

UNIVERSITY OF CALIFORNIA, BERKELEY

Department of Economics

Berkeley, California 94720

Working Paper 8741

EMPLOYMENT CONTRACTS, INFLUENCE ACTIVITIES
AND EFFICIENT ORGANIZATION DESIGN

Paul R. Milgrom

May 13, 1987

Key words: Rent seeking, influence, transaction costs, organization, incentives, labor contracts.

JEL Classification: 511, 026, 053

EMPLOYMENT CONTRACTS, INFLUENCE ACTIVITIES
AND EFFICIENT ORGANIZATION DESIGN

Paul R. Milgrom

When changing jobs is costly, efficient employment contracts rarely compensate workers fully for the effects of post-hiring events and decisions. If there are executives and managers with authority to make discretionary decisions that other employees care about, those other employees will be led to waste valuable time attempting to influence decisions. Efficient organization design counters this tendency by limiting the discretion of decision-makers, especially for those decisions with large distributional consequences but little importance to the organization.

1. Introduction

Experience suggests -- and most Western economists believe -- that some degree of market-like decentralization is necessary to encourage innovation and efficient resource use in an economic system. The source of these advantages, however, has proved difficult to pinpoint. Why can't a centrally planned, socialist economy mimic a decentralized one whenever that is desirable? In his study of "The Nature

An earlier draft of this paper was entitled "Quasirents, Influence and Organization Form."

I owe a special debt of thanks to John Roberts for allowing me to incorporate the fruits of joint research into this paper, for suggesting some of the reported applications and references, and for his thorough and constructive commentary on the previous versions. For their helpful comments, I also thank Drew Fudenberg, Bengt Holmstrom, Roger Kormendi, Ed Lazear, Maggie Levenstein, Rick Levin, Meg Myer, Jim Mirrlees, Charlie Plott, Ariel Rubinstein, and Jean Tirole, my research assistants Byung-Il Choi and Bernard Desgagne, and two anonymous referees. I gratefully acknowledge the financial support of the National Science Foundation, the Guggenheim Foundation and the Sloan Foundation.

of the Firm," Coase [1937] posed the corresponding question: "Why is not all production carried out by one big firm?"

I shall argue that there are costs, called "influence costs," that attend any increase in centralized control, whether in a firm or in a larger economic system. These costs arise because participants inevitably care about the decisions that the central authority can take, and so tend to spend too much time trying to influence the authority's decisions. That time, of course, is valuable; if it were not wasted on influence activities, it could be used for directly productive activities or simply consumed as leisure.

The fact that centralization entails costs does not mean that centralizing decision authority is never desirable. Central planning and decisionmaking may improve coordination among the diverse actors in an economic system sufficiently to make bearing the attendant influence costs worthwhile. In general, when the potential benefits of central control are slight and the influence costs are great, the discretion of the central authority should be restricted. Since influence costs tend to be greater when the members of the organization have larger stakes in the decision to be taken, efficiently designed organizations limit the discretion of decisionmakers in those matters that are of little importance to the organization (in terms of the potential for improved decisionmaking to advance the

organization's objectives) but of great importance to individual organization members.

Although the foregoing themes appear to be general ones, we shall limit our formal analysis of them to the important special case where the organization is a profit-maximizing firm and the interested parties are the firm's employees. This focus forces us to face certain issues squarely. First: Why do employees care about the decisions made by their employers? Under the traditional spot contracting equilibrium theory of the labor market, prevailing wages always leave each employee just indifferent between his current (best) job and his next best job alternative. According to that theory, jobs that are unpleasant or dangerous pay higher wages than those with more desirable characteristics. In practice, employers do pay some compensating differentials: Premium wages have often been paid for hazardous duty, overseas assignments, and late night shifts. Why aren't these practices even more extensive, fully compensating all employees for all variations in job characteristics? Are the uncompensated job characteristics found in practice just an unimportant residual? These questions are of central importance for our theory, for if wages did fully compensate employees for all variations in job characteristics, then employees would have no interest in influencing employer decisions.

We have no formal evidence to offer concerning the magnitudes of the failure of compensating differential theories, though it is clear from casual observation that actual wages are normally adjusted only for substantial and long-lasting changes in job attributes. The costs of writing detailed contracts may be part of the reason for this incompleteness of compensating differentials.

In Section 2, we offer some alternative explanations. We first examine an optimal contracting model in which the wage paid can depend on all the attributes of a worker's assignment; the assignment itself is assumed to be determined only after the worker is hired. We further assume that there are some restrictions on worker mobility, such as relocation or training costs, that free the employer from the absolute need to compensate employees fully for every variation in their work environment. Still, under the terms of an optimal contract, risk neutral employers always insure risk averse employees against income fluctuations, and one might guess that employers would also insure employees against other sources of fluctuations in their welfare. Such a guess would be far off the mark. For example, with a Cobb-Douglas specification of preferences over working conditions and wages, an optimal contract will specify that higher wages be paid to employees enjoying better working conditions! More generally, when employees care about both working conditions and consumption and

consumption is a normal good, employees will prefer assignments with good working conditions, because under an optimal contract poorer working conditions are not fully compensated by higher wages. The magnitudes of these effects depend on employee risk aversion: As risk aversion increases, the optimal wage schedule is transformed toward one with fully compensating differentials.

