The optimal structure of incentives and authority within an organization

James A. Mirrlees

Professor of Economics
Nuffield College, Oxford

Two kinds of models for a productive organization are presented. In the first, both production and rewards are based on the performance of individuals, which is perfectly observed. Their abilities are not observable. Despite this, theorems are proved giving strong grounds for the equality of wages and marginal products unless there is monopsony in the labor market. This latter case is also discussed. The second model, which focuses on the imperfect observation of performance, allows interesting deductions about optimal payment schedules and organizational structure.

1. Introduction

■ The usual idea of an organization is that it is a group of people (or roles) within which a structure of authority is defined. In other words, the actions of each member of the organization are constrained by certain of the decisions made by other members. Simon (1957) has developed this view, and it has been taken up more recently by Arrow (1974). Simon and Arrow, in common with many political theorists, emphasize the existence and desirability of limits on authority, which are termed "responsibility." More generally, it is clear that relations of authority can, and usually do, operate in both directions between any two members or subgroups within an organization. But, even allowing for this, the possibilities of organizational structure are perhaps rather richer than those suggested by the term "authority," as Marschak and Radner (1972, p. 313) have indicated. In their work on the "theory of teams," they choose to concentrate on a different, though related, aspect of organizations—the diversity of information available to the members. At the same time, they narrow their atten-

---

James A. Mirrlees received the M.A. from the University of Edinburgh (1957), the B.A. from the University of Cambridge (1959), and the Ph.D. from the University of Cambridge (1963). His research centers on systems of incentive and control under uncertainty.

I am most grateful to Oliver Williamson for suggesting the subject and inviting me to present the paper at the Symposium on The Economics of Internal Organization, which was held at the University of Pennsylvania, September 19–21, 1974. The Symposium was supported by grants from the National Science Foundation (NSF-GS-35889X) and the General Electric Foundation. I am grateful to the participants in that conference for their comments on the first draft of this paper, and particularly to Michael Spence, whose perception of a serious error, and constructive suggestions, stimulated a complete revision of Section 2 of the paper.

tion, and propose a correspondingly narrow definition of "team," to organizations whose members have common preferences. Some authors have talked more loosely of teams as groups of people who together can achieve more than if they act separately [Mirrlees (1972a), Alchian and Demsetz (1972)]. Where such a coalition is possible, one can consider the possibility of the group's agreeing to work together without setting up a system of authority. Alchian and Demsetz propose a model of the firm in which the function of the management hierarchy is simply to measure the labor input of members of the group, payments being in accordance with contractual agreement. They even seem to claim that this is the only model appropriate to firms, and that authoritative relationships do not occur.

This last claim cannot be accepted, but it may be agreed that there are interesting possibilities of organization which one would find difficult to describe in terms of authority, and that these possibilities can be applied to the same production possibilities as can more authority-based modes of organization. The common feature is personal relationship between members of the group, with an established pattern of interaction. Thus conceived, organizations may include sharecropping tenancy of a man's land, or even bank lending; for in these cases, the contract governing the use of the lender's asset is based on personal information, and may specify aspects of individual behavior [cf. Cheung (1968), Stiglitz (1974)]. Indeed, there is no sharp line to be drawn between perfect-market relationships and intraorganization relationships: arrangements may vary in a continuous manner from pure trading at an exogenously established price to near perfect obedience by one party to the command of the other. One wants to consider the whole of this spectrum, so as to explain actual organizational structures, and prescribe better ones.

It is hard to specify the spectrum of possible economic relationships in a way that begins to be adequate. In this paper, I make the task easier by completely ignoring all bargaining. It might be thought that this ignores too much; but there are interesting suggestions in the literature, including some already mentioned, which avoid game theory. For example, Williamson (1970) advances a model of organizations, based on imperfect communication, from which is derived an optimum size of firm even when the technology exhibits constant or increasing returns to scale. The models to be developed in the present paper provide a similar theory, based on a more detailed theory of the relationships between the parties involved; but with rather different consequences.

In these models, the members of the organization have different interests, and behave in accordance with their own interests. We are therefore not dealing with a team in the Marschak-Radner sense. Members of the organization take independent, but related, decisions. Thus, though related to Wilson's theory of syndicates (1968) (where diverse individuals share in the consequences of a single decision), our theory is essentially different. The models are closely related to the theory of land tenure systems (mentioned above), the theory of agency [Ross (1973)], and models of behavior subject to moral hazard [Pauly (1968), Zeckhauser (1970), Spence and Zeckhauser (1971)]. This being so, it is as well to repeat the point that Arrow has made about the theory of moral hazard (1971, p. 220), that in situations where moral hazard arises, there is, potentially, general advantage in moral behavior, i.e., behavior not motivated by narrow self-interest; and that such behavior occurs. The models used in this paper assume that contracts, explicit and implicit, are exploited by the parties in their own interest, so that promises and claims about unobservable behavior, for instance, are not admissable. A theory that overemphasizes self-interested behavior in this way deserves to fail in predicting various features of actual organizations; but it would be surprising if it were wholly irrelevant.

The aim of the models in the paper is to explain the distribution of incomes within the firm, to explain the existence of authoritative relationships, and to derive a hierarchical structure and investigate its effects on production. Naturally much of this program is imperfectly worked out, and important elements are missing. Throughout, uncertainty plays an essential role. There may be uncertainty about the tasks (or the value of the tasks) that any member of the organization will be undertaking, about his capability to undertake these tasks, and about the way in which he will or has undertaken these tasks. Uncertainties of these kinds suggest an analogy with the theory of public finance, where the government is assumed to have limited information about the characteristics of the population it rules. Section 2 develops the theory of a profit-maximizing organization, which has imperfect information about the qualities of its workers when they are recruited, and uses the model to elucidate the relevance of incentive considerations to payment schedules. The model assumes, however, that workers are perfectly able to predict their own activity and rewards within the organization. It is therefore impossible to discuss many interesting questions, particularly the structure of authority.

These issues are taken up in Sections 3–5, where the models allow some scope for imperfect reward administration and monitoring. Section 3 deals with the organization that, from the point of view of this paper, seems to be the simplest: an organization consisting of two members, one of whom is subordinate (in a sense to be made precise) to the other. This example is helpful in allowing one to model and explore the various kinds of uncertainty mentioned above. The model is then extended to become a two-level organization in Section 4, and a multilevel organization in Section 5. In Section 6—where we return to a less technical level after the mathematical complications of the previous three sections—features of the models are briefly related to the concepts of control-loss and control-span and their consequences. Section 7 summarizes the argument.

2. The pay structure for unknown abilities

■ It is commonly asserted that one reason for the hierarchical pay structure within organizations is to be found in the incentives they provide. The person whose work is found good is promoted, and the managers and owners of the organization are thus enabled to discover which members of the organization do good work. None of this makes sense for the firm in perfect competition. But that is because perfect competition assumes that the properties, including the abilities, of each potential worker are public knowledge. For incentives to have a role, it is necessary that the management have only imperfect information about the abilities and willingness of men and women to work: then they are interested in the way that a structure of wages and salaries elicits much valuable activity from the able and the energetic, while allowing others to choose a more moderate work level and position. It may suit to pay more than the market requires. At any rate, it appears that consideration of the incentives created by the pay structure should be an important element in explaining it.

To explore these matters, a reasonably simple and extreme model is suggested. In it, workers are completely aware of their own abilities, and choose how hard to work on that basis. Employers, on the other hand, cannot distinguish among job applicants, and therefore set pay schedules which are applied to all comers. There are many firms in the economy, all wanting to maximize profits. The behavior of workers, which they determine themselves in the light of the pay structure, is assumed to determine the firm's output. Thus there is no explicit role for authoritative relations, and the model looks much like the orthodox monopsony model with the firm setting prices and the workers choosing labor supply. The case of a perfectly elastic supply of workers to the firm will, however, be discussed, as well as the monopsonistic case. The question to be chiefly considered is the relation, in equilibrium, between the wages which workers of different skills receive and their marginal productivities.

□ **Technologies.** Production possibilities will be assumed to take a form which considerably generalizes the special model which has been used in recent years to analyze the corresponding problem in public finance, that of optimal income taxation. The generalization does not greatly affect the details of the analysis, except at one important point; but it is important to establish that the analysis applies to a model wherein a hierarchical structure of employment is natural, since that is the object of study in the present paper.

Each firm in this model produces a single kind of output, in amount y, with fixed factors (which are ignored notationally) and different kinds of labor. The work done by a worker is denoted by z: different workers choose to provide different z. z is a measure of the quantity and quality of work, one-dimensional for simplicity. Output is taken to be a function of the *distribution* of z. This is best illustrated by some examples, where z is distributed as a continuous variable with density function f:

$$y = H(\int zf(z)dz), \tag{1}$$

$$y = h \exp (\int \alpha(z) \log f(z)dz), \tag{2}$$

$$y = H(\int \alpha(z)H^{-1}(f(z))dz), \tag{3}$$

or

$$y = G(\int_{z\epsilon A}\beta(z)f(z)dz, \int_{z\epsilon A}\gamma(z)f(z)dz). \tag{4}$$

(1) is the function that has usually been used, for simplicity, in income-tax theory, where the contributions of different workers are perfectly substitutible for one another. (2) generalizes the Cobb-Douglas production function to a continuum of inputs. (3) is a more general class, including (1), (2), and the generalized CES function as special cases. (4) captures the notion that workers may be used in either of two different kinds of activity (e.g., manual or supervisory), the levels of each contributing to total output through the function G. This last form is particularly worthy of attention because it can

capture the idea that some workers are promoted after a time when the employer has acquired sufficient information about their working performance; and promotion is to a different kind of activity. This kind of arrangement is frequently mentioned as a reason for paying supervisory workers more: it can provide an incentive for those in less senior occupations to work hard. One of our aims is to check this argument. Perhaps a more precise statement of the situation would be

$$y = G(b + \lambda c, (1 - \lambda)c),\ b = \int_{z \in A} \beta f dz,\ c = \int_{z \notin A} \gamma f dz, \quad \text{(4a)}$$

where λ is the (discounted) proportion of career spent in the "lower" occupation.

