationale for doing this is not primarily that the end result is more descriptive of the corporation. The major loss in retaining the concept of the entrepreneur is that one is prevented from developing a perspective on management and riskbearing as separate factors of production, each faced with a market for its services that provides alternative opportunities and, in the case of management, motivation toward performance.

Thus, any given set of contracts, i.e., a particular firm, is in competition with other firms, which are likewise teams of cooperating factors of production. If there is a part of the team that has a special interest in its viability, it is not obviously the riskbearers. It is true that if the team does not prove viable, factors like labor and management are protected by markets in which rights to their future services can be sold or rented to other teams. The riskbearers, as residual claimants, also seem to suffer the most direct consequences from the failings of the team. However, the riskbearers in the modern corporation also have markets for their services, capital markets, which allow them to shift among teams with relatively low transaction costs and to hedge against the failings of any given team by diversifying their holdings across teams.

Indeed, modern portfolio theory tells us

that the optimal portfolio for any investor is likely to be diversified across the securities of many firms. Since he holds the securities of many firms precisely to avoid having his wealth depend too much on any one firm, an individual securityholder generally has no special interest in personally overseeing the detailed activities of any firm. In short, efficient allocation of riskbearing seems to imply a large degree of separation of security ownership from control of a firm.

On the other hand, the managers of a firm rent a substantial lump of wealth, their human capital, to the firm, and the rental rates for their human capital signaled by the managerial labor market are likely to depend on the success or failure of the firm. The function of management is to oversee the contracts among factors and to ensure the viability of the firm. For the purposes of the managerial labor market, the previous associations of a manager with success and failure are information about his talents. The manager of a firm, like the coach of any team, may not suffer any immediate gain or loss in current wages from the current performance of his team, but the success or failure of the team impacts his future wages, and this gives the manager a stake in the success of the team.

The firm's securityholders provide important but indirect assistance to the managerial labor market in its task of valuing the firm's management. A securityholder wants to purchase securities with confidence that the prices paid reflect the risks he is taking, and that the securities will be priced in the future to allow him to reap the rewards (or punishments) of his riskbearing. Thus, although an individual securityholder may not have a strong interest in directly overseeing the management of a particular firm, he has a strong interest in the existence of a

capital market which efficiently prices the firm's securities. The signals provided by an efficient capital market about the values of a firm's securities are likely to be important for the managerial labor market's revaluations of the firm's management.

We come now to the central question. To what extent can the signals provided by the managerial labor market and the capital market, perhaps along with other market-induced mechanisms, discipline managers? I shall first discuss, still in general terms, the types of discipline imposed by managerial labor markets, both within and outside the firm. I shall then analyze conditions under which this discipline is sufficient to resolve potential incentive problems that might be associated with the separation of security ownership and control.

IV. The Viability of Separation of Security Ownership and Control of the Firm: General Comments

The outside managerial labor market exerts many direct pressures on the firm to sort and compensate managers according to performance. One form of pressure comes from the fact that an ongoing firm is always in the market for new managers. Potential new managers are concerned with the mechanics by which their performance will be judged, and they seek information about the responsiveness of the system in rewarding performance. Moreover, given a competitive managerial labor market, when the firm's reward system is not responsive to performance, the firm loses managers, and the best are the first to leave.

There is also much internal monitoring of managers by managers themselves. Part of the talent of a manager is his ability to elicit and measure the productivity of lower managers, so there is a natural process of

monitoring from higher to lower levels of management. Less well appreciated, however, is the monitoring that takes place from bottom to top. Lower managers perceive they can gain by stepping over shirking or less competent managers above them. Moreover, in the team or nexus of contracts view of the firm, each manager is concerned with the performance of managers above and below him since his productivity (and ultimately his wage) is likely to depend on theirs. Finally, although higher managers are affected more than lower managers, all managers realize that the managerial labor market uses the performance of the firm to determine each manager's outside opportunity wage. In short, each manager has a stake in the performance of the managers above and below him, and, as a consequence, undertakes some amount of monitoring in both directions.