Two additional contracting models are also analyzed in Section 2. In these models, unlike the one just discussed, job characteristics matter to employees only to the extent that they affect income. In each model, we compute the optimal contract and then study the income streams attached to different assignments. In the first, we find that employees prefer assignments that build their human capital, because these raise future wages with no offsetting current wage reduction. In the second, we find that employees prefer "critical" jobs -- defined as those for which quits are especially costly to the employer, because critical jobs pay higher wages in order to reduce turnover. In all three models, employees care about events that occur after the date of hiring. And, in all three, an employee's ranking of these events bears no necessary relation to the ranking based on employer net profits.

Influence activities and the optimal limitation of executive and management discretion is the subject of Section 3. There, we study a model in which employees

allocate their time between influence activities and some directly productive activity. In the model, the firm can use its wage policy to alleviate influence costs, but it will sometimes prefer to restrict the decisionmaker's discretion instead. There are two key parameters in the model which we vary to study when the discretion permitted decisionmakers should be restricted. The first parameter measures the importance of the decision being modeled; it is essentially the excess of the expected payoff from making an informed decision over that from holding unconditionally to the status quo. The second measures the utility that would be transferred from one employee to another if a change from the status quo were authorized. In an efficiently designed organization, management will be allowed no discretion over those decisions that are relatively unimportant to the organization but that have large potential redistributive consequences.

As an illustration of efficient design, consider American Airlines' procedure for assigning flight attendants to routes. Once a month, flight attendants bid for the routes they prefer, with conflicts resolved on the basis of seniority.¹ Management exercises no discretion over the assignment decision. This, of course, is perfectly appropriate: The airline cares little about which attendants are

¹For international flights, some positions are reserved for suitably multilingual flight attendants, and only those with certified fluency are permitted to bid for those positions.

assigned to which routes, but the flight attendants care a great deal. American Airlines's practice, like many standard operating procedures, can be understood as an attempt to avoid the influence activities that would result if management exercised discretion in assigning flight attendants to routes.

Rosenberg and Birdzell [1986] have emphasized the importance for Western economic growth over the past several centuries of the "immunity [of innovators] from interference by the formidable social forces opposed to change, growth, and innovation" (p. 24). In terms of our theory, the social costs of an incorrect decision to allow, say, a new steelmaking process or a new sailing ship design were small compared to the potential redistributive consequences of such an innovation. Thus, it was wise or lucky that Western governments were unwilling or unable to establish an agency to review and approve innovations. By contrast, in China, which was more centralized and more scientifically advanced than the West in the Middle Ages (when an important period of Western economic growth began), the mandarins exercised a control over the means of production that allowed entrenched interests to slow the pace of innovation.

A brief review of some related theoretical literature is given in Section 4; applications are suggested in Section 5; and concluding remarks are offered in Section 6.

2. Why Full Compensating Differentials Are Not Paid

Following Coase [1937] and Simon [1951], let us suppose that at the time of contracting neither the employer nor the employee knows precisely what conditions will prevail at the time that work must actually be performed. In an academic job market, a new professor may not know who his colleagues will be, which courses he will teach, what his committee and administrative responsibilities will be, which office and secretary will be assigned to him, who his research assistant will be, etc. These characteristics of the job, to be determined after the employment relation begins, will be denoted by x . The employment contract specifies a wage that may be a function of the undetermined characteristic: $w = w(x)$.

To build a simple formal model of this situation, we assume that the possible circumstances $\{x_1, \dots, x_N\}$ and their probabilities $\{p_1, \dots, p_N\}$ are given exogenously. Let w_i denote the wage paid in circumstances x_i . Suppose that the employee's preferences are given by the von Neumann-Morgenstern utility function $u(x, w)$. For brevity, let us write $u_i(w)$ for $u(x_i, w)$. We assume each u_i is twice continuously differentiable with $u_i' > 0$. The employer is a risk neutral expected net profit maximizer; it receives revenues of π_i in event x_i . Suppose that, at the time of contracting, labor market conditions require the employer to offer the agent an expected utility of at least \bar{u} . Further

suppose that the employee, after signing the contract and learning that the job is x , will quit and reenter the labor market unless $u(x, w(x))$ is at least some reservation level $\hat{u} < \bar{u}$, where $\bar{u} - \hat{u}$ reflects mobility costs. The employer, however, is assumed always to be bound by the contract. An efficient contract, subject to the employee's "no quitting" constraint, solves:

$$\begin{aligned}
 \text{(CP)} \quad & \text{Maximize } \sum_{i=1}^N p_i (\pi_i - w_i) \\
 & \text{subject to} \\
 & \sum_i p_i u_i(w_i) \geq \bar{u} \\
 & u_j(w_j) \geq \hat{u} \quad \text{for all } j=1, \dots, N.
 \end{aligned}$$

We consider a family of problems like (CP), parameterized by \bar{u} . Take \hat{u} to be any function of \bar{u} such that $\hat{u}(\bar{u})$ is always less than \bar{u} . When does the optimal contract pay full compensating differentials, leaving the employee indifferent among post-hiring events?