One wants to extend these definitions to general distributions of z, where, say, a nonzero proportion of the labor force all supply the same value of z. The extension is done by continuity. For example, in case (2), a group of workers concentrated at a single value of z has no effect on output.

The assumption that output depends only on work levels, z, and not on ability, n, distinguishes this model from that of Spence (1973), and others, where a worker's productivity also depends on n. This dependence is crucial for Spence's results. Work by Rothschild and Stiglitz on Spence's model shows that the n-independent productivity assumption is important for the results to be proved below.

In each of the examples above (with suitable differentiability of the unspecified functions) output is a differentiable function of labor inputs. This means, in the present context, that when the distribution f depends differentiably on a parameter ϵ, there exists a function p such that, if y is the output resulting from f

$$\frac{d}{d\epsilon} y = \int p(z) \frac{\partial}{\partial \epsilon} f(z,\epsilon) dz. \quad (5)$$

$p(z)$ is, in a natural sense, the marginal product of a worker providing work z. Note that in general (i.e., apart from the simple case (1)) p depends on $f(\cdot)$, the distribution of z within the labor force. It will be assumed in what follows that p exists.

□ **Workers.** Members of the labor force are characterized by a single parameter n. When faced with a payment schedule $w(z)$, a worker of type n chooses $z(n)$ so as to maximize a function, strictly concave in w and z,

$$u(w(z),z,n). \quad (6)$$

$w(z)$ is to be thought of as the present value of earnings in the firm, possibly through a series of promotions, if the worker behaves in the way described by z. It is assumed that $u_w > 0$, $u_z < 0$, $u_n > 0$. Though other interpretations are possible, n will here be interpreted as "skill." The maximization of (6) implies, if everything is differentiable, and z nonzero, that

$$u_w w'(z) + u_z = 0. \quad (7)$$

This can be written in the form

$$w'(z) = s(w,z,n), \quad (8)$$

where $s = -u_z/u_w$ is the marginal rate of substitution. The interpretation of n as skill suggests that we assume $s_n < 0$. The distribution of skills, n, within the labor force of the firm is described by a density function $\phi(n)$.

An alteration of the pay schedule w changes the function $z(n)$. We must see how this change affects output, the labor force being given. Let F be the distribution function of z. Then $F(z(n))$ is unchanged, so that variations in F and z are related by

$$\delta F(z) + F'(z)\delta z(n) = 0,$$

i.e.

$$\delta F = - f(z)\delta z(n). \tag{9}$$

From (5), $\delta y = \int p(z)\delta F'(z)dz = -\int p'(z)\delta F(z)dz$, integrating by parts. Therefore, using (9),

$$\begin{aligned}\delta y &= \int p'(z)\delta z(n)f(z)dz \\ &= \int p'(z(n))\delta z(n)\phi(n)dn.\end{aligned} \tag{10}$$

The specification of the model is now complete, except for the conditions of supply of workers to the firm. The whole spectrum of possibilities will be considered, from the firm with a given labor force to the firm facing a perfectly elastic supply of workers. The case of perfectly elastic supply is that where workers of type n are available to the firm provided they are assured of some threshold utility level (their supply price). Formally, the firm will be said to operate in a competitive labor market when it can recruit and keep workers of type n if and only if

$$\operatorname*{Max}_{z}\, u(w(z),z,n) \geq \bar{u}(n). \tag{11}$$

$\bar{u}(n)$ will be the utility available to the worker of type n in alternative employment. If all inequalities (11) hold, the firm knows that each recruit is drawn at random from a population with a given distribution described by $\phi(n)$. If some are violated, the recruit is drawn from that part of the population where (11) is satisfied. The extension of the model to an intermediate case, to be called the monopsonistic case, will be explained later.

The analysis proceeds by means of a series of propositions, the third of which is followed by an important remark.

Proposition 1. For a firm operating in a competitive labor market, with a pay schedule which satisfies (11), in equilibrium

$$\int\{p(z) - w(z)\}\phi(n)dn = 0. \tag{12}$$

Proof. Consider the effect of recruiting one more worker into the firm. The expected change in output per recruit is, by (5),

$$\int p(z)f(z)dz = \int p(z(n))\phi(n)dn,$$

where $z(n)$ is the work-level chosen by a worker of type n. The average labor cost per recruit is

$$\int w(z)f(z)dz = \int w(z(n))\phi(n)dn.$$

It is clear from these two equations that equilibrium is possible only if (12) holds.

In a competitive equilibrium, all firms (at least if identical) pay the same wage schedules, and the constraints (11) hold with equality. Rather than attempt to approach the question of the shape of the equilibrium pay schedule under these conditions directly, I look first at what is, in effect, the opposite extreme. This is a firm that has a given labor force; with the supplementary consideration that workers would leave if wages fell too low, a constraint that will not apply to most of them in equilibrium.

Proposition 2. If a firm maximizes profits for a given labor force, subject to the constraints (11); and if for all n,

$$u(p(z(n)),z(n),n) \geq \bar{u}(n), \tag{13}$$

then for all n

$$w(z(n)) \leq p(z(n)). \tag{14}$$

Proof. For simplicity, proof is confined to the case where a density $f(z)$ exists. Suppose, in order to obtain a contradiction, that the firm has a wage schedule satisfying

$$w(z) > p(z), \text{ for } a < z < b, \tag{15}$$

with $w = p$ at $z = a, b$. Construct a family of schedules $w(z,\epsilon)$ depending differentiably on ϵ, with

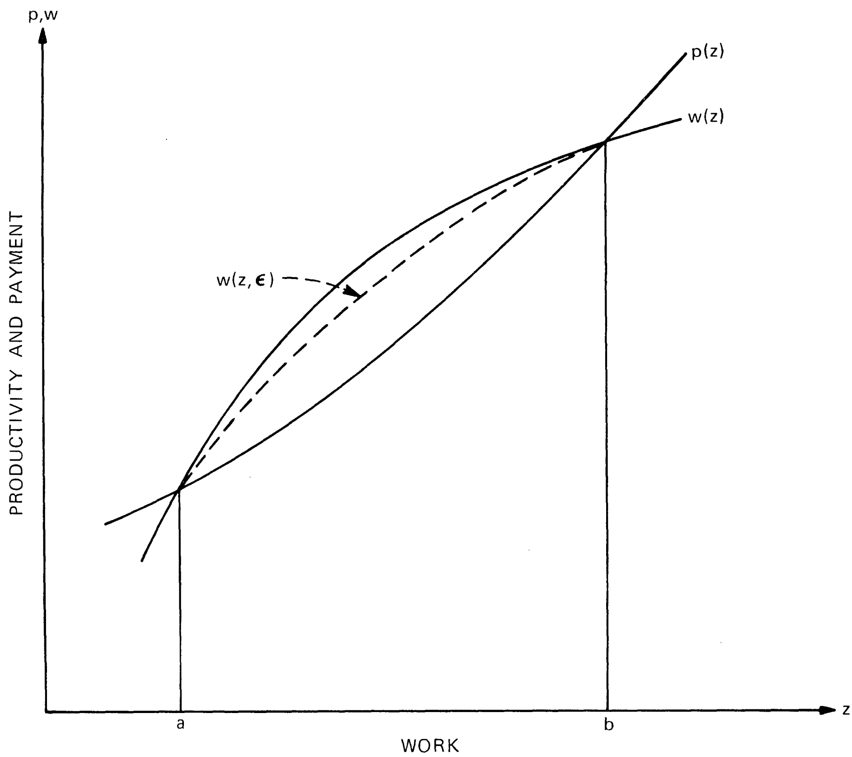
$$\left.\begin{array}{c} w(z,0) = w(z) \\ w(z,\epsilon) = w(z) \text{ when } z \leq a \text{ and } b \leq z \\ w_\epsilon(z,\epsilon) < 0 \text{ when } a < z < b \\ w_z(z,\epsilon) = w_z(z) \text{ when } z = a,\ b. \end{array}\right\} \tag{16}$$

This variation of the pay schedule is illustrated in Figure 1. (13) and (16) imply that, for all small enough ϵ,

$$u(w(z(n),\epsilon),z(n),n) > \bar{u}(n) \qquad (a < z(n) < b).$$

FIGURE 1

VARIATION OF THE PAY SCHEDULE



Therefore, the proposed variation is consistent with the constraints (11), and has no effect on the membership of the labor force.

The variation is to be chosen in such a way that a density function $f(z,\epsilon)$ exists for the distribution of work levels, and so that $f(z,\epsilon)$ is a differentiable function of ϵ. That is why we must insist, in (15), that w remains a smooth function of z as ϵ increases.

