

Agency Problems and the Theory of the Firm

Eugene F. Fama

University of Chicago

This paper attempts to explain how the separation of security ownership and control, typical of large corporations, can be an efficient form of economic organization. We first set aside the presumption that a corporation has owners in any meaningful sense. The entrepreneur is also laid to rest, at least for the purposes of the large modern corporation. The two functions usually attributed to the entrepreneur—management and risk bearing—are treated as naturally separate factors 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 of its individual members. 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 the firm.

Economists have long been concerned with the incentive problems that arise when decision making in a firm is the province of managers who are not the firm's security holders.¹ One outcome has been the development of "behavioral" and "managerial" theories of the firm which reject the classical model of an entrepreneur, or owner-

This research is supported by the National Science Foundation. Roger Kormendi has contributed much, and the comments of A. Alchian, S. Bhattacharya, G. Becker, F. Black, M. Blume, M. Bradley, D. Breeden, N. Gonedes, B. Horwitz, G. Jarrell, E. H. Kim, J. Long, H. Manne, W. Meckling, M. H. Miller, M. Scholes, C. Smith, G. J. Stigler, R. Watts, T. Whisler, and J. Zimmerman are gratefully acknowledged. Presentations at the finance, labor economics, and industrial organization workshops of the University of Chicago and the workshop of the Managerial Economics Research Center of the University of Rochester have been helpful. The paper is largely an outgrowth of discussions with Michael C. Jensen.

¹ Jensen and Meckling (1976) quote from Adam Smith (1776). The modern literature on the problem dates back at least to Berle and Means (1932).

[*Journal of Political Economy*, 1980, vol. 88, no. 2]

© 1980 by The University of Chicago. 0022-3808/80/8802-0005\$01.65

manager, who single-mindedly 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." Examples of this approach are Baumol (1959), Simon (1959), Cyert and March (1963), and Williamson (1964).

More recently the literature has moved toward theories that reject the classical model of the firm but assume classical forms of economic behavior on the part of agents within the firm. The firm is viewed as a set of contracts among factors of production, with each factor motivated by its self-interest. Because of its emphasis on the importance of rights in the organization established by contracts, this literature is characterized under the rubric "property rights." Alchian and Demsetz (1972) and Jensen and Meckling (1976) are the best examples. The antecedents of their work are in Coase (1937, 1960).

The striking insight of Alchian and Demsetz (1972) and Jensen and Meckling (1976) is in viewing the firm as a set of contracts among factors of production. In effect, the firm is viewed as a team whose members act from 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 risk bearer. 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 fails to explain the large modern corporation in which control of the firm is in the hands of managers who are more or less separate from the firm's security holders.

The main thesis of this paper is that separation of security ownership and control can be explained as an efficient form of economic organization within the "set of contracts" perspective. We first set aside the typical presumption that a corporation has owners in any meaningful sense. The attractive concept of the entrepreneur is also laid to rest, at least for the purposes of the large modern corporation. Instead, the two functions usually attributed to the entrepreneur, management and risk bearing, are treated as naturally separate factors 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 of 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.

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 risk bearing 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 contracts agreed among inputs, all of which can be characterized as “decision making.” To explain the role of the risk bearers, assume for the moment that the firm rents all other factors of production and that rental contracts are negotiated at the beginning of each production period with payoffs at the end of the period. The risk bearers then contract to accept the uncertain and possibly negative 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 risk bearers to invest anything in the firm at the beginning of the period. Most commonly, however, the risk bearers guarantee performance of their contracts by putting up wealth *ex ante*, with this front money used to purchase capital and perhaps also the technology that the firm uses in its production activities. In this way the risk bearing function is combined with ownership of capital and technology. We also commonly observe that the joint functions of risk bearing 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 risk bearing and ownership of capital with a low amount of risk bearing relative to the combination of risk bearing and ownership of capital inherent in the common stock. Unless the bonds are risk free, the risk bearing function is in part borne by the bondholders, and ownership of capital 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. Dispelling the tenacious notion that a firm is owned by its security holders is important because it is a first step toward understanding that control over a firm’s decisions is not necessarily the province of security holders. The second step is setting aside the equally tenacious role in the firm usually attributed to the entrepreneur.