Theorem 1. A solution to (CP) exists and makes the employee indifferent among outcomes and just willing to work ($u_i(w_i^*) \equiv \bar{u}$) for every \bar{u} in the range of u_1 if and only if u_1 is concave and for all i there exists g_i such that:

$$(1) \quad u_i(w) \equiv u_1(w + g_i) \quad \text{for all } w.$$

Proof. It is routine to check that (1) and the concavity of u_1 imply that the optimal contract exists and satisfies $u_i(w_i^*) \equiv \bar{u}$; we focus attention on the reverse implication. Writing $w_i^* = w_i^*(\bar{u})$, the hypothesis is

$u_i(w_i^*(\bar{u})) \equiv \bar{u}$ for all i and all \bar{u} in the range of u_1 , that is, $w_i^* = u_i^{-1}$. The first-order necessary conditions for optimality in (CP) imply that, for all i and all \bar{u} in the range of u_1 :

$$(2) \quad u_i'(w_i^*(\bar{u})) = u_1'(w_1^*(\bar{u})) .$$

Then, $w_i^*(\bar{u}) = w_1^*(\bar{u})$. Integrating this identity, there exists g_i such that $w_i^*(\bar{u}) \equiv w_1^*(\bar{u}) + g_i$ for all \bar{u} in the range of u_1 . Then, for any fixed w , we have $u_1(w+g_i) = u_1[w_1^*(u_i(w))+g_i] = u_1[w_1^*(u_i(w))] = u_i(w)$. This holds for all w , as required.

Given the identity just derived, the second order necessary conditions imply that $u_1''(w_1^*(\bar{u})) \leq 0$ for all \bar{u} , which establishes concavity. \square

An optimal contract equates a risk averse employee's marginal utility of income in the different events x ; it does not also equate his utilities in the different events unless the employee has ordinal preferences that can be represented by vertically parallel indifference curves in (x,w) -space. This characterization of ordinal preferences is quite restrictive. When it fails, the optimal contract will not leave the employee indifferent among assignments.

Now we make an obvious but quite important observation: At the optimal contract, all the income risk is borne by the employer, that is, the employee's wage w_i does not depend at all on π_i or on any π_j ($j \neq i$). Consequently, there is no necessary relationship between the employer's ranking of

outcomes and the employee's. Later, when we introduce the possibility that the employee can influence x , this possible divergence of rankings will become quite important.

Theorem 1 is just a starting point. It tells us that full compensating differentials are rarely paid in a large class of contracting models. The remainder of this section is devoted to the development of examples to illustrate the following points: (1) there is not even a general tendency for optimal wage schedules to compensate for job characteristics, so that employee job concerns under optimal contracts may be quite pronounced; (2) increases in risk aversion tend to lead to the payment of fuller compensating differentials; (3) employees may care about job attributes under optimal contracts even when, contrary to the simple model just presented, job attributes are not an argument of employee utility functions; and (4) these models lead to plausible predictions about the kinds of preferences among job characteristics that employees may systematically show.

Example 1: Preference for good working conditions.

Let $x \geq 0$ denote either working conditions or on-the-job consumption² and let $w \geq 0$ denote the wage or at-home consumption. Suppose that the employee's ordinal preferences have the Cobb-Douglas form $x^\alpha w$ and that his coefficient of relative risk aversion for wage gambles is the constant

²Stafford and Cohen [1974] supplied one of the earliest economic treatments of on-the-job consumption in a study of how work effort varies during the workday.

$\beta > 0$, so that the employee is risk averse. These cardinal preferences are represented by $U(x,w) = \alpha \ln(x) + \ln(w)$ in case $\beta = 1$ and by $U(x,w) = (x^\alpha w)^{1-\beta} / (1-\beta)$ in case $\beta \neq 1$. Suppose that the employee's initial reservation utility level is $\bar{u}^{1-\beta} / (1-\beta)$, with $\bar{u} > 0$ and $\hat{u} = -\infty$. Then the solution to the contracting problem (CP) is $w(x) = \lambda x^{\alpha(1-\beta)/\beta}$ for some constant λ that depends on the parameters α , β , \bar{u} , and (p_1, \dots, p_N) . Notice in particular that if $\beta < 1$, then $w(\cdot)$ is actually an increasing function of x : This establishes that there is no general tendency for optimal contracts to pay even partial compensation for unfavorable working conditions.

In this example, the ordinal utility associated with job x can be measured by $x^\alpha w(x) = \lambda x^{\alpha/\beta}$; it increases in x for any level of risk aversion. This last observation is a special case of a general result that has been derived by several authors including Bergstrom [1984], Chari [1983], and Green and Kahn [1983]. Their result, applied to this model, holds that if on-the-job consumption is a normal good, then the optimal wage contract will always lead employees to prefer jobs with higher x .

In Example 1, as the coefficient of relative risk aversion β increases, the ordinal utility measure $x^\alpha w(x) = \lambda x^{\alpha/\beta}$ becomes increasingly flat and converges to the constant \bar{u} . With increases in risk aversion, the optimal contract pays higher wages in bad jobs and lower wages in

good jobs until, in the limit as relative risk aversion tends to infinity, full compensating differentials are paid. The proposition proved below generalizes the example and establishes that, for fixed ordinal preferences represented by a smooth utility functions $U(x,w)$ that is concave in w , if increases in risk aversion cause the wages to rise in some jobs and to fall in others, then the wages rise in "poor jobs" and fall in "good jobs."

Let $U(x,w)$ represent the preferences of the less risk averse employee and $V(U(x,w))$ the preferences of the more risk averse employee. Assume that $U_w > 0$, $U_{ww} < 0$, $V' > 0$, and $V'' < 0$. In our finite state model, we may assume without loss of generality that $V'(u) \rightarrow -\infty$ as $u \rightarrow -\infty$ and $V'(u) \rightarrow 0$ as $u \rightarrow +\infty$. The reservation utility levels for the two problems are \bar{u} and \bar{v} ; we make no assumption about how they are related. Assuming an interior optimum to the two optimal contracting problems, the marginal utilities of income across assignments are equalized for each of the two agents: $U_w(x, w(x)) = \lambda$ and $V'(U(x, \hat{w}(x)))U_w(x, \hat{w}(x)) = \mu$ for all x , where $w(\cdot)$ and $\hat{w}(\cdot)$ are the respective optimal wage schedules.