Applying equation (5), we have

$$\frac{d}{d\epsilon}\{y - \int w(z,\epsilon)f(z,\epsilon)dz\} = \int[(p - w)f_\epsilon - w_\epsilon f]dz$$

$$= \frac{d}{d\epsilon}\int(p - w)f(z,\epsilon)dz$$

$$= \frac{d}{d\epsilon}\int\{p(z(n,\epsilon)) - w(z(n,\epsilon),\epsilon)\}\phi(n)dn,$$

where $z(n,\epsilon)$ is the work-level chosen by an n-worker when faced with $w(\cdot,\ \epsilon)$. Therefore, the firm would like to increase ϵ if for each n

$$\frac{\partial}{\partial\epsilon}\{p(z(n,\epsilon)) - w(z(n,\epsilon),\epsilon)\} > 0. \tag{17}$$

In order to calculate this expression, we need to know the partial derivative of $z(n,\epsilon)$ with respect to ϵ. A worker of type n chooses z so that $w' = s$ (equation (8)). Differentiating this relationship partially with respect to ϵ, one obtains

$$w'_\epsilon + w'' \cdot z_\epsilon = s_w \cdot (w_\epsilon + w' \cdot z_\epsilon) + s_z \cdot z_\epsilon,$$

where the ϵ-subscript denotes differentiation: w'_ϵ, for example, means $\partial^2 w/(\partial\epsilon\partial z)$. From this it follows that

$$z_\epsilon = \frac{w'_\epsilon - s_w \cdot w_\epsilon}{s_w \cdot w' + s_z - w''}. \tag{18}$$

Using this result, we have

$$\frac{d}{d\epsilon}(w - p) = w_\epsilon + (w' - p')z_\epsilon$$

$$= \frac{(w' - p')w'_\epsilon + (s_w p' + s_z - w'')w_\epsilon}{s_w w' + s_z - w''}, \tag{19}$$

using (18) to substitute for z_ϵ. We want to be able to choose w_ϵ satisfying (16) in such a way that this expression is negative.

Utility maximization implies that the denominator in (19) is non-negative, for this is the condition that the indifference curve lie (locally) above the budget constraint $w(z)$. Let us assume slightly more, *viz.*,

$$s_w w' + s_z - w'' > 0 \quad (a \leq z \leq b). \tag{20}$$

Other (exceptional) cases are presumably easily dealt with. If we now try defining $w(\cdot,\ \epsilon)$ by

$$w_\epsilon = -(w - p)^r \quad (a \leq z \leq b) \tag{21}$$

with $r > 1$, it is readily checked that (16) is satisfied; and substitution in (19) yields

$$\frac{d}{d\epsilon}(w - p) = -\frac{(w - p)^{r-1}}{s_w w' + s_z - w''}\{r(w' - p')^2$$
$$- s_w(w - p)(w' - p') + (w - p)(s_w w' + s_z - w'')\}.$$

The expression within braces is a quadratic form in $w' - p'$, and therefore positive definite if

$$4r(s_w w' + s_z - w'') > s_w^2(w - p). \tag{22}$$

Because of our assumption (20), we can find r to satisfy (22) for all z in the interval. Therefore, the variation defined by (21) does have the desired property when r is chosen large enough.

Thus inequality (17) is established, and it follows that the original payment schedule cannot have been profit-maximizing for the firm. The proposition is proved.

This proposition is best appreciated if we consider one firm in an environment of other firms all paying less than the marginal products for laborers. For this situation, the proposition asserts that the firm will not choose to pay any of its workers more than his marginal product. Thus we may say that incentive considerations in themselves give no reason for paying anyone more than his marginal product. If any worker is to be paid more than his marginal product, it must be by reason of some constraint imposed on the firm. It is not easy to see where such a constraint could come from in an industry of like firms. Indeed, in a special case, a much stronger result can be proved.

Proposition 3. Under competitive conditions, with production possibilities available to each firm described by

$$y = H(\int zf(z)dz), \tag{1}$$

in equilibrium,

$$w(z) = p(z) = H'(\int z(n)\phi(n)dn)z. \tag{23}$$

Proof. In equilibrium with a competitive labor market, all firms have the same pay schedule, and $\bar{u}(n) = \max_z u(w(z),z,n) = u(w(z(n)),z(n),n)$. It will first be shown that, for this pay schedule, $w(z) \leq p(z)$ for all z. The reason is essentially simple: any firm constrained (by competition) to pay more than the marginal product for some work level z_0 can increase its profits by offering less than the marginal product and doing without work-levels in the neighborhood of z_0.

Formally, let us suppose that

$$w(z_0) > p(z_0). \tag{24}$$

z_0 is the work-level supplied workers of type n_0, and we can choose z_0 satisfying (24) so that no other workers supply z_0, if necessary by slightly changing it. Furthermore, we can take it that z_0 is the unique work-level workers of type n_0 are prepared to choose, because of the following lemma.

Lemma. For the case described by (1), the equilibrium wage schedule has the property that for each n, $u(w(z),z,n)$ is maximized by a single value of z.

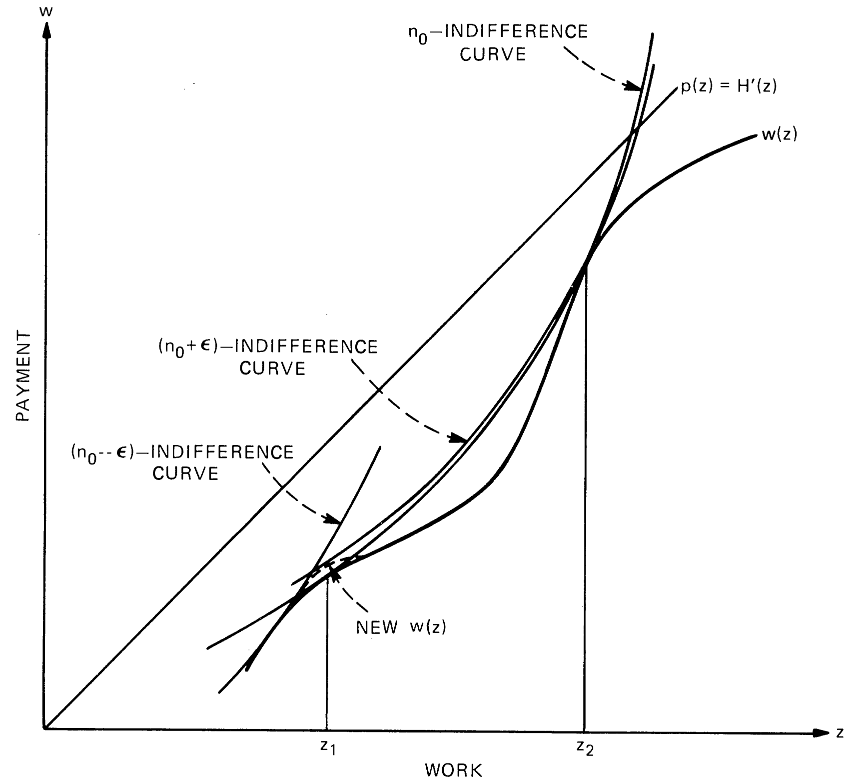
Proof of lemma. Suppose, contrary to the stated result, that z_1 and z_2 both maximize $u(w(z),z,n_0)$. If $w(z_2) - w(z_1) \neq (z_2 - z_1)H'$, suppose without loss of generality that $z_1H' - w(z_1) > z_2H' - w(z_2)$. If w is slightly increased in the neighborhood, a small group with $n \leq n_0$ change from approximately z_2 to approximately z_1, and a small group with $n \leq n_0$ change, but remain close to z_1. The effect on profits of ϵ workers changing from z_2 to z_1 is

$$\epsilon(z_1 - z_2)H' - \epsilon(w(z_1) - w(z_2)),$$

a positive number of order ϵ. The other effects are of order ϵ^2. Thus, profits can be increased by the proposed perturbation. See Figure 2.

FIGURE 2

INCREASING PROFITS WHEN $w(z_2) - w(z_1) \neq (z_2 - z_1)H'$



If $w(z_2) - w(z_1) = (z_2 - z_1)H'$, profits can be increased by raising the pay schedule slightly above the n_0-indifference curve joining $(z_1,w(z_1))$ and $(z_2,w(z_2))$, as shown in Figure 3. This completes the proof of the lemma.

Returning to the proof of the proposition, we consider the effect of reducing the pay schedule for the firm in the neighborhood of z_0 (Figure 4). This can be done in such a way that precisely those n between $n_0 - \epsilon$ and $n_0 + \epsilon$ are now unable to attain $\bar{u}(n)$. For all other n, z remains unchanged. $\int zfdz$ is reduced by $\int_{n_0-\epsilon}^{n_0+\epsilon} z\phi dn$ and the wage bill is reduced by $\int_{n_0-\epsilon}^{n_0+\epsilon} w(z)\phi dn$. Thus, profits are increased by

$$\int_{n_0-\epsilon}^{n_0+\epsilon} w\phi dn - \int_{n_0-\epsilon}^{n_0+\epsilon} z\phi dn \cdot H'(\int z\phi dn). \qquad (25)$$

Now in the case of the production function (1), it is easily seen that

$$p(z) = zH'.$$

Therefore, (25) can be written as

$$\int_{n_0-\epsilon}^{n_0+\epsilon} \{w(z(n)) - p(z(n))\}\phi dn,$$

which is positive, by (24) and the continuity of w and p.