Management and Risk Bearing: A Closer Look

The entrepreneur (manager–risk bearer) is central in both the Jensen-Meckling and Alchian-Demsetz analyses of the firm. For

example, Alchian-Demsetz state: "The essence of the classical firm is identified here as a contractual structure with: 1) joint input production; 2) several input owners; 3) one party who is common to all the contracts of the joint inputs; 4) who has the right to renegotiate any input's contract independently of contracts with other input owners; 5) who holds the residual claim; and 6) who has the right to sell his central contractual residual status. The central agent is called the firm's owner and the employer" (1972, p. 794).

To understand the modern corporation, it is better to separate the manager, the agents of points 3 and 4 of the Alchian-Demsetz definition of the firm, from the risk bearer described in points 5 and 6. The rationale for separating these functions is not just that the end result is more descriptive of the corporation, a point recognized in both the Alchian-Demsetz and Jensen-Meckling papers. The major loss in retaining the concept of the entrepreneur is that one is prevented from developing a perspective on management and risk bearing 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, 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 risk bearers. 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 risk bearers, as residual claimants, also seem to suffer the most direct consequences from the failings of the team. However, the risk bearers 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, portfolio theory tells us that the optimal portfolio for any investor is likely to be diversified across the securities of many firms.² Since he holds the securities of many firms precisely to avoid having his wealth depend too much on any one firm, an individual security holder generally has no special interest in personally overseeing the detailed activities of any firm. In short, efficient allocation of risk bearing 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

² Detailed discussions of portfolio models can be found in Fama and Miller (1972, chaps. 6 and 7), Jensen (1972), and Fama (1976, chaps. 7 and 8).

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 security holders provide important but indirect assistance to the managerial labor market in its task of valuing the firm's management. A security holder 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 risk bearing. Thus, although an individual security holder 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? We first discuss, still in general terms, the types of discipline imposed by managerial labor markets, both within and outside of the firm. We then analyze specific conditions under which this discipline is sufficient to resolve potential incentive problems that might be associated with the separation of security ownership and control.

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 that 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 marginal product is likely to be a positive function of 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 security holders does not seem optimal or endowed with good survival properties. Diffuse ownership of securities is beneficial in terms of an optimal allocation of risk bearing, but its consequence is that the firm's security holders 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 security holder 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 mana-

gers. In a state of advanced evolution of the external markets that buttress the corporate firm, 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 security holders, it probably has good survival properties.³

This 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 cannot 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 marginal products are affected by those of the top decision makers. On the other hand, one also cannot 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 reordering 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.

The perspective suggested here 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 risk bearers, the firm's security holders, 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 man-

³ 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 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 (e.g., security holders) to the benefit of other factors (e.g., management). Like the professional outside director, the welfare of the outside auditor depends largely on "reputation."

agement the province of the firm's risk bearers, 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 risk bearing are naturally separate 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 risk bearing are viewed as naturally separate factors of production, looking at the market for risk bearing from the viewpoint of portfolio theory tells us that risk bearers 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 risk bearers, 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.

The Viability of Separation of Security Ownership and Control: Details

The preceding is a general discussion of how pressure from managerial labor markets helps to discipline managers. We now examine somewhat more specifically conditions under which the discipline imposed by managerial labor markets can resolve potential incentive problems associated with the separation of security ownership and control of the firm.

To focus on the problem we are trying to solve, let us first examine the situation where the manager is also the firm's sole security holder, so that there is clearly no incentive problem. When he is sole security holder, a manager consumes on the job, through shirking, perquisites, or incompetence, to the point where these yield marginal expected utility equal to that provided by an additional dollar of wealth usable for consumption or investment outside of the firm. The man-

ager is induced to make this specific decision because he pays directly for consumption on the job; that is, as manager he cannot avoid a full ex post settling up with himself as security holder.

In contrast, when the manager is no longer sole security holder, and in the absence of some form of full ex post settling up for deviations from contract, a manager has an incentive to consume more on the job than is agreed in his contract. 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 other factors. Rational managerial labor markets understand any shortcomings of available mechanisms for enforcing ex post settling up. Assessments of ex post deviations from contract will 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.

Three general conditions suffice to make the wage revaluation imposed by the managerial labor market a form of full ex post settling up which resolves the managerial incentive problem described above. The first condition is that a manager's talents and his tastes for consumption on the job are not known with certainty, are likely to change through time, and must be imputed by managerial labor markets at least in part from information about the manager's current and past performance. Since it seems to capture the essence of the task of managerial labor markets in a world of uncertainty, this assumption is no real restriction.