All managers below the very top level have an interest in seeing that the top managers choose policies for the firm which provide the most positive signals to the managerial labor market. But by what mechanism can top management be disciplined? Since the body designated for this function is the board of directors, we can ask how it might be constructed to do its job. A board dominated by securityholders does not seem optimal or endowed with good survival properties. Diffuse ownership of securities is beneficial in terms of an optimal allocation of riskbearing, but its consequence is that the firm's securityholders are generally too diversified across the securities of many firms to take much direct interest in a particular firm.

If there is competition among the top managers themselves (all want to be the boss of bosses), then perhaps they are the best ones to control the board of directors. They

are most directly in the line of fire from lower managers when the markets for securities and managerial labor give poor signals about the performance of the firm. Because of their power over the firm's decisions, their market-determined opportunity wages are also likely to be most affected by market signals about the performance of the firm. If they are also in competition for the top places in the firm, they may be the most informed and responsive critics of the firm's performance.

Having gained control of the board, top management may decide that collusion and expropriation of securityholder wealth are better than competition among themselves. The probability of such collusive arrangements might be lowered, and the viability of the board as a market-induced mechanism for low cost internal transfer of control might be enhanced, by the inclusion of outside directors. The latter might best be regarded as professional referees whose task is to stimulate and oversee the competition among the firm's top managers. In a state of advanced evolution of the external markets that buttress the corporate form the outside directors are in their turn disciplined by the market for their services which prices them according to their performance as referees. Since such a system of separation of security ownership from control is consistent with the pressures applied by the managerial labor market, and since it likewise operates in the interests of the firm's securityholders, it probably has good survival properties.

In an important paper, Watts and Zimmerman (1978) provide a similar description of the market-induced evolution of "independent" outside auditors whose function is to certify, and, as a consequence, stimulate the economic viability of the set of contracts called the firm. Like the outside directors,

the outside auditors are policed by the market for their services which prices them in large part on the basis of how well they resist perverting the interests of one set of factors, for example, securityholders, to the benefit of other factors, for example, management. Like the professional outside director, the wealth of the outside auditor depends largely on "reputation."

My analysis does not imply that boards of directors are likely to be composed entirely of managers and outside directors. The board is viewed as a market-induced institution, the ultimate internal monitor of the set of contracts called a firm, whose most important role is to scrutinize the highest decision makers within the firm. In the team or nexus of contracts view of the firm, one can't rule out the evolution of boards of directors that contain many different factors of production (or their hired representatives), whose common trait is that their productivity is affected by that of the top decision makers. On the other hand, one also can't conclude that all such factors will naturally show up on boards since there may be other market-induced institutions, for example, unions, that more efficiently monitor managers on behalf of specific factors. All one can say is that in a competitive environment, lower cost sets of monitoring mechanisms are likely to survive. The role of the board in this framework is to provide a relatively low-cost mechanism for replacing or re-ordering top managers; lower cost, for example, than the mechanism provided by an outside takeover, although, of course, the existence of an outside market for control is another force which helps to sensitize the internal managerial labor market.

It is well to emphasize that in a competitive environment, the need to survive forces the evolution of devices for efficiently

monitoring the performance of the set of contracts called a firm, and of its separate team members. For example, an individual firm may not always choose the optimal mix of insiders and outsiders for its board of directors, but the incentive toward efficiency provided by competition should cause firms in general to do so. Thus, there seems no compelling reason for government intervention in these matters. Indeed, since efficient forms of monitoring can vary across firms and industries and change through time, anything short of prescient intervention is likely to disrupt the natural evolution of such monitoring devices. The need for a government prescription seems no greater here than in the evolution of, say, computer technology, which is likewise of substantial interest to society as a whole as well as to the investors directly involved.