Theorem 2. There exists u^* such that for all x , $u^* \leq U(x, \hat{w}(x))$ if and only if $w(x) \geq \hat{w}(x)$. That is, as the employee grows more risk averse, the ordinal utility levels associated with each assignment are contracted toward a level u^* by raising wages in assignments with lower utility and reducing wages in assignments with higher utility.

Proof. Choose u^* so that $V'(u^*) = \mu/\lambda$. There are two cases. If $U(x, \hat{w}(x)) \geq u^*$, then since U_w is positive and V' is decreasing,

$$\mu = V'(U(x, \hat{w}(x)))U_w(x, \hat{w}(x)) \leq (\mu/\lambda)U_w(x, \hat{w}(x)).$$

So, $U_w(x, \hat{w}(x)) \geq \lambda = U_w(x, w(x))$, which (since $U_{ww} < 0$) implies that $\hat{w}(x) \leq w(x)$ and hence that $u^* \leq U(x, \hat{w}(x)) \leq U(x, w(x))$. The case for $U(x, \hat{w}(x)) \leq u^*$ is similar. \square

Under the additional assumptions that $\bar{v} = V(\bar{u})$ and that $-V''(\cdot)/V'(\cdot)$ is bounded below by a constant r , it can be shown that $U(x, \hat{w}(x)) \rightarrow \bar{u}$ as $r \rightarrow \infty$; that is, as the lower bound on the coefficient of absolute risk aversion tends to infinity, wages tend to compensate fully for variations in working conditions.

Example 2: Preference to Accumulate Human Capital.³

This example is a variation on Example 1 in which the relevant attribute of the job, contribution to general human capital, is not a direct argument of the worker's utility function.

We suppose that the employee has a two-period life. His productivity in period 1 is p ; in period 2 it is either p again or, if he has incremented his human capital in the first period, it is $q > p$. There is no firm-specific human capital, so that the worker's productivity does not depend on whether he remains with his initial employer. There are

³This analysis of this model was motivated by (and bears a strong resemblance to) that of Harris and Holmstrom [1982].

two possible events in the first period. In the first event, which arises with probability r , the employer will assign the worker to a task that increases his second period productivity to an amount $q > p$. In the second, which occurs with probability $1-r$, the worker is assigned to a task in which human capital is unchanged, so the worker's second period marginal product will be p .

At the beginning of the second period, the worker is free to quit the firm and go to work elsewhere for a wage equal to his current marginal product. This mobility imposes a lower bound on the wage the worker can be paid in the second period. However, there are some market frictions: The employee cannot leave during the first period after learning his job assignment.

Let π_j be an increment to the firm's revenues when event j occurs and w_{ij} the corresponding period i wage. We assume that competition among similar firms drives the expected wage over the two period contract to be equal to the worker's expected marginal product over that period, which is $rq + (2-r)p$. Our model also assumes that the worker can neither borrow nor save (although only the no borrowing constraint is in fact binding), so that his consumption is equal to his income in each period. Competition among employers will lead them to offer an efficient contract -- one that maximizes the worker's

utility subject to the maximum expected wage constraint and the constraints on second period wages:

$$(3) \quad \text{Maximize } r[u(w_{11}) + u(w_{21})] + (1-r)[u(w_{12}) + u(w_{22})]$$

subject to

$$r[w_{11} + w_{21}] + (1-r)[w_{12} + w_{22}] \leq rq + (2-r)p$$

$$w_{21} \geq q$$

$$w_{22} \geq p$$

where u is some strictly concave function.

This is a concave maximization problem with linear constraints, so its optimal solution is fully characterized by a first-order condition. It is not hard to verify that the unique optimal solution has $w_{21} = q$ and $w_{11} = w_{12} = w_{22} = p$ (with Lagrange multipliers of $u'(p)$, $r[u'(p) - u'(q)]$, and zero, respectively, on the three constraints.) Thus, in each period the employee is paid his current marginal product. An employee who is fortunate enough to be assigned to job 1 acquires valuable human capital but suffers no offsetting wage reduction under the terms of the optimal contract. Consequently, employees prefer job assignments that increment their human capital. The employer's net profit under the contract in event j is precisely π_j -- an amount unrelated to the employee's human capital acquisition. So, the employee's interests may conflict with the employer's.

Example 3: Preference for "Critical" Jobs.

Our final example is a simple "efficiency wage" model. According to efficiency wage theories, the productivity of

an employee is an increasing function of his wage, so employers may find it optimal to pay a wage exceeding the market clearing level. Higher wages may increase productivity for a wide range of reasons; for example they may encourage employees to work more diligently or they may attract better applicants or reduce employee turnover. Several of the important papers in the efficiency wage literature are reprinted in a volume edited by Akerlof and Yellen [1986] together with a helpful survey by the editors.

Our purpose here is to note that the same factors that make an employer choose to pay wages in excess of market clearing may also make him choose to pay different wages for different jobs in a way unrelated to employee qualifications, so that employees will care about how those jobs are assigned. In particular, we shall show that wages are positively related to the costs of job turnover, since higher wages reduce costly turnover.