The second assumption is that managerial labor markets appropriately use current and past information to revise future wages and understand any enforcement power inherent in the wage revision process. In short, contrary to much of the literature on separation of security ownership and control, we impute efficiency or rationality in information processing to managerial labor markets. In defense of this assumption, we note that the problem faced by managerial labor markets in revaluing the managers of a firm is much entwined with the problem faced by the capital market in revaluing the firm itself. Although we do not understand all the details of the process, available empirical evidence (e.g., Fama 1976, chaps. 5 and 6) suggests that the capital market generally makes rational assessments of the value of

the firm in the face of imprecise and uncertain information. This does not necessarily mean that information processing in managerial labor markets is equally efficient or rational, but it is a warning against strong presumptions to the contrary.

The final and key condition for full control of managerial behavior through wage changes is that the weight of the wage revision process is sufficient to resolve any potential problems with managerial incentives. In this general form,⁴ the condition amounts to assuming the desired result. More substance is provided by specific examples.

Example 1: Marketable Human Capital

Suppose a manager's human capital, his stream of future wages, is 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 security holders, because of his current deviations from contract. Then, as long as the manager is not a risk preferrer, these revaluations of his 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. Note first that the market's assessment of such wealth changes is also its assessment of the difference between the manager's ex post marginal product and the marginal product he contracted to deliver ex ante. However, any assessment of the manager's marginal product is likely to include extraneous noise which has little to do with his talents and efforts. Without specific details on what the market takes to be the statistical process governing the evolution of the manager's talents and his tastes for consumption on the job, one cannot say exactly how far it will go in adjusting his future wages to reflect its most recent measurement of his marginal product. Assuming the market uses information rationally, the adjustment is closer to complete the larger the signal in the most recent measurement relative to the noise, but as long as there is some noise in the process, the adjustment is less than complete.⁴

Although his next wage may not adjust by the full amount of an unbiased assessment of the current cost of his deviations from con-

⁴ Specific illustrations of this point are provided later.

tract, a manager with a multiperiod horizon may perceive that the implied current wealth change, the present 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.

Of course, changes in a manager's wealth as a consequence of anticipated future wage revisions are not always equivalent to full ex post settling up. When a manager does not expect to be in the 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. 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. Since many readers have had trouble with this point, it is well to bring it closer to home.

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.

Example 2: Stochastic Processes for Marginal Products

The next example of ex post settling up through the wage revision process is somewhat more formal than that described above. We make specific assumptions about the stochastic evolution of a manager's measured marginal product and about how the managerial labor market uses information from the process to adjust the manager's future wages—in a manner which amounts to precise, full ex post settling up for the results of past performance.

Suppose the manager's measured marginal product for any period t is composed of two terms: (i) an expected value, given his talents, effort exerted during t , consumption of perquisites, etc.; and (ii)

random noise. The random noise may in part result from measurement error, that is, the sheer difficulty of accurately measuring marginal products when there is team production, but it may also arise in part from the fact that effort exerted and talent do not yield perfectly certain consequences. Moreover, because of the uncertain evolution of the manager's talents and tastes, the expected value of his marginal product is itself a stochastic process. Specifically, we assume that the expected value, \bar{z}_t , follows a random walk with steps that are independent of the random noise, ϵ_t , in the manager's measured marginal product, z_t . Thus, the measured marginal product,

$$z_t = \bar{z}_t + \epsilon_t, \quad (1)$$

is a random walk plus white noise. For simplicity, we also assume that this process describes the manager's marginal product both in his current employment and in the best alternative employment.

The characteristics (parameters) of the evolution of the manager's marginal product depend to some extent on endogenous variables like effort and perquisites consumed, which are not completely observable. Our purpose is to set up the managerial labor market so that the wage revision process resolves any potential incentive problems that may arise from the endogeneity of z_t in a situation where there is separation of security ownership and control of the firm.

Suppose next that risk bearers are all risk neutral and that 1-period market interest rates are always equal to zero. Suppose also that managerial wage contracts are written so that the manager's wage in any period t is the expected value of his marginal product, \bar{z}_t , conditional on past measured values of his marginal product, with the risk bearers accepting the noise ϵ_t , in the ex post measurement of the marginal product. We shall see below that this is an optimal arrangement for our risk-neutral risk bearers. However, it is not necessarily optimal for the manager if he is risk averse. A risk-averse manager may want to sell part of the risk inherent in the uncertain evolution of his expected marginal product to the risk bearers, for example, through a long-term wage contract.