This proves that whenever (24) holds, profits can be increased. Therefore, in equilibrium,

FIGURE 3

INCREASING PROFITS WHEN $w(z_2) - w(z_1) = (z_2 - z_1)H'$

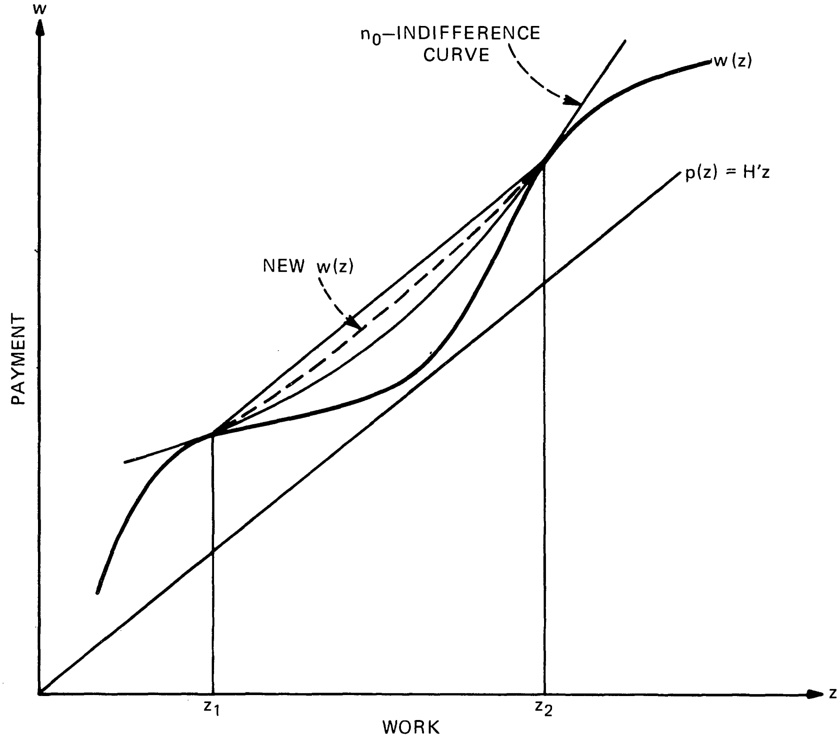
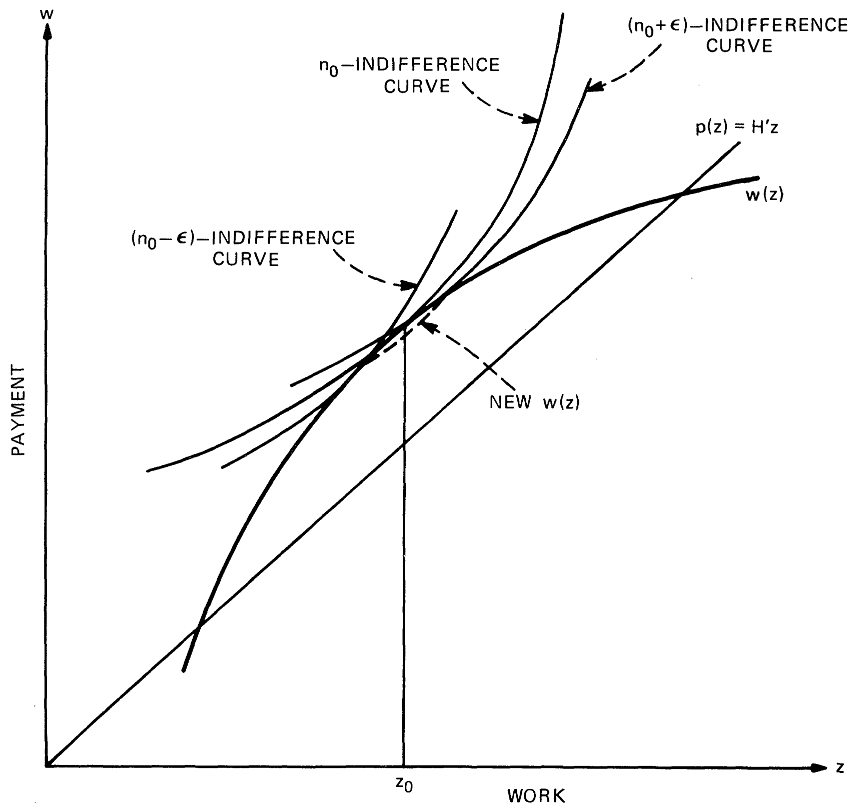


FIGURE 4

LOCAL REDUCTION OF THE PAY SCHEDULE



$$w(z) \leq p(z) \qquad (26)$$

for all relevant z. Combining (26) with the result (12) of Proposition 1, we have in fact

$$w(z) = p(z),$$

as was to be proved.

The proof of this proposition—apart from the lemma which is merely disposing of a special case—does not consider the incentive effects of the pay schedule, but only the effects on the supply of laborers. Indeed, it is little more than the usual argument for equality between wages and marginal products, with special care taken to check that no incentive effects occur when a wage is reduced. But the production function to which the argument applies is a very special one, which, by assuming perfect substitutibility among the different work-activities in the firm, excludes just those features of industrial production that one associates with the existence of job hierarchies. Probably the proof of the proposition could be extended to cover such cases as are suggested in (4) and (4a) above. The argument should fail only when a large change in work-level by a small number of workers has a large effect on marginal productivities. But that is, surely, an interesting and large class of cases.

For the general case, a less precise, but almost equally compelling, argument can be offered. Suppose that, for a range of work-levels, wage-rates are above marginal products. The firm does not reduce these wage-rates if that would lose all the corresponding workers, because it needs such workers. But, to be a little more realistic, a small reduction in wage rates will not suddenly lose all such workers to the firm. A temporary reduction would simply reduce recruiting for a time. Then the change in other marginal products is small and can properly be neglected: the wage reduction increases profits, so long as it does not, and is not expected to, go on too long. There is, then, a general weakness in wage levels if they ever get above marginal products, so that we can expect them to crumble. When this argument is coupled with the implication of Proposition 1, that incentive considerations will not in themselves push wages above marginal products, it is hard to resist the conclusion that, *under competitive conditions, with unidentifiable labor skills, wages and marginal products are equal in equilibrium.*

This conclusion has been based on an industry with identical firms. We should also discuss the equilibrium of a group of firms with *different* production functions taking their workers from the same labor market. If all were recruiting the same mixture of abilities, they would, in general, have different implied marginal productivities, and it would follow from (23) that different firms were paying according to different schedules. If workers are well informed about payment schedules, this is impossible: different wage schedules would sort worker-types among firms. A disequilibrium process of this kind could provide different firms with labor forces of different composition until $p(\cdot)$ is the same for all, and the wage-payment schedule uniform. Thus, equilibrium for an industry is characterized by the orthodox uniformity of marginal productivities and wages. The result, it should be reiterated, has been obtained without the usual assumption that employers know the abilities of those they are hiring.

Since the result is not, despite appearances, the standard one, it is subject to anomalies. The proposition tells us what equilibrium must be like if the firm is recruiting in a perfect labor market, but not that equilibrium exists. I have in mind the possibility that a firm with an established labor force may be able to increase its profits by changing to a payment schedule which will no longer attract new recruits. This will be a possibility if employees cannot expect the same present value of earnings in another firm as those who occupy similar positions to their own. The firm they are working for has acquired information about them (represented by the z they supply) and has placed them in the organization appropriately. They may have spent an initial period with the firm at lower pay, establishing their credentials for promotion. It is not in the interests of their employer to tell a new employer what grades or positions they had reached.

Thus, at least in the case of employees with some seniority, the employer has a monopoly advantage. He may, of course, be unable to exercise it in the face of collective action by his employees. If he succeeds in exercising it, he will cut out recruitment of workers who have sufficient ability to aspire to higher paid positions. The policy is appropriate, therefore, only for a declining firm. It may not be a very realistic possibility. Where it occurs, it means that declining firms have a less steep wage payment schedule than growing firms.

□ **Imperfect labor markets.** The question is thus raised how flat a payment schedule it could profit the firm to adopt bearing in mind the disincentive effects. I examine this in a more general setting, where the labor market is imperfect. Imperfection will be captured by extending assumption (11): the utility level required to attract another worker of ability n is assumed to be an increasing function of the number already employed. Formally, the firm is constrained by

$$\max_z u(w(z),z,n) \geq \bar{u}(n,\phi(n)), \qquad (27)$$

where the utility-threshold $\bar{u}$ is an increasing function of ϕ (as well as of n). As before, the firm wishes to maximize

$$y - \int w(z)\phi(n)dn,$$

where $z = z(n)$ is the value chosen by a worker of ability n.

The techniques appropriate to this problem are those developed for the theory of optimal income taxation [Mirrlees (1971)]. The following nonrigorous development makes it reasonably easy to see where the various conditions come from. First, I express the constraints that, for each n, $z(n)$ maximizes $u(w(z),z,n)$, in a way convenient for the problem. Assuming—as one has no right to do, but we are not being rigorous—that w is differentiable, $u_w w'(z) + u_z = 0$. (Note that $w(\cdot)$ will be so chosen that $z(n) > 0$.) Therefore,

$$\frac{d}{dn} u(w(z),z,n) = u_n(w,z,n), \qquad (28)$$

the *partial* derivative with respect to n. This "envelope condition" is plainly equivalent to the first-order condition; and it is convenient because utility also appears in the constraint (27). Define

$$v(n) = u(w(z(n)),z(n),n), \qquad (29)$$

and express w and u_n as functions of v, z, and n, defined by $u(w,z,n) = v$:

$$w = w(v,z,n), \qquad u_n = \psi(v,z,n). \tag{30}$$

Differentiation of $u(w,z,n) = v$ with respect to v and z shows that

$$w_v = 1/u_w, \qquad w_z = -u_z/u_w = s. \tag{31}$$

Using these results, we then calculate

$$\psi_v = u_{nw}/u_w, \qquad \psi_z = u_{nw}s + u_{nz} = -u_w s_n. \tag{32}$$

After these preliminaries, we can express the problem as

$$\text{maximize } y - \int w\phi(n)dn,$$

subject to

$$v \geq \bar{u}(n,\phi(n)), \quad v'(n) = \psi(v,z,n),$$

and introduce Lagrange multipliers $\lambda(n)$, $\mu(n)$ for the two sets of constraints. Thus,

$$\begin{aligned} L &= y - \int w\phi dn + \int \lambda\{v - \bar{u}(n,\phi)\}dn + \int \mu(n)\{v'(n) - \psi\}dn \\ &= y + \int\{- w\phi + \lambda(v - \bar{u}) - \mu'(n)v - \mu\psi\}dn + \mu(\infty)v(\infty) \\ &\quad - \mu(0)v(0) \end{aligned}$$

is to be stationary when $v(\cdot)$, $z(\cdot)$, and $\mu(\cdot)$ are varied. Differentiating with respect to $z(n)$, we have

$$p'(z)\phi - w_z\phi - \mu\psi_z = 0,$$

or

$$(p'(z) - s)\phi = -\mu u_w s_n. \tag{33}$$

Differentiation with respect to $v(n)$ yields

$$-w_v\phi + \lambda - \mu' - \mu\psi_v = 0$$

or

$$u_w\mu' + u_{nw}\mu = u_w\lambda - \phi; \tag{34}$$

and, finally, differentiation with respect to ϕ gives

$$p - w = \lambda\bar{u}_\phi. \tag{35}$$

(If ϕ were zero, we could have $p - w < \lambda\bar{u}_\phi$.) The way in which p' and p came out in these differentiations may not be quite clear, although it is readily justified by the arguments given earlier. Heuristically, we can take it that *locally* (i.e., for first-order changes), y is a constant plus $\int p(z)\phi(n)dn$. Once this is granted, it is clear where (33) and (35) come from.