Finally, the perspective I am suggesting owes much to, but is nevertheless different from, existing treatments of the firm in the property rights literature. Thus, Alchian (1969) and Alchian and Demsetz (1972) comment insightfully on the disciplining of management that takes place through the inside and outside markets for managers. However, they attribute the task of disciplining management primarily to the riskbearers, the firm's securityholders, who are assisted to some extent by managerial labor markets and by the possibility of outside takeover. Jensen and Meckling (1976) likewise make control of management the province of the firm's riskbearers, but they do not allow for any assistance from the managerial labor market. Of all the authors in the property rights literature, Manne (1965, 1967) is most concerned with the market for corporate control. He recognizes that with diffuse security ownership, management and riskbearing are naturally sepa-

rate functions. But for him, disciplining management is an "entrepreneurial job" which in the first instance falls on a firm's organizers and later on specialists in the process of outside takeover.

When management and riskbearing are viewed as naturally separate factors of production, looking at the market for riskbearing from the viewpoint of portfolio theory tells us that riskbearers are likely to spread their wealth across many firms and so not be interested in directly controlling the management of any individual firm. Thus, models of the firm, like those of Alchian-Demsetz and Jensen-Meckling, in which the control of management falls primarily on the riskbearers, are not likely to allay the fears of those concerned with the apparent incentive problems created by the separation of security ownership and control. Likewise, Manne's approach, in which the control of management relies primarily on the expensive mechanism of an outside takeover, offers little comfort. The viability of the large corporation with diffuse security ownership is better explained in terms of a model where the primary disciplining of managers comes through managerial labor markets, both within and outside of the firm, with assistance from the panoply of internal and external monitoring devices that evolve to stimulate the ongoing efficiency of the corporate form, and with the market for outside takeovers providing discipline of last resort.

V. The Viability of Separation of Security Ownership and Control: Some Details

The preceding is a general discussion of how pressure from managerial labor markets helps to discipline managers. I now examine somewhat more specifically conditions under which the discipline imposed by managerial

labor markets through the wage revision process can resolve potential incentive problems associated with the separation of security ownership and control of the firm.

To understand better the problem we are trying to solve, let us first examine the situation where the manager is also the firm's sole securityholder, so that there is clearly no incentive problem. When he is sole securityholder, a manager consumes on the job—through shirking, perquisites, or incompetence—to the point where these yield marginal utility equal to that provided by an additional dollar of wealth usable for consumption or investment outside of the firm. The manager is induced to make this specific decision because he pays for consumption on the job directly; that is, as manager he cannot avoid a full ex post settling up with himself as securityholder.

In contrast, when the manager is no longer sole securityholder, and in the absence of some form of full ex post settling up, a manager has an incentive to consume more on the job than is agreed in his contract; that is, the manager perceives that, on an ex post basis, he can beat the game by shirking or consuming more perquisites than previously agreed. This does not necessarily mean that the manager profits at the expense of the other factors. Rational managerial labor markets would understand any shortcomings of available mechanisms for enforcing ex post settling up. Assessments of ex post deviations from contract would be incorporated into contracts on an ex ante basis, for example, through an adjustment of the manager's wage. Nevertheless, a game which is fair on an ex ante basis does not induce the same behavior as a game in which there is also ex post settling up. Herein lie the potential losses from separation of security ownership and control of a firm. There are

situations where, with less than complete ex post settling up, the manager is induced to consume more on the job than he would like, given that on average he pays for his consumption ex ante.

In Fama (1980), I analyze many specific situations where the type of incentive problem described above is resolved through ex post settling up imposed by the managerial labor market through the wage revision process. Here I shall just review the simplest example, which nevertheless captures much of the essential flavor of the arguments.

Pushing a bit on reality for the sake of clarity, suppose we consider a manager's human capital, his stream of future wages, as a marketable asset. Suppose the manager perceives that, because of the consequent revaluations of future wages, the current value of his human capital changes by at least the amount of an unbiased assessment of the wealth changes experienced by other factors, primarily the securityholders, because of his current deviations from previously agreed levels of consumption on the job. Such revaluations of the manager's human capital are a form of full ex post settling up. The manager need not be charged ex ante for presumed ex post deviations from contract since the weight of the wage revision process is sufficient to neutralize his incentives to deviate.