We avoid this issue by assuming that, perhaps because of the more extreme moral hazard problems in long-term contracts (remember that \bar{z}_t is in part under the control of the manager) and the contracting costs to which these moral hazard problems give rise, simple contracts in which the manager's wage is reset at the beginning of each period are dominant, at least for some nontrivial subset of firms and managers.⁵ If we could also assume away any remaining moral hazard

⁵ Institutions like corporations, that are subject to rapid technological change with a large degree of uncertainty about future managerial needs, may find that long-term

(managerial incentive) problems, then with risk-averse managers, risk-neutral risk bearers, and the presumed fixed recontracting period, the contract which specifies ex ante that the manager will be paid the current expected value of his marginal product dominates any contract where the manager also shares the ex post deviation of his measured marginal product from its ex ante expected value (see, e.g., Spence and Zeckhauser 1971).

However, contracts which specify ex ante that the manager will be paid the current expected value of his marginal product seem to leave the typical moral hazard problem that arises when there is less than complete ex post enforcement of contracts. The noise ϵ_t in the manager's marginal product is borne by the risk bearers. Once the manager's expected marginal product \bar{z}_t (= his current wage) has been assessed, he seems to have an incentive to consume more perquisites and provide less effort than are implied in \bar{z}_t .

A mechanism for ex post enforcement is, however, built into the model. With the expected value of the manager's marginal product wandering randomly through time, future assessments of expected marginal products (and thus of wages) will be determined in part by ϵ_t , the deviation of the current measured marginal product from its ex ante expected value. In the present scenario, where \bar{z}_t is assumed to follow a random walk, Muth (1960) has shown that the expected value of the marginal product evolves according to

$$\bar{z}_t = \bar{z}_{t-1} + (1 - \phi)\epsilon_{t-1}, \quad (2)$$

where the parameter ϕ ($0 < \phi < 1$) is closer to zero the smaller the variance of the noise term in the marginal product equation (1) relative to the variance of the steps in the random walk followed by the expected marginal product.

In fact, the process by which future expected marginal products are adjusted on the basis of past deviations of marginal products from their expected values leads to a precise form of full ex post settling up. This is best seen by writing the marginal product z_t in its inverted form, that is, in terms of past marginal products and the current noise. The inverted form for our model, a random walk embedded in random noise, is

$$z_t = (1 - \phi)z_{t-1} + \phi(1 - \phi)z_{t-2} + \phi^2(1 - \phi)z_{t-3} + \dots + \epsilon_t, \quad (3)$$

managerial contracts can only be negotiated at high cost. On the other hand, institutions like governments, schools, and universities may be able to forecast more reliably their future needs for managers (and other professionals) and so may be able to offer long-term contracts at relatively low cost. These institutions can then be expected to attract the relatively risk-averse members of the professional labor force, while the riskier employment offered by corporations attracts those who are willing to accept shorter-term contracts.

so that

$$\bar{z}_t = (1 - \phi)z_{t-1} + \phi(1 - \phi)z_{t-2} + \phi^2(1 - \phi)z_{t-3} + \dots \quad (4)$$

(see, e.g., Nelson 1973, chap. 4, or Muth 1960).

For our purposes, the interesting fact is that, although he is paid his ex ante expected marginal product, the manager does not get to avoid his ex post marginal product. For example, we can infer from (4) that z_{t-1} has weight $1 - \phi$ in \bar{z}_t ; then it has weight $\phi(1 - \phi)$ in \bar{z}_{t+1} , $\phi^2(1 - \phi)$ in \bar{z}_{t+2} , and so on. In the end, the sum of the contributions of z_{t-1} to future expected marginal products, and thus to future wages, is exactly z_{t-1} . With zero interest rates, this means that the risk bearers simply allow the manager to smooth his marginal product across future periods at the going opportunity cost of all such temporal wealth transfers. As a consequence, the manager has no incentive to try to bury shirking or consumption of perquisites in his ex post measured marginal product.

Since the managerial labor market is presumed to understand the weight of the wage revision process, which in this case amounts to precise full ex post settling up, any potential managerial incentive problems in the separation of risk bearing, or security ownership, from control are resolved. The manager can contract for and take an optimal amount of consumption on the job. The wage set ex ante need not include any allowance for ex post incentives to deviate from the contract since the wage revision process neutralizes any such incentives. Note, moreover, that the value of ϕ in the wage revision process described by (4) determines how the observed marginal product of any given period is subdivided and spread across future periods, but whatever the value of ϕ , the given marginal product is fully accounted for in the stream of future wages. Thus, it is now clear what was meant by the earlier claim that although the parameter ϕ in the process generating the manager's marginal product is to some extent under his control, this is not a matter of particular concern to the managerial labor market.