There are two further conditions to note, arising because L must be stationary as $v(0)$ and $v(\infty)$ are varied:

$$\mu(0) = \mu(\infty) = 0. \tag{36}$$

Also it should be noted that, since the multiplier λ is associated with the inequality $v \geq \bar{u}$, $\lambda \geq 0$, and vanishes if $v > \bar{u}$. Since by assumption, $\bar{u}_\phi > 0$, (35) implies that, ϕ being positive,

$$p \geq w, \text{ with equality if } v > \bar{u}. \tag{37}$$

(It is pretty clear that equality would be very exceptional.) This confirms and generalizes the earlier result. Intuitively one expects that the firm feels constrained by the need to have $v'(n)$ at least as great as ψ. If it were so, μ would also be nonnegative. But this depends of course on the way in which the other constraint operates.

To see what is going on, let us adopt a definition of the elasticity of supply of workers of type n:

$$\eta(n) = wu_w/(\phi\bar{u}_\phi). \tag{38}$$

This is the percentage increase in $\phi(n)$ obtained for a one percent increase in w; that increase being paid only for $z(n)$. (We ignore in the definition the consequent effect on supplies of other labor-types.) Using (35) and (38), (34) becomes

$$u_w\mu' + u_{nw}\mu = \left(\eta\,\frac{p - w}{w} - 1\right)\phi. \tag{39}$$

The firm's optimal payment schedule is defined by (39) with (33) in the form

$$p'(z) - w'(z) = -\mu u_w s_n, \tag{40}$$

the worker's maximization condition $w'(z) = s$, the labor supply condition $u(w,z,n) = \bar{u}(n,\phi(n))$, and the boundary conditions $\mu(0) = 0 = \mu(\infty)$.

Since n describes a worker's capabilities, we are assuming that

$$s_n < 0, \tag{41}$$

meaning that a man with greater n is more able (or willing) to substitute labor for consumption. Under this assumption, it is clear that w and z must be increasing functions of n. It may not be possible to measure n in such a way that (41) holds; but for the remainder of this section, I assume that it is. Then (40) shows that $\mu > 0$ is equivalent to the wage schedule's being less steep at $z(n)$ than the marginal productivity schedule.

Solving (39) for μ, we obtain

$$\mu(n) = \int_0^n \left(\eta\,\frac{p - w}{w} - 1\right)\beta dm, \tag{42}$$

where

$$\beta = \exp \int_n^m (u_{nw}/u_w)dv \cdot \frac{\phi}{u_w} > 0. \tag{43}$$

$\mu(0) = 0$ is used in deriving (42). $\mu(\infty) = 0$ implies that

$$\int_0^\infty (\eta\,\frac{p - w}{w} - 1)\beta dm = 0. \tag{44}$$

It follows that for some n, $\eta(p - w)/w$ exceeds 1, while for others it is less than 1. It cannot, in interesting cases, equal 1 for all n because that would imply $\mu = 0$, and therefore $p' = w'$, i.e., $w(z) - p(z) =$ constant. This would mean $w =$ constant $x\ \eta$, which we may reject since it is not likely that η increases with n, whereas w certainly does (granted assumption (41)).

Let us assume—as seems plausible—that

$$\eta \text{ is a decreasing function of } n. \tag{45}$$

Is it possible that μ is negative for small n? If it were,

$$\alpha = \eta \frac{p - w}{w} - 1 \tag{46}$$

would be negative for small n. Now, since μ is negative, $p' - w' < 0$, and, therefore, $p - w$ is decreasing in n; while w is increasing in n, and η decreasing. Therefore, α is decreasing, and remains negative. The argument would continue to apply, and α would be negative for all n. This is, because of (44), impossible. The argument used actually tells us that, once μ is negative for some n_0, it remains negative for all $n \geq n_0$. Thus we have the following:

Proposition 4. Assuming (41) and (45), there exists n_0 such that (α being defined by (46))

$$\left.\begin{array}{ll} \alpha \geq 0 & (n \leq n_0) \\ \alpha \leq 0 & (n \geq n_0) \end{array}\right\} . \tag{47}$$

Furthermore, for all n,

$$\mu \geq 0; \tag{48}$$

and, for all z,

$$w'(z) \leq p'(z). \tag{49}$$

The second part of the proposition follows at once from the first, for (by (44)) we can write μ either as $\int_0^n \alpha\beta dm$ or $-\int_n^\infty \alpha\beta dm$. (47) has an interesting interpretation, obtained by writing (46) in the form

$$\frac{p - w}{w} = \frac{1}{\eta}(1 + \alpha). \tag{50}$$

If incentive considerations were ignored, the optimal mark-up of marginal product above wage would be $1/\eta$ for each n. (47) and (50) tell us that the incentive considerations imply a larger mark-up than monopsony theory suggests for the less skilled, and a lower mark-up for the more skilled. At the same time, the absolute difference between marginal product and wage is higher for the more skilled: this is the result of assumption (45).

This section has been addressed to the importance of incentive considerations in determining the pay structures of profit-maximizing firms. It has been shown that the firm must have the usual kind of monopsony power in the labor market before it has much reason to pay attention to the incentive effects of its pay structure. The one exception to this is the stagnant firm, and that can be regarded as a case of monopsony where there is some inelasticity in response to wage reductions but not to wage increases. In the monopsony case, which I take to be in some degree realistic, no worker receives more than his marginal product, but workers of higher capability receive a wage which is less close to their marginal product than those of lower ability. The last conclusion depends on assuming that higher ability is associated with lower elasticity of supply; which should be true at least because of the investment in reputation, and knowledge of the firm, which a worker of higher ability normally makes. It is interesting that, if firms have imperfect knowledge of their recruits' capabilities, and recruit in imperfect markets, they should in theory apply a progressive tax to marginal products before paying wages, and thus do some of an egalitarian government's work for it.

One can use the equations presented to compute optimum payment schedules in particular cases, but I do not pursue the model further here. It leaves out a very great deal of what interests us about organizations. Presumably this model ought to be developed in the direction of allowing the firm some limited information about the men it hires; but it is unlikely that the principal proposition would be much affected. In any case, it would take us further away from the internal organization of the firm. The chief shortcoming of the model is the lack of an explicit reason for workers to work together, leaving information considerations highly implicit: therefore, there are no reasons for the hierarchical organization that one observes, and would like to explain, for itself, and as an important influence on the pay structure of firms. In fact, when they join a firm, workers are uncertain what they will be doing, and what reward they will receive; but their relationships are quite mechanical. In the remaining sections of the paper, I study models that do not deal explicitly with the distribution of skills, but do deal with the uncertainties of production activity and supervision.

3. Principal and agent

■ Consider a man with von Neumann-Morgenstern utility function $u(x,z)$, where x is the payment received as a result of the work he does and z is the work done. Payment is taken to be an uncertain consequence of the work done. It is a function of the man's performance as observed by his principal. Observed performance may or may not accurately represent the value to the principal of the agent's work; and the value of the work, which I shall call output and denote by y, may itself be a random variable conditioned by z. The accuracy of the principal's observation of y depends upon the time devoted to making the observation. (It should also depend upon the agent's own efforts to affect the observation; but I shall ignore this, important though it may be. I am not sure how best to model the effect of possibly competing efforts to influence observational precision.)

The formal model of the agent is that

$$z \text{ maximizes } U(z) = E_\epsilon E_{Y|z} u\left\{\phi\left(Y + \frac{1}{\theta}\epsilon\right), z\right\}$$
$$= \int\int u\{\phi(y + \frac{1}{\theta}\epsilon), z\} f(y,z) g(\epsilon) dy d\epsilon, \qquad (51)$$

where ϕ is the schedule governing payment to the agent, and θ^2 is the time spent (by the principal) on observation—the idea being that he samples repeatedly. Observational errors and output uncertainty are assumed independent. It should also be emphasized that all functions and random variables occurring are supposed to be nice: points of rigor will be ignored, although they are very important. Necessary conditions for maximization in (51) are

$$U'(z) = 0 \qquad (52)$$

and

$$U''(z) \leq 0. \qquad (53)$$

The principal handles the output of the agent and makes payment to him. It is convenient to assume that the two are commensurable, and to give the principal a utility function $v(y - \phi, \theta^2)$. Thus, he would like to choose $\phi(\cdot)$ and θ to maximize

$$E_\epsilon E_{Y|z} v\left\{Y - \phi\left(y + \frac{1}{\theta}\epsilon\right), \theta^2\right\}$$
$$= \int\int v\left\{y - \phi\left(y + \frac{1}{\theta}\epsilon\right), \theta^2\right\} f(y,z) g(\epsilon) dy d\epsilon, \quad (54)$$

given that z is determined by (51).