Thus, assume that the gross profits earned when x_i occurs are π_i if the employee works and $\pi_i - \Delta_i$ if he quits. The agent is assumed to be a risk neutral expected wage maximizer: His utility is w if he works in job i at wage w , $g + b$ if he is laid off and receives a layoff bonus b , and $g + b_Q$ if he quits and receives bonus b_Q . The variable g -- the employee's outside opportunities -- is privately observed by the employee after the job is assigned and is drawn from a distribution F with a density function f that

is continuous and positive on the interval $(0, \bar{g})$. There is no bonding of employees and the employer cannot penalize the employee for quitting, that is, $b, b_Q \geq 0$.

To insure an interior optimum for our contracting problem, assume that $\bar{g} > \max_i \Delta_i > \min_i \Delta_i > 0$. To have the optimum characterized by first-order conditions, we also assume strict quasiconcavity of the objective (4) below for all values of Δ ; this amounts to the assumption that $w + F(w)/f(w)$ is increasing in w .

If the employer's only instrument were to set wages w_i to pay in each event and a termination bonus $b = b_Q$ to pay to departing employees, the employee would quit whenever his outside opportunities were at least $w_i - b$. The problem could then be written in the form:

$$(4) \quad \text{Max} \sum p_i [(\pi_i - w_i)F(w_i - b) + (\pi_i - \Delta_i - b)(1 - F(w_i - b))] \\ \text{subject to}$$

$$\sum p_i [w_i F(w_i - b) + \int_{w_i - b}^{\bar{g}} (g + b)f(g)dg] \geq \bar{u}.$$

The wage policy w_i that maximizes (4) takes the form $w_i = w(\Delta_i)$. Since w_i does not depend on π_i , there is no necessary relation between the interests of the employer and employee. Thus, as in the previous models, arbitrarily severe conflicts of motives can arise between the employer and employee. The heuristic optimal wage policy satisfies the rearranged first-order condition

$$(5) \quad \Delta_i = w_i + (1 - \lambda)F(w_i)/f(w_i)$$

where λ , the Lagrange multiplier of the single constraint, is the marginal cost of providing an extra dollar of expected income to the employee. It is clear that λ cannot exceed one. Hence, the right-hand side of (5) is increasing in w_i , so wages increase with Δ_i : Employees prefer to occupy "critical" jobs in which turnover is costly to the employer.

In Appendix A, we give a full formal analysis of this problem without the restriction to simple wage policies used above. In the full model, the employee may report his outside opportunities to his employer, but the truthfulness of any report cannot be assured. The employer can take account of the report in setting wages, quit bonuses, layoffs, and layoff bonuses; it can also randomize among policy options. The upshot is that none of these additional options are used by the employer and that the heuristic analysis given above yields the right answer:

Theorem 3. The employer has an optimal policy that requires no randomization or reporting by the employee. The policy establishes for each assignment a fixed wage w_i to be paid if the employee works and a bonus b_Q to be paid if he quits (which occurs whenever his outside option pays more than $w_i - b_Q$). Under this optimal policy, the wage $w_i = w(\Delta_i)$ is an increasing function of Δ_i .

3. When Does It Pay to Restrict Management Discretion?

We consider a simple model of influence in which the employee allocates his available time T between two activities, a directly productive activity and an influence activity. If the employee spends time t at the directly productive activity, then his output will be "high" with probability $p(t)$ and "low" with probability $1-p(t)$. The organization will earn an extra profit of π if output is "high." We assume that $p'(t)$ is continuous and strictly positive and $0 < p(t) < 1$ on $[0, T]$.

If the employee spends time s at influence activities and the central decisionmaker has discretion to authorize a change from the status quo, then the change will be authorized with probability $q(s)$ and the expected increment to profits from added flexibility in decisionmaking will be $I\gamma(s)$. We assume that $q'(s)$ and $\gamma(s)$ are continuous and strictly positive on $[0, T]$. The positive parameter I measures the "importance" of the decision in terms of its potential to improve profits.

The employee's preferences are specified by a utility function that provides utility of $u(w)$ for a wage w in the status quo and $u(w) + k$ for a wage w when a change in conditions is approved; we take $k > 0$. We assume that u is defined on $[0, \infty)$, that $u' > 0$, and that $u'' < 0$. With these preferences, the employee has no actual aversion to spending time in productive activities. Formally, that distinguishes

this model from moral hazard models such as those studied by Grossman and Hart [1983], Harris and Raviv [1979], Holmstrom [1979], or Holmstrom and Milgrom [1987]. However, this is a moral hazard model because if management has discretion to change the status quo, then there is an opportunity cost to other workers' time: Time spent in production is unavailable for influence activities. This is represented by the constraints $s + t \leq T$ and $s, t \geq 0$.

The wage paid may depend on the decision (change or no change) and on the employee's output performance (high or low). There are four possible decision-performance outcomes. Individual outcomes are denoted by i and their corresponding probabilities and wages are denoted by $p_i(s, t)$ and w_i .

When the executive has discretion, a rational, self-interested employee will seek to:

$$(6) \quad \text{Max}_{s, t} \sum p_i(s, t) u(w_i) + q(s)k$$

subject to $s + t \leq T$
and $s, t \geq 0$.

The social objective is given by:

$$(7) \quad \sum p_i u(w_i) + \lambda [I\gamma(s) + p(t)\pi - \sum p_i w_i],$$

where $\lambda > 0$. This objective is a positively weighted combination of the firm's profits and employee's utility, but it excludes the employee's utility increment k .