A somewhat evident qualification is in order. The smoothing process described by (4) contains an infinite number of terms, whereas any manager has a finite working life. For practical purposes, full ex post settling up is achieved as long as the manager's current marginal product is "very nearly" fully absorbed by the stream of wages over his future working life. This requires a value of ϕ in (4) which is sufficiently far from 1.0, given the number of periods remaining in the manager's working life. Recall that ϕ is closer to 1.0 the larger the variance of the noise in the manager's measured marginal product relative to the variance of the steps of the random walk taken by the expected value of his marginal product. Intuitively, when the variance

of the noise term is large relative to that of the changes in the expected value, the current measured marginal product has a weak signal about any change in the expected value of the marginal product, and the current marginal product is only allocated slowly to expected future marginal products.

Some Extensions

Having qualified the analysis, let us now indicate some ways in which it is robust to changes in details of the model.

1. More Complicated Models for the Manager's Marginal Product

The critical ingredient in enforcing precise full ex post settling up through wage revisions on the basis of reassessments of expected marginal products is that when the marginal product and its expected value are expressed in inverted form, as in (3) and (4), the sum of the weights on past marginal products is exactly 1.0. This will be the case (see, e.g., Nelson 1973, chap. 4) whenever the manager's marginal product conforms to a nonstationary stochastic process, but the changes from period to period in the marginal product conform to some stationary ARMA (mixed autoregressive moving average) process. The example summarized in equations (1)–(4) is the interesting but special case where the expected marginal product follows a random walk so that the differences of the marginal product are a stationary, first-order moving average process. The general case allows the expected value of the marginal product to follow any more complicated nonstationary process which has the property that the differences of the marginal product are stationary, so that the marginal product and its expected value can be expressed in inverted form as

$$z_t = \pi_1 z_{t-1} + \pi_2 z_{t-2} + \dots + \epsilon_t \quad (5)$$

$$\bar{z}_t = \pi_1 \bar{z}_{t-1} + \pi_2 \bar{z}_{t-2} + \dots \quad (6)$$

with

$$\sum_{i=1}^{\infty} \pi_i = 1. \quad (7)$$

These can be viewed as the general conditions for enforcing precise full ex post settling through the wage revision process when the

manager's wage is equal to the current expected value of his marginal product.⁶

2. Risk-Averse Risk Bearers

In the framework summarized in equations (5)–(7), if the manager switches firms, the risk bearers of his former firm are left with the remains of his measured marginal products not previously absorbed into the expected value of his marginal product. Nevertheless, in the way we have set up the world, the risk bearers realize that the manager's next firm continues to set his wage according to the same stochastic process as the last firm. Since this results in full ex post settling up on the part of the manager, the motive for switching firms cannot be to avoid perverse adjustments of future wages on the basis of past performance. On average, the switching of managers among firms does not result in gains or losses to risk bearers, which means that the switches are a matter of indifference to our presumed risk-neutral risk bearers.

It is, however, interesting to examine how the analysis might change when the risk bearers are risk averse and switching of managers among firms is not a matter of indifference. Suppose, for the moment, that the risk bearers offer managers contracts where, as before, the manager's wage tracks the expected value of his marginal product, but each period there is also a fixed discount in the wage to compensate the risk bearers for the risks of unfinished ex post settling up with the firm as a consequence of a possible future shift by the manager to another firm. Such an arrangement may satisfy the risk bearers, but it will not be acceptable to the manager. As long as his marginal product evolves according to equations (5)–(7), both in his current firm and in the best alternative, the manager is subject to full ex post settling up. Thus, any risk adjustment of his wage to reflect the fact that the settling up may not be with his current firm is an uncompensated loss which he will endeavor to avoid.

The manager can avoid any risk discount in his wage, and maintain complete freedom to switch among firms, by himself bearing all the risk of his marginal product; that is, he contracts to accept, at the end of each period, his ex post measured marginal product rather than its ex ante expected value so that there is, period by period, full ex post settling up with his current firm. There is such a presumption against

⁶ When \bar{z}_t follows a stationary process, the long-run average value toward which the process always tends will eventually be known with near perfect certainty. Thus, the case of a stationary expected marginal product is of little interest, at least for the purposes of ex post settling up enforced by the wage revision process.

the optimality of immediate, full ex post settling up in the literature on optimal contracting that it behooves us to examine how and why it works, and is optimal, in the circumstances under examination.