It is important to consider why the manager might perceive that the value of his human capital changes by at least the amount of an unbiased assessment of the wealth changes experienced by other factors due to his deviations from contract. Although his next wage may not adjust by the full amount of an unbiased assessment of the current cost of his deviations from contract, a manager with a multiperiod horizon may perceive that the implied current wealth change, the discounted value of likely changes in the stream

of future wages, is at least as great as the cost of his deviations from contract. In this case, the contemporaneous change in his wealth implied by an eventual adjustment of future wages is a form of full ex post settling up which results in full enforcement of his contract. Moreover, the wage revision process resolves any potential problems about a manager's incentives even though the implied ex post settling up need not involve the firm currently employing the manager; that is, lower or higher future wages due to current deviations from contract may come from other firms.

It is also interesting to consider how the managerial labor market might come to an unbiased assessment of the cost of a manager's deviations from the terms of his contract. If the manager consumes more on the job than previously agreed, he has no incentive to advertise his underperformance. However, managers who fulfill or more than fulfill their contracts have an incentive to advertise their performance, and that of the underperforming managers, to the managerial labor markets both within and outside of the firm. Thus, the performing managers help these markets to revalue nonperforming managers. Moreover, the information provided to markets (financial statements, capital expenditure proposals, etc.) by firms themselves allows comparisons across firms (profits, unit costs, etc.) which are useful for evaluating manager performance. There is also much information about manager performance coming from other markets, for example, the capital markets and the markets for the firm's outputs.

Much of the monitoring of managers that takes place is a public good in the sense that the firm and the manager benefit from the monitoring without paying its direct cost. For example, the signals provided by the outside managerial labor market about a

manager's value in alternative jobs are without direct cost to the manager or the firm. The firm contributes to this outside monitoring process, and pays part of its costs when it enters the managerial labor market to bid for the managers of other firms. An analogous phenomenon exists in financial markets where the holder of securities gets the benefit of continuous repricing as a consequence of the purchase and sale decisions of other investors, and he contributes to the process when he likewise engages in portfolio rebalancing.

There is also much information about a manager's performance which is generated by other managers, both within and outside of the firm, at little direct expense to the manager or the firm. The members of any given team scrutinize the performance of other members of the team in part because they perceive that such monitoring will allow them to step over less talented or more consumption-prone managers above them, and in part because they perceive that their own marginal products are a positive function of the performance of other team members. They also scrutinize the performance of their own team and other teams because they enjoy the game. Given some sort of communication system, the end result is a large body of information about performance available at low cost. The markets for ball players and economists provide ready examples of such information networks. Since the game is not obviously less interesting and the monetary stakes are higher, it would be surprising if such an information network does not also exist in the managerial labor market.

VI. Conclusions

I do not claim that the wage revision process imposed by the managerial labor market always results in a full ex post settling up on

the part of the manager. For example, for an older manager who does not expect to be in the managerial labor market for many future periods, the weight of future wage revisions due to current assessments of performance may amount to substantially less than full ex post settling up. There are certainly other situations where the weight of anticipated future wage changes is insufficient to counterbalance the gains to be had from shirking, or perhaps outright theft, in excess of what was agreed ex ante in a manager's contract.

However, it is just as important to recognize that the weight of anticipations about future wages may amount to more than full ex post settling up. There may be situations where the personal wealth change perceived by the manager as a consequence of deviations from contract is greater than the wealth change experienced by other factors, especially securityholders. Since many of my colleagues have had trouble with this point, it is well to provide an example. Economists (especially young economists) easily imagine situations where the effects of higher or lower quality of a current article or book on the market value of human capital, through enhancement or lowering of "reputation," are in excess of the effects of quality differences on the market value of the specific work to any publisher. Managers can sometimes have similar perceptions with respect to the implications of current performance for the market value of their human capital.

The extent to which the wage revision process imposes ex post settling up in any particular situation is, of course, an empirical issue. But it is probably safe to say that the general phenomenon is at least one of the ingredients in the survival of the large corporation, characterized by diffuse security ownership and the separation of security ownership and control, as a viable form of economic organization.