Excluding k from the social objective represents our assumption that this employee's gain is a loss to some other employee who is accorded equal weight in the social

calculus. Thus, k denotes the potential magnitude of the distributional effect of the decision.

Fix a time allocation (s, t) and let $V(s, t)$ be the optimal value of the corresponding "Implementation Problem":

$$(8) \quad \begin{array}{l} \text{Maximize } \sum_{\{w_i\}} p_i [u(w_i) - \lambda w_i] + \lambda p(t) \pi \\ \text{subject to } (s, t) \text{ solves (6).} \end{array}$$

In standard fashion, $V(s, t)$ is an upper semicontinuous function on $[0, T] \times [0, T]$.

With this notation, the social problem can be expressed as $\text{Max}_{s, t} V(s, t) + \lambda I \gamma(s)$. The value of this transformed social objective is increasing in I , so the optimal value is increasing in I as well.

When there is no decisionmaker with authority to alter the status quo, the maximal social payoff is

$$\bar{V} = \text{Max}_w u(w) - \lambda w + p(T) \pi .$$

Lemma. $\bar{V} > \text{Max} \{V(s, t) \mid s + t \leq T, s, t \geq 0\}$.

Proof. First, we claim that $\bar{V} > V(0, T)$. Indeed, \bar{V} is the maximal value of the relaxed version of (8) with $s = 0$, $t = T$, and the incentive constraint (that (s, t) maximizes (6)) omitted. The unique optimum of the relaxed problem has $u'(w_i) = \lambda$ for all i . But then $(0, T)$ does not maximize (6), so the optimal value of the constrained problem is less than the optimal value of the relaxed problem: $\bar{V} > V(0, T)$.

Next, we claim that $\bar{V} > V(s, t)$ for all (s, t) with $t < T$. Indeed, the optimal value of the relaxed version of (8) with the incentive constraint omitted is obtained by

setting $u'(w_i) = \lambda$ for all i , which yield the optimal value $\bar{V} + \lambda[p(t) - p(\bar{T})]\pi$. Since $\lambda\pi p' > 0$, this is less than \bar{V} for all $t < T$, as claimed.

Finally, since V is upper semicontinuous, there exists a pair (s^*, t^*) such that

$$V(s^*, t^*) = \text{Max} \{V(s, t) \mid s + t \leq T, s, t \geq 0\}.$$

Whatever that pair is, $V(s^*, t^*) < \bar{V}$. \square

\bar{V} is the optimal value achieved when management has no discretion to authorize a change and $\text{Max}_{s, t} V(s, t) + \lambda I \gamma(s)$ is the optimal value when management does have discretion. In view of the lemma and the boundedness of $\gamma(\cdot)$, it is clear that as I approaches zero it is best to restrict management discretion.

Theorem 4. There exists a pair of parameters (I, k) such that when these parameters prevail, it is better to eliminate discretion than to provide wage incentives to limit influence activities. Moreover, if (I, k) is such a pair and if $I' \leq I$ and $k' \geq k$, then (I', k') is another such pair.

The arguments preceding the Theorem establish all the Theorem's assertions except the assertion about how the optimal policy varies in the parameter value k . Formal proof of that assertion is contained in Appendix B.

A number of assumptions have been incorporated into the model to keep things simple, and one may well wonder: How far can these be relaxed? First, the restriction to two

output levels (high and low) is plainly dispensable; what is important for our argument is only that the moral hazard problem be severe enough that the first-best is unattainable when management discretion is unlimited.

Second, we have assumed that the $q(s)$ and $r(s)$ functions are given exogenously, so that the decision criterion to be used by management is not a choice variable of the problem. If there are several possible decision criteria but these cannot be committed to ex ante (perhaps because it is hard even to describe a standard of evidence that will be required), then once again the decision criterion is not a choice variable and Theorem 4 holds precisely as stated.

Third, we have set up our model with k as a purely distributional parameter — of no import for the social objective. We would have reached a conclusion similar to Theorem 4 if we included k in the social objective in the following way. Let I , formerly a positive parameter, be allowed to take negative values as well. Define a social importance function $\hat{I}(s) = kq(s) + \lambda I r(s)$. The costs of unlimited discretion still depend only on k and the gains only on $\hat{I}(\cdot)$. Then, a result resembling Theorem 4 can be obtained in terms of the social importance function \hat{I} and the real parameter k .

Finally, several of the assumptions made here have been relaxed by Milgrom and Roberts [1987b], who study influence

activities by workers seeking a desirable job assignment. Their model includes the possibilities of competition among workers, promotions as rewards for past performance, and decision rules that are chosen in advance by management. Despite these differences, their conclusions reinforce our general finding that central decisionmakers ought not always be allowed full discretion to make optimal decisions given the facts at hand, since that leads to excessive influence activity.

4. Related Literature.

Others have offered alternative explanations of the diseconomies of central control by executives in firms or by regulators in the public sector. Williamson [1985] and Grossman and Hart [1986], writing in the framework of transaction cost economics,¹ emphasize the hazards that arise from opportunistic behavior by the owner-managers² of integrated firms as a source of diseconomies. More closely related to our theory is the rent-seeking argument developed by Bhagwati [1982], Buchanan [1980], Krueger [1974], Posner [1985], and Tullock [1967], among others. This theory holds that government granted subsidies, tariffs, and monopolies

¹See Milgrom and Roberts [1987a] for a survey of transaction costs economics that integrates influence costs into the theory.