Contractual Settling Up

The literature on optimal contracting, for example, Harris and Raviv (1978, 1979), Holmström (1979), and Shavell (1979), suggests uniformly that when there is noise in the manager's marginal product, that is, when the deviation of measured marginal product from its expected value cannot be traced unambiguously and costlessly to the manager's actions (talents, effort exerted, and consumption on the job), then a risk-averse manager will always choose to share part of the uncertainty in the evaluation of his performance with the firm's risk bearers. He will agree to some amount of ex post settling up, but always less than 100 percent of the deviation of his measured marginal product from its ex ante expected value. In short, the contracting models suggest that we must learn to live with the incentive problems that arise when there is less than complete ex post enforcement of contracts.

The contracting literature is almost uniformly concerned with 1-period models. In a 1-period world, there can be no enforcement of contracts through a wage revision process imposed by the managerial labor market. The existence of this form of ex post settling up in a multiperiod world affects the manager's willingness to engage in explicit contractual ex post settling up.

For example, in the model summarized in equations (5)–(7), the manager's wage in any period is the expected value of his marginal product assessed at the beginning of the period, and the manager does not immediately share any of the deviation of his ex post marginal product from its ex ante expected value. However, because it contains information about future expected values of his marginal product, eventually the manager's current measured marginal product is allocated in full to future expected marginal products. Equivalently, in the wage revision process described by equations (5)–(7), the managerial labor market in effect acts as a financial intermediary. It withdraws portions of past accumulated measured marginal products to pay the manager a dividend on his human capital equal to the expected value of his marginal product, and implicitly provides the lending arrangements which allow the manager to spread his current measured marginal product over future periods in precisely the way the current marginal product will contribute to expected future marginal products.

Looked at from this perspective, however, the manager might simply contract to take the ex post measured value of his marginal product as his wage and then himself use the capital market to smooth his measured marginal product over future periods. Since the same asset (his human capital) is involved, the manager should be able to carry out these smoothing transactions via the capital market on the same terms as can be had in the managerial labor market. The advantage to the manager in smoothing through the capital market, however, is that he can then contract to accept full ex post settling up period by period (he is paid his measured marginal product), which means he can avoid any risk discount in his wage that might be imposed when he is paid the expected value of his marginal product with the possibility of unanticipated switches to other firms.⁷

It is important to recognize that using the capital market in the manner described above allows the manager to "average out" the random noise in his measured marginal product. Thus, when he is instead paid the expected value of his marginal product each period, and when the process generating his marginal product is described by equations (5)–(7), the manager's current measured marginal product is eventually allocated in full to future expected marginal products. This happily, but only coincidentally, resolves incentive problems by imposing full ex post settling up. The allocation of the current marginal product to future expected marginal products in fact occurs because the current marginal product has information about future expected marginal products. The weights π_i in equations (5)–(7) are precisely those that optimally extract this information and so optimally smooth or average out the purely random noise in the manager's measured marginal product. The manager can achieve the same result by contracting to be paid the measured value of his marginal product and then using the capital market to smooth his marginal product. This power of the capital market to reduce the terror in full contractual ex post settling up is lost in the 1-period models that dominate the contracting literature.

⁷ With positive interest rates, contracting to be paid his measured marginal product and then using the capital market to smooth the marginal product over future periods dominates the contract in which the manager is paid the expected value of his marginal product. Equivalence can be restored by adjusting the expected marginal product \bar{z}_i in eq. (6) for accumulated interest on the past marginal products, z_{i-1}, z_{i-2}, \dots , or by prepaying the present value of interest on the deferrals of the current marginal product over future periods. Suffice it to say, however, that either accumulation or prepayment of interest complicates the problems posed by possible shifts of the manager to other firms and so may lean the system toward contracts in which the manager is paid his measured marginal product and then uses the capital market to achieve optimal smoothing.

Conclusions

The model summarized by equations (5)–(7) is one specific scenario in which the wage revision process imposed by the managerial labor market amounts to full ex post settling up by the manager for his past performance. The important general point is that in any scenario where the weight of the wage revision process is at least equivalent to full ex post settling up, managerial incentive problems—the problems usually attributed to the separation of security ownership and control of the firm—are resolved.