One of the odd features of this assumption is that a principal who benefits from y cannot observe it, although one might think that a man can benefit only from what he observes. But notice that the assumption would look much more reasonable if the principal had many agents all contributing to the output he receives—and we shall come to that case. In any event, it is common for benefits to accrue long after payments have been made irreversibly; and $Y + \frac{1}{\theta}\epsilon$ should properly be regarded as the agreed basis for payment, which has to be adhered to whatever output may actually be.

Another odd feature is that there is no very evident reason for the principal-agent relationship in the model. But it is implicit that the agent uses assets owned by the principal. There seems to be no advantage in making that explicit unless one wants to consider the decision of how much to let the agent use. Interesting though that may sometimes be, it adds nothing to the picture of the relationship that I want to convey.

Returning to the mathematical problem, we must include another constraint on the principal's maximization—that the agent gets expected utility sufficient to induce him to accept the contract. In other words, there should be a supply price for the kind of labor we are considering. (In the case of slavery, utility would not be the relevant variable constrained; and indeed morale, motivation, and health are considerations that should perhaps be brought in, but they would complicate matters further.) The constraint is

$$U(z) \geq A, \quad (55)$$

A being a fixed number.

The analysis now proceeds along lines I have used elsewhere.[1] Consider the Lagrangean form obtained by assigning undetermined multipliers λ and μ to (52) and (55):

$$L = \int\int \{v(y - \phi, \theta^2) + \lambda u(\phi, z) + \mu(u_2 + u f_z/f)\} f g \, dy d\epsilon.$$

(Here and below, numerical subscripts to u and v denote differentiation with respect to the indicated argument.) Differentiating with respect to the scalars z, and the function $\phi(\cdot)$, we obtain first-order conditions:

$$\frac{\partial L}{\partial z} = \int\int v f_z g \, dy d\epsilon + \lambda U'(z) + \mu U''(z) = 0,$$

i.e.,

$$\mu = - U''(z)^{-1} \int\int v f_z g \, dy d\epsilon. \quad (56)$$

Before taking the other derivatives, it is advantageous to change a variable of integration and write

$$L = \theta \int\int \{v(y - \phi(x), \theta^2) + \lambda u(\phi(x), z) + \mu(u_2 + u f_z/f)\} f(y,z) g(\theta(x - y)) dy dx. \quad (57)$$

[1] Mirrlees (1972b, 1974). See also Spence and Zeckhauser (1971).

Then

$$\frac{\partial L}{\partial \theta} = 0. \tag{58}$$

(Since I shall not be computing solutions, I do not evaluate this derivative explicitly: it is clearly rather a complicated expression.)

$$\frac{\partial L}{\partial \phi(x)}$$
$$= \theta \int \{-v_1 + \lambda u_1 + \mu(u_{12} + u_1 f_z/f)\} f(y,z) g(\theta(x - y)) dy = 0.$$

Since u and its derivatives are independent of y, this may be rewritten in the form

$$\frac{1}{u_1} \int v_1 fg dy / \int fg dy = \lambda + \mu \left\{ \frac{u_{12}}{u_1} + \int f_z g dy / \int fg dy \right\},$$

or equivalently, using the notation $E(\cdot\,|x)$ for expectations conditional upon a given value x of observed (apparent) output,

$$\frac{E(v_1|x)}{u_1} = \lambda + \mu \left\{ \frac{u_{12}}{u_1} + E\left(\frac{f_z}{f}\middle| x\right) \right\}. \tag{59}$$

(56), (58), and (59), along with (52), effectively determine equilibrium, which is characterized by numbers θ and z, and the function ϕ. Given θ and z, (59) determines ϕ. Thus, although explicit solution of actual cases is difficult, one can use (59) to find out something about the shape of ϕ.

Before discussing that, we need some assumptions about f. Since larger z is supposed to mean increased effort, and therefore, on average, output, it should decrease f for small y and increase it for large. I shall furthermore assume that

$$f_z/f \text{ is an increasing function of } y. \tag{60}$$

Since, for all z, $\int f dy = 1$, $\int f_z dy = 0$; thus (60) implies that f_z/f is negative for small y, positive for large. A similar assumption is made for g, modelled on the case where ϵ is a normal random variable:

$$\log g \text{ is a concave function of } \epsilon. \tag{61}$$

Lemma. $E(f_z/f|x)$ is a nondecreasing function of x.

Proof.

$$\frac{\partial}{\partial x} E\left(\frac{f_z}{f}\middle| x\right) = \frac{\partial}{\partial x} \{\int f_z g(\theta(x - y)) dy / \int fg dy\}$$
$$= \theta\{f_z g' dy / \int fg dy - \int f_z g dy \int fg' dy / (\int fg dy)^2\}$$
$$= \theta\left\{ E\left(\frac{f_z}{f}\frac{g'}{g}\middle| x\right) - E\left(\frac{f_z}{f}\middle| x\right) E\left(\frac{g'}{g}\middle| x\right) \right\}. \tag{62}$$

Now by assumptions (60) and (61), f_z/f and g'/g are both increasing functions of y, and, therefore, positively correlated, given x. Therefore, by (62)

$$\frac{\partial}{\partial x} E\left(\frac{f_z}{f}\middle| x\right) \geq 0, \tag{63}$$

as was claimed. Q.E.D.

This allows us to deduce that, if (i) $\mu > 0$, (ii) u_{12}/u_1 is a nonincreasing function of ϕ, and (iii) u and v are concave in their first arguments, then

$$\phi'(x) > 0. \tag{64}$$

To prove, we consider $E(v_1|x)/u_1$ as a function of ϕ and x, $a(\phi,x)$. An argument exactly similar to that of the lemma shows that $a_x \leq 0$, while concavity of u and v implies that $a_\phi > 0$. From (59), we have

$$\left\{a_\phi - \mu\frac{\partial}{\partial\phi}(u_{12}/u_1)\right\}\phi'(x) = \mu\left\{\frac{\partial}{\partial x} E(f_z/f|x) - a_x\right\}, \tag{65}$$

which, with all our assumptions, implies as desired that $\phi' \geq 0$. Assumption (iii) is fairly unexceptionable; and something like (ii) is bound to be required. But (i) is not the kind of thing we should assume: it should be deduced. From (56)—bearing in mind that $U'' \leq 0$ (and ignoring the exceptional possibility that $U'' = 0$)—we see that $\mu > 0$ if and only if the principal would like z to be further increased, if he could control it directly, the payment schedule being fixed. This is plausible, certainly: one expects the payments to be so arranged that the agent does not do all the principal would like. But I have been unable to find general assumptions that exclude the possibility $\mu < 0$.

Certain special cases are of interest. Since it is necessary to simplify drastically in order to get much further, I shall from now on assume that

$$u_{12} = 0. \tag{66}$$

Case I: Only output is uncertain.

In this case, ϵ is equal to zero with probability one, and (59) becomes, since x and y are identical,

$$\frac{v_1(y - \phi(y))}{u_1(\phi(y))} = \lambda + \mu\frac{f_z(y,x)}{f(y,z)}. \tag{67}$$

In this case, if $\mu < 0$, ϕ must be a decreasing function of y; and

$$\begin{aligned}\int\int v f_z g\, dy\epsilon &= \int v(y - \phi(y))f_z(y,z)dy \\ &= \int\{v(y - \phi) - v(y_0 - \phi(y_0))\}f_z dy \text{ (since } \int f_z dy = 0),\end{aligned}$$

where y_0 is a value of y for which $f_z = 0$. Thus by (60), $\int\int v f_z g\, dy d\epsilon > 0$ if ϕ is decreasing. But this by (56) implies that $\mu > 0$—a contradiction. Thus, in this case ϕ necessarily increases in y. Indeed, since $\mu > 0$, we can say more. Differentiating,

$$\frac{v_{11}}{u_1}(1 - \phi') - \frac{v_1 u_{11}}{u_1^2}\phi' > 0.$$

Therefore,

$$\begin{aligned}\phi' &> \frac{-v_{11}/v_1}{-v_{11}/v_1 - u_{11}/u_1} \\ &= \frac{\text{absolute risk aversion of principal}}{\text{sum of absolute risk aversions}}.\end{aligned} \tag{68}$$

The marginal share going to the agent will be low only if his absolute risk aversion is significantly greater than that of the principal. (70), of course, allows us to look at the schedule in more detail. But the second case is perhaps more interesting for the internal organization of firms.

Case II: Only performance observation is uncertain.

In this case, $y = z$ with probability one. This is a limiting case of the general analysis, and one must either go back to the beginning, or confidently substitute a delta-function for f in (59). In any case the result is

$$\frac{v_1(z - \phi(x), \theta^2)}{u_1(\phi(x))} = \lambda - \mu \frac{g'(\phi(x - z))}{g}. \tag{69}$$

By the same kind of argument as in Case I, but using assumption (61), we can show again that $\mu > 0$.