²Both Williamson's argument and Grossman and Hart's apply also to include hazards of opportunism by the central office executives when the owner can effectively motivate them to act in his interests.

impose welfare losses on society because they lead businesses to waste resources in attempts to win tariff protection or monopoly rights for themselves. These analyses lend support to our view that government interventions ought sometimes to be limited in order to discourage wasteful rent-seeking activity.

Our argument extends that of the "rent-seeking" literature in two principal respects. First, we explicitly treat the possibility that influence activities may improve decisionmaking. The parties most affected by a decision are frequently among those best informed about the alternatives and their consequences, or are at least best motivated to discover and analyze thoroughly the alternatives and their likely consequences. For that reason, decisionmakers benefit by having access to the information, opinions, analyses, and concerns of interested parties. Second, influence activities are not the exclusive concern of the public sector. There are tremendous payoffs in the private sector from "salesmanship" -- both the actual commissions earned by salesmen and more generally the payoffs to having one's ideas accepted or projects adopted or performance evaluated favorably. Our analysis replaces the earlier emphasis on the costs of government intervention and the need to reduce public sector influence activities with a new and much broader emphasis on the costs of centralized

authority and the need to design organizations that provide incentives for only appropriate influence activities.

Within firms, influence activities can be controlled by wage policy, by limiting access to the decisionmaking process, or by limiting management discretion. In public sector decisionmaking by regulatory bodies and legislatures -- especially in a society where public access to decisionmakers is regarded as a matter of right -- the corresponding instruments of to control lobbying and influence are weaker. Consequently, influence costs are likely to be higher in the public sector than within firms. This observation may help to explain the exclusive focus of earlier writers on rent-seeking behavior within the public sector.

5. Applications.

Our approach points to possible economic explanations and analyses for phenomena traditionally studied by sociologists as well as to new analyses of some traditional economic problems. Here are just a few examples.

1. Resistance to change. As we have seen, employees in even the best run firms are rarely indifferent about matters that affect their working conditions or job content. Employees can be expected to resist those changes that threaten to leave them less well-off by failing to cooperate in the search for better ways to do business or by subverting changes in the hope of restoring the older order. This rent-seeking theory contrasts with non-economic theories in

the way it identifies the sources of resistance, the kinds of changes that it predicts will be most vigorously opposed, and the strategies that it predicts will be adopted to overcome resistance by successful firms in rapidly changing environments. (See Milgrom and Roberts [1987b] for a more extensive analysis).

2. Vertical integration. When a firm's key suppliers are not perfect competitors (i.e., their prices exceed their marginal costs), they may incur excessive selling costs and impose decision costs on the buyer in their attempts to earn the rents associated with marginal sales. All these costs are influence costs that can be reduced or eliminated by vertical integration (which restricts the buyer's discretion about from whom to purchase). Any gains realized in this way must be balanced against the losses from reduced discretion and the costs of newly centralized authority over other decisions in the integrated organization.

3. Takeover Bids/Golden Parachutes. The evidence concerning takeovers indicates that the stockholders of the acquirer do not earn conspicuous excess returns.³ Thus the economic motive for takeovers may well be the increased rents earned by the management of the acquirer, for example because their increased authority in the merged firm makes their jobs more "critical" in the sense of Example 3 of

³See Jensen and Ruback (1983).

Section 2. Mere transfers of the rents earned by the former management to the shareholders and new management do not enhance efficiency, and that part of takeover activity by the acquiring firm and defensive activity by target firm's management that is simply redistributive is wasteful. Golden parachutes, properly designed, are executive compensation packages that force potential acquirers to reimburse former managers for any lost rents when there is a transfer of control. These discourage inefficient takeovers and reduce both rent-seeking by potential acquirers and rent-protecting behavior by existing management. The consequent efficiency gains ultimately benefit the shareholders.

4. Litigation Policy. A court trial is a centralized decision process in which the disputants often incur enormous influence costs to effect a redistribution of wealth. As "bright-line" law fades and parties become less sure of the likely outcome of litigation, the discretion of juries and judges correspondingly rise.⁴ Damage rules play the role that wages played in our study of influence within firms: Rules limiting damages reduce influence costs

⁴Discretion varies in degree. The greatest discretion is that of a decisionmaker who has authority to pursue whatever objectives he wishes or even to make decisions on impulse. A lesser, but still important, discretion is that of the decisionmaker who is permitted to exercise judgment in some aspect of the decisionmaking process -- for example about what information to consider, how to evaluate information or weigh evidence, etc. Judges who are bound to follow the law or managers who are obliged to serve the interests of owners can still be said to have discretion of this second kind.

possibly at the expense of such other objectives as paying just compensation or creating efficient incentives for contractual performance.

6. Concluding Remarks.

The economic environment we have described differs markedly from the neoclassical, perfectly competitive spot contracting environment in which buyers are indifferent at the margin about what they buy, sellers are indifferent about incremental sales, and workers are indifferent about employer decisions. In our conception, people care about decisions and attempt to influence them. When decisionmakers are honest and rational, influence takes the form of suggesting alternatives and supplying information, opinion, and analysis; when they are not, influence may take more insidious forms. Efficient organization design in this conception seeks to do what the system of prices and property rights does in the neoclassical conception -- to channel the self-interested behavior of individuals away from purely redistributive activities and into socially productive ones. The success that our institutions have in achieving this objective is a major determinant of our economic welfare.