No claim is made that the wage revision process always results in a full ex post settling up on the part of the manager. There are certainly situations where the weight of anticipated future wage changes is insufficient to counterbalance the gains to be had from ex post shirking, or perhaps outright theft, in excess of what was agreed ex ante in a manager's contract. On the other hand, precise full ex post settling up is not an upper bound on the force of the wage revision process. There are certainly situations where, as a consequence of anticipated future wage changes, a manager perceives that the value of his human capital changes by more than the wealth changes imposed on other factors, and especially the firm's security holders, by his current deviations from the terms of his contract.

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 modern large corporation, characterized by diffuse security ownership and the separation of security ownership and control, as a viable form of economic organization.

References

- Alchian, Armen A. "Corporate Management and Property Rights." In *Economic Policy and the Regulation of Corporate Securities*, edited by Henry G. Manne. Washington: American Enterprise Inst. Public Policy Res., 1969.
- Alchian, Armen A., and Demsetz, Harold. "Production, Information Costs, and Economic Organization." *A.E.R.* 62 (December 1972): 777–95.
- Baumol, William J. *Business Behavior, Value and Growth*. New York: Macmillan, 1959.
- Berle, Adolph A., Jr., and Means, Gardiner C. *The Modern Corporation and Private Property*. New York: Macmillan, 1932.
- Coase, Ronald H. "The Nature of the Firm." *Economica*, n.s. 4 (November 1937): 386–405.
- . "The Problem of Social Cost." *J. Law and Econ.* 3 (October 1960): 1–44.
- Cyert, Richard M., and March, James G. *A Behavioral Theory of the Firm*. Englewood Cliffs, N.J.: Prentice-Hall, 1963.

- Fama, Eugene F. *Foundations of Finance*. New York: Basic, 1976.
- Fama, Eugene F., and Miller, Merton H. *The Theory of Finance*. New York: Holt, Rinehart & Winston, 1972.
- Harris, Milton, and Raviv, Artur. "Some Results on Incentive Contracts with Applications to Education and Employment, Health Insurance, and Law Enforcement." *A.E.R.* 68 (March 1978): 20-30.
- . "Optimal Incentive Contracts with Imperfect Information." Working Paper no. 70-75-76, Carnegie-Mellon Univ., Graduate School of Indus. Admin., April 1976 (rev. January 1979), forthcoming in *J. Econ. Theory*.
- Holmström, Bengt. "Moral Hazard and Observability." *Bell J. Econ.* 10 (Spring 1979): 74-91.
- Jensen, Michael C. "Capital Markets: Theory and Evidence." *Bell J. Econ. and Management Sci.* 3 (Autumn 1972): 357-98.
- Jensen, Michael C., and Meckling, William H. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *J. Financial Econ.* 3 (October 1976): 305-60.
- Manne, Henry G. "Mergers and the Market for Corporate Control." *J.P.E.* 73, no. 2 (April 1965): 110-20.
- . "Our Two Corporate Systems: Law and Economics." *Virginia Law Rev.* 53 (March 1967): 259-85.
- Muth, John F. "Optimal Properties of Exponentially Weighted Forecasts." *J. American Statis. Assoc.* 55 (June 1960): 299-306.
- Nelson, Charles R. *Applied Time Series Analysis for Managerial Forecasting*. San Francisco: Holden-Day, 1973.
- Shavell, Steven. "Risk Sharing and Incentives in the Principal and Agent Relationship." *Bell J. Econ.* 10 (Spring 1979): 55-73.
- Simon, Herbert A. "Theories of Decision Making in Economics and Behavioral Science." *A.E.R.* 49 (June 1959): 253-83.
- Smith, Adam. *The Wealth of Nations*. 1776. Cannan ed. New York: Modern Library, 1937.
- Spence, Michael, and Zeckhauser, Richard. "Insurance, Information and Individual Action." *A.E.R.* 61 (May 1971): 380-87.
- Watts, Ross L., and Zimmerman, Jerold. "Auditors and the Determination of Accounting Standards, an Analysis of the Lack of Independence." Working Paper GPB 7806, Univ. Rochester, Graduate School of Management, 1978.
- Williamson, Oliver E. *The Economics of Discretionary Behavior: Managerial Objectives in a Theory of the Firm*. Englewood Cliffs, N.J.: Prentice-Hall, 1964.