This is the model I shall use in the rest of the paper, so it is time I attended to one important feature of the analysis that has been neglected. If one considers a normal distribution for ϵ, $g(\epsilon) = (2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}\epsilon^2}$ and

$$\frac{g'}{g} = -\epsilon. \tag{70}$$

Since ϵ varies from $-\infty$ to $+\infty$, the right-hand side of (69) is sometimes negative. This is a trifle upsetting, since the left-hand side is always positive (assuming u and v are increasing functions of income). It might be thought that this trouble arises because, with a normal error, observed output can be negative; but the same thing happens if $x = ze^{\epsilon/\theta}$. What we have to do is allow for the possibility $\phi = 0$ explicitly. Maximization of the Lagrangean with respect to ϕ yields an inequality if $\phi = 0$:

$$-v_1 + \lambda u_1 - \mu u_1 g'/g \leq 0. \tag{71}$$

If $\lambda - \mu g'/g < 0$, (71) holds with strict inequality, and consequently $\phi = 0$.

We can see this explicitly if we consider an example where both principal and agent have constant absolute risk aversion:

$$u_1 = e^{-\alpha\phi}, \qquad v_1 = e^{-\beta(z-\phi)}. \tag{72}$$

Using normal g, we find that

$$\left.\begin{aligned} \phi(x) &= \frac{\alpha z}{\alpha + \beta} + \frac{1}{\alpha + \beta}\log\left(\lambda + \mu\theta(x - z)\right) \\ &\qquad\qquad (x \geq x_0 = z + (e^{-\alpha z} - \lambda)/(\mu\theta)) \\ &= 0 \qquad\qquad (x \leq x_0) \end{aligned}\right\}. \tag{73}$$

The form of this schedule is shown in Figure 5. In passing, it should be noted that the case where u is $-\infty$ when $\phi = 0$, which I have discussed elsewhere in the context of insurance and incentives (1974), is even more peculiar: the risk of receiving no payment is made to do all the work of providing incentives. But this is not the appropriate assumption in the present context.

To show that the upper part of the payment schedule is not necessarily concave, I note two other special cases:

(i) *Principal risk-neutral; agent constant relative risk aversion* ρ:

$$\phi(x) = [(B + Cx)^{1/\rho}]_+ . \tag{74}$$

B and C are positive constants, determined by other aspects of the model.

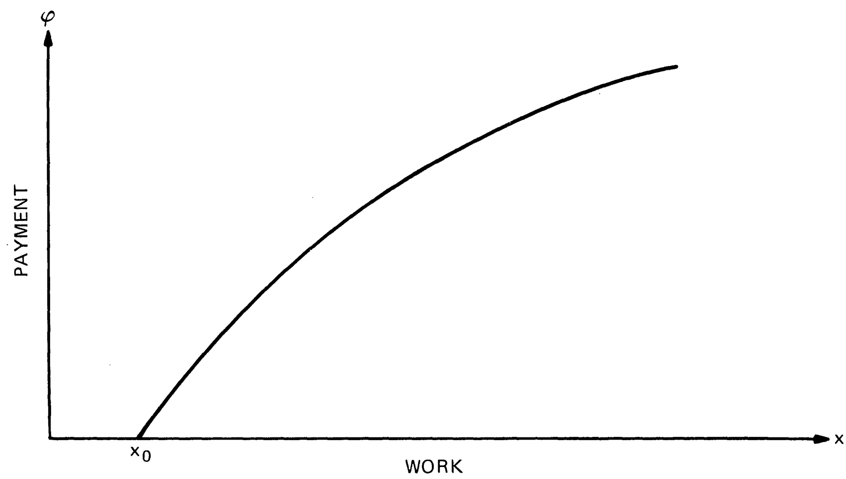
(ii) *Principal and agent have same constant relative risk aversion* ρ:

$$\phi(x) = [z\{1 + (B + Cx)^{-1/\rho}\}^{-1}]_+ . \qquad (75)$$

Both (74) and (75) are concave in their positive sections if $\rho \geq 1$.

FIGURE 5

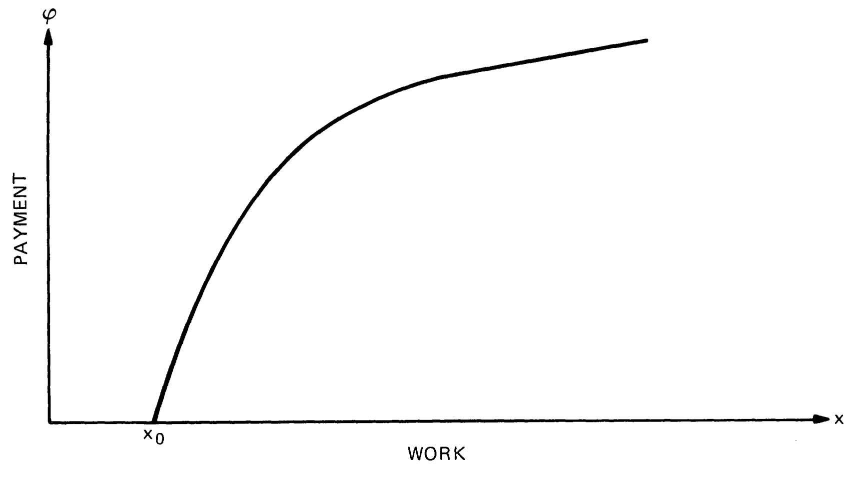
AN OPTIMAL PAYMENT SCHEDULE



One wants to know how the schedules vary as the chosen θ, and the agent's elasticity of substitution between income and effort, vary. Unfortunately that is complicated in all the examples I have looked at, because of the awkward form of condition (56). From (73), it can be seen that the payment schedule takes a fairly sharp "logistic" form as in Figure 6 if $\frac{\mu\theta}{\alpha + \beta} e^{\alpha z}$ and $e^{\alpha z}$ are both large. (These are the absolute values of the first and second derivatives of ϕ at x_0.) This suggests that large risk aversion on the part of the agent combined with accurate observation by the principal (reflected in large θ) give a schedule that is a bit like the effect of an *instruction*. An instruction is, in effect, a promise that rewards will vary rapidly in the neighbor-

FIGURE 6

A PAYMENT SCHEDULE APPROXIMATING AN INSTRUCTION



hood of a particular output level, but not elsewhere. Unfortunately, this suggestion could be confirmed only if we knew how μ depended on α, β, and θ: that is the difficult mathematical problem.

In any case, it is clear that the schedule will only exceptionally take the really sharp form of an instruction. It is not surprising that such policies are generally suboptimal where men are uncertain about what they may achieve and about what they may be seen to have achieved. What is notable is that the optimal payment schedule is usually as "unfair" as that shown in Figure 5, in circumstances where one might have expected liberal allowance for unfavorable observational inaccuracies.

In the remaining sections of the paper I shall outline how this simple two-person model can be used as a building block in constructing a model of a complicated organization.

4. Two-level organizations

■ Suppose now that the principal has n men to supervise, all identical, who as before choose z to maximize $U(z)$; and suppose outputs add (i.e., constant returns). The principal (following Case II of the previous model) receives $nz - \sum_{i=1}^{n} \phi(z + \epsilon_i/\theta)$, there being independent observation of the n agents. He would like to choose ϕ and θ so as to maximize

$$V = Ev(nz - \Sigma\phi, n\theta^2). \tag{76}$$

A small change in $\phi(x)$ changes V by an amount proportional to

$$-nE_2 \ldots E_n v_1(nz - \phi(x) - \sum_{i=2}^{n} \phi(X_i)), \tag{77}$$

E_i being the expectation operator for the random variable $X_i = z + \epsilon_i/\theta$. Therefore the first-order condition for ϕ is (cf. (59))

$$\frac{nE_2 \ldots E_n v_1}{u_1} = \lambda + \mu\left(\frac{u_{12}}{u_1} - \frac{g'}{g}\right) \quad (x > 0). \tag{78}$$

The effect of having numerous subordinate workers is that the payment schedule is governed by an average of v_1.

If the principal has constant absolute risk aversion, the averaging drops out conveniently:

$$n\{Ee^{\beta\phi(X)}\}^{n-1}e^{-\beta(nz-\phi(x))}\frac{1}{u_1} = \lambda + \mu\left(\frac{u_{12}}{u_1} - \frac{g'}{g}\right). \tag{79}$$

In this case the optimal payment schedule has the same form as that discussed at the end of the previous section. There is no reason to think it would get flatter or steeper as the number of subordinates increases. But if absolute risk aversion is decreasing, it is a different matter, because the relative aggregate riskiness of the principal's receipts diminishes as n increases. In that case, for large n, we can regard the principal as effectively risk-neutral. He chooses ϕ and θ to maximize

$$v(nz - nE\phi(z + \epsilon/\theta), n\theta^2), \tag{80}$$

and the first-order condition for the payment schedule is of the form

$$\frac{1}{u_1} = \lambda' + \mu'\left(\frac{u_{12}}{u_1} - \frac{g'}{g}\right) \quad (x > 0). \tag{81}$$

The effect of this averaging is, roughly, to reduce the concavity of the payment schedule. With our earlier examples,

$$u_{12} = 0, \quad g = (2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}\epsilon^2},$$

relative risk aversions ρ, σ

$$\phi'' = -\phi'^2\left[\frac{\rho(\rho - 1)}{\phi^2} + \frac{\sigma(\sigma + 1)}{(z - \phi)^2} + \frac{2\rho\sigma}{\phi(z - \phi)}\right]. \tag{82}$$

Thus $\rho < 1$ implies convexity if $\sigma = 0$, but concavity (except for small x) if $\sigma > 0$.

Without doing the mathematics, we can see that the model in this form provides a reason for not increasing the size of the organization beyond a certain point. If the principal chooses n as well as θ, he will not be willing to increase n without limit. If n is given, there may be no solution that gives the principal sufficient utility to undertake the task. The same would be true with increasing returns to the number of agents (workers) cooperating. A model of this kind would also enable one to consider under what circumstances it would pay a group of workers to have one of their number undertake all the performance observation, and when it would pay instead to have a symmetric solution in which each worker devotes some of his time to "monitoring" (one of the others, presumably). It is not obvious that the asymmetric solution outlined here, and assumed optimal by Alchian and Demsetz (1972), is in fact optimal when the means of production are owned in common.

5. Hierarchical organizations

■ When the number of workers becomes large, it presumably pays to establish an inspection hierarchy. It can be modelled as follows. I begin with the worker level, and then continue through successive supervisory levels:

(i) $$z \max Eu^1(\phi_1(X^1),z),$$

where

$$X^1 = z + \frac{1}{\theta_1}\epsilon^1.$$

(ii) $$\phi_1,\theta_1 \max Eu^2(\phi_2(X^2),n_1\theta_1^2),$$

where

$$X^2 = n_1 z - \sum_{i=1}^{n_1}\phi_1(X_i^1) + \frac{1}{\theta_2}\epsilon^2 = Z^2 + \frac{1}{\theta_2}\epsilon^2.$$

(iii) $$\phi_2,\theta_2 \max Eu^3(\phi_3(X^3),n_2\theta_2^2),$$

where

$$X^3 = \sum_{i=1}^{n_2} Z_i^2 - \sum_{i=1}^{n_2}\phi_2(X_i^2) + \frac{1}{\theta_3}\epsilon^3;$$

and so on, until the last level

$$\phi_{t-1},\ \theta_{t-1} \max\ Eu^t(Z^t, n_{t-1}\theta_{t-1}{}^2),$$

where

$$Z^t = \sum_{i=1}^{n_{t-1}} Z_i^{t-1} - \sum_{i=1}^{n_{t-1}} \phi_{t-1}(X_i^{t-1}).$$

At each stage, the maximization is carried out subject to the maximizations at earlier stages, and subject to the supply price (in utility terms) of the men at each level being satisfied.

In this specific model, it is assumed that aggregate output is correctly measured at each level, and that efficient accountants set the wage bill off against output in every department. The n_i may also be chosen, either at each level, or at the top.

It is plain without undertaking mathematics that a larger viable organization can be created in this way than is possible when restricted to two levels, because, by fixing the numbers supervised by any individual, and the supervision time per subordinate, and setting up some simple payment rule (such as fixed proportions of the departmental net income), we should be able to satisfy the utility constraints. There is then no reason why the organization should not be increased in size without limit.

It requires more careful analysis to see whether, despite the obvious disadvantage of increasing supervisory staff, a large organization of the kind under discussion may have compensating advantages. It can be shown that if (i) each supervisor has the same number of immediate subordinates, (ii) errors of observation are proportional to the mean size of the observed variable, (iii) the payment rule is proportional at each level, and (iv) all individuals are identical, then there is a payment system such that each supervisor gets at least as much expected utility as a (first level) worker and the income of the proprietor (at the top) is approximately

$$Z^t \sim (AN + B + N^{\frac{1}{2}}\,\epsilon)\ z\ (z = \text{value of output per worker}),$$

where N is the number of workers, A and B are constants, and ϵ is a random variable that is, likewise, independent of N. Optimization would, of course, achieve more. (I hope to present proof of this result, and further analysis of the model elsewhere.)

Ignoring errors, this result implies that profit per worker diminishes as N increases. But if the cost of other inputs (such as capital) is less than zA per man, the entrepreneur will prefer a larger organization, provided that his relative risk aversion does not decrease too rapidly. Perfect competition could (if B is large enough) force zA and other costs into equality and encourage small firms, but very little weakening of competitive forces is required to render the larger organization not only viable but more profitable. Any increasing returns would strengthen the tendency.

The model suggests another interesting conclusion. If we suppose that the members of the organization have constant or decreasing relative risk aversion (not that that is so very likely), it seems to follow, since relative income riskiness is smaller at higher levels, that payment schedules should be less concave at the higher levels—less like instructions, and more like profit-sharing. The reader will not need to be reminded that, considering the conjectural basis of the reasoning, and the restrictive nature of the model, this is not a very well-founded conclusion. But it is the kind of conclusion that one could hope to obtain from such a model as this one.

6. The scale of firms

■ It has been suggested that the scale of firms is limited by the capabilities of managers, by the intrinsic difficulty of controlling large organizations, and by the communication loss resulting from long chains of authority. The model of the previous section made no explicit allowance for the managerial skills that may be required of men in charge of large departments. In fact the model was motivated by a desire to see what kinds of complexity might arise in a large organization. It gives a fairly routine role to everyone in the firm: men either work or observe the results of what others have done; and of course they choose what to do. In this way information about the value of work—which may be interpreted as the value of different kinds of work—gets conveyed through the organization, from apex to workers. There is no apparent reason why the top man's job should be any harder than that of anyone else; no reason why, in the optimum policy, middle managers should get more than workers. If there are reasons why higher-level jobs in a hierarchy require more ability than lower-level ones, they have largely escaped the model. Perhaps the best candidate is the ability to take decisions: for the choice of payment schedule—which stands, of course, for advice and instruction, redeployment, and even encouragement and discipline—surely has greater effects at higher levels than at lower ones. This consideration may also support the second reason for diminishing returns mentioned above. An adequate model for skill in decision-taking remains to be built.

Communication losses are the basis of Williamson's analysis of the size of firms (1970). The evidence quoted by him was experimental; but it seems plausible that information is lost in long chains. Yet, the model in the previous section gives an example of how information may be conveyed where at each link in the chain, observation is imperfect, everyone is choosing what to do on the basis of self-interest, and as a result creates an incentive system which conveys information to the next man down. The transmission of information is imperfect, because, for example, no one has an incentive simply to obey an instruction or to pass it on. Yet it turns out that the contribution of the organizational structure to diminishing returns is very weak, becoming for large organizations essentially negligible. Uncertainty in communication does not necessarily imply increased losses in proportion to the size of the organization.

7. Final remarks

■ Two different models of an organization have been discussed, complementing other kinds of models that have appeared in the literature. In these models, imperfect information binds the organization together. They are therefore very much in the spirit of the Alchian-Demsetz paper referred to at the beginning. The models—particularly in the latter part of the paper—seem at present too cumbersome to answer many of the questions one would ask of them. For example, the analysis of optimal payment schedules is seriously incomplete, and nothing has been said about the optimum shape of the organization (how $n_1, n_2, \ldots, n_{t-1}$ should compare to one another).

Apart from the interest that the models may have as pictures—

which I hope is not negligible—the most important conclusions (and suggestions) of this paper may be summarized as follows:

(1) Imperfect information about employees is not in itself enough to explain deviations from the elementary conclusions of the theory of the firm, that wage rates are equal to marginal products.
(2) Where (because of labor market monopsony) a firm's pay structure is devised to encourage work and the demonstration of ability, the marginal product of a more highly skilled man exceeds his wage by more than for less skilled workers, although the payment schedule is steeper than pure monopsony theory implies.
(3) For the firm as a whole, a hierarchical structure does not necessarily impose decreasing returns to scale.
(4) Within the firm, optimal payment schedules for individual employees would normally pay nothing whenever apparent performance falls below a certain point (this was a Pareto-optimality result—it is in the worker's interest, given the supervisor's utility level).
(5) The optimizing model used in the paper suggests some tendency for the giving of orders (rather than allowing "initiative") to be more nearly optimal at lower levels in a hierarchy, and wherever the principal is strongly risk averse.

References

ALCHIAN, A. A. AND DEMSETZ, H. "Production, Information Costs and Economic Organization." *The American Economic Review,* Vol. 62, No. 5 (December 1972).

ARROW, K. J. *Essays in the Theory of Risk-Bearing.* London and Amsterdam: North-Holland, 1972.

———. *The Limits of Organization.* New York: W. W. Norton, 1974.

CHEUNG, S. "Private Property Rights and Share-Cropping." *Journal of Political Economy* (1968).

MARSCHAK, J. AND RADNER, R. *Economic Theory of Teams.* New Haven: Yale Univ. Press, 1972.

MIRRLEES, J. A. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* (April 1971).

———. "On Producer Taxation." *Review of Economic Studies* (January 1972a).

———. "Population Policy and the Taxation of Family Size." *Journal of Public Economics* (August 1972b).

———. "Notes on Welfare Economics, Information and Uncertainty" in M. S. Balch, D. L. McFadden and S. Y. Wu, eds., *Contributions to Economic Analysis.* Oxford and Amsterdam: North-Holland, 1974.

PAULY, M. "The Economics of Moral Hazard." *The American Economic Review,* Vol. 58, No. 3 (June 1968).

ROSS, S. A. "The Economic Theory of Agency: The Principal's Problem." *The American Economic Review,* Vol. 63, No. 2 (May 1973).

SIMON, H. A. *Administrative Behavior.* New York: Free Press, 1957.

SPENCE, A. M. "Job Market Signalling." *Quarterly Journal of Economics* (August 1973).

——— AND ZECKHAUSER, R. "Income, Information, and Individual Action." *The American Economic Review,* Vol. 61, No. 2 (May 1971), pp. 380–387.

STIGLITZ, J. E. "Incentives and Risk Sharing in Sharecropping," *Review of Economic Studies* (April 1974).

WILLIAMSON, O. E. "Hierarchical Control and Optimum Firm Size." *Journal of Political Economy* (April 1967). Reprinted in D. Needham, ed., *Readings in the Economics of Industrial Organization.* New York: Holt, Rinehart and Winston, 1970.

WILSON, R. "The Theory of Syndicates." *Econometrica* (January 1968).

ZECKHAUSER, R. "Mutual Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory* (March 1970).