REFERENCES

- Akerlof, George and Janet Yellen, Efficiency Wage Models of the Labor Market, Cambridge, England: Cambridge University Press, 1986.
- Bergstrom, Theodore, "Soldiers of Fortune?" in Equilibrium Analysis: Essays in Honor of Kenneth Arrow, Volume II, Walter Heller, Ross Starr, and David Starret (eds), Cambridge: Cambridge University Press, 1986: 57-80.
- Bhagwati, Jagdish, "Directly Unproductive, Profit-Seeking (DUP) Activities," Journal of Political Economy, 90, 1982: 988-1002.
- Buchanan, James, "Rent Seeking and Profit Seeking," in Toward a Theory of the Rent Seeking Society, James Buchanan, Robert Tollison, and Gordon Tullock (eds), College Station, Texas: Texas A&M University Press, 1980: 3-15.
- Chari, V.V., "Involuntary Unemployment and Implicit Contracts," Quarterly Journal of Economics, XCVIII (1983), 107-22.
- Coase, Ronald, "The Nature of the Firm," Economica, 1937: 386-405.
- Green, Jerry and Charles Kahn, "Wage-Employment Contracts," Quarterly Journal of Economics, XCVIII (1983), 173-87.
- Grossman, Sanford and Oliver Hart, "An Analysis of the Principal-Agent Problem," Econometrica 51 (1983), 7-45.
- , "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," Journal of Political Economy, 1986: 691-719.
- Harris, Milton and Arthur Raviv, "Optimal Incentive Contracts with Imperfect Information," Journal of Economic Theory, 1979: 231-59.
- and Bengt Holmstrom, "A Theory of Wage Dynamics," Review of Economic Studies, 1982: 315-33.
- Holmstrom, Bengt, "Moral Hazard and Observability," Bell Journal of Economics, 1979: 79-91.

- and Paul Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives," Econometrica, 1987.
- Jensen, Michael and Richard Ruback, "The Market for Corporate Control: The Scientific Evidence," Journal of Financial Economics 11, 1983: 5-50.
- Krueger, Anne, "The Political Economy of the Rent-Seeking Society," American Economic Review 64, 1974: 291-303.
- Milgrom, Paul and John Roberts, "Bargaining and Influence Costs and the Organization of Economic Activity," Stanford GSB Working Paper, 1987a.
- , "An Economic Approach to Influence Activities in Organizations," unpublished working paper, 1987b.
- Posner, Richard, "The Social Costs of Monopoly and Regulation," Journal of Political Economy 83, 1975: 807-27.
- Simon, Herbert, "A Formal Theory of the Employment Relationship," Econometrica, 1951: 293-305.
- Stafford, Frank and Malcom Cohen, "A Model of Work Effort and Productive Consumption," Journal of Economic Theory, 1974: 333-47.
- Tullock, Gordon, "The Welfare Costs of Tariffs, Monopolies, and Theft," Western Economic Journal 5, 1967: 224-32.
- Williamson, Oliver, The Economic Institutions of Capitalism, New York: The Free Press, 1985.

APPENDIX A

Proof of Theorem 3.

By standard arguments in the theory of optimal contracts, the firm can (without loss of optimality) represent its design problem as follows. The employee, observing the job x_i and his outside opportunity g , reports that his outside opportunity is \hat{g} . Prior to receiving the employee's report, the firm commits itself to a strategy that specifies for each job i and report \hat{g} a wage $w_i(\hat{g})$ and a probability $q_i(\hat{g})$ that the employee will be laid off and paid a bonus of $b_i(\hat{g})$. Employees who quit from job i are paid $b_{iQ}(\hat{g})$. Without loss of generality, we may restrict attention to contracts for which employees never want to quit.⁵ The functions w , b , and q are "incentive compatible," that is, they always provide the employee with an incentive to report g truthfully. In this form, the firm's problem is:

$$(A1) \quad \text{Max}_{w, b, q} \sum_i p_i \left[\pi_i - \int_0^{\bar{g}} [q_i(g)(A_i + b_i(g)) + (1 - q_i(g))w_i(g)] f(g) dg \right]$$

subject to

⁵ Given any contract in which employees sometimes quit, there is an "equivalent" contract without quits. Under the equivalent contract, the employer lays off the employee in just the circumstances when he would have quit under the originally given contract, and pays a layoff bonus equal to the quitting bonus the employee would otherwise have received. The equivalent contract simply relabels quits as layoffs.

$$(XA-PC) \quad \sum_i p_i \cdot \int_0^{\bar{g}} U^i(g; g) f(g) dg \geq \bar{u}$$

$$(XP-PC) \quad w_i(g) \geq g + b_{iQ}(g) \quad \text{for all } i, g,$$

$$(IC) \quad U^i(g; g) \geq U^i(\hat{g}; g) \quad \text{for all } i, g, \text{ and } \hat{g}, \text{ and}$$

$$w_i(g), b_i(g), b_{iQ}(g) \geq 0 \quad \text{for all } i \text{ and } g,$$

where

$$(A2) \quad U^i(\hat{g}; g) = q_i(\hat{g})(g + b_i(\hat{g})) + (1 - q_i(\hat{g}))w_i(\hat{g})$$

Equation (A2) defines $U^i(\hat{g}; g)$ to be the employee's expected utility when he reports that his outside opportunity is \hat{g} but it is actually g , and the employee does not quit. The firm's objective (A1) is one of profit maximization. It consists of expected gross profits minus the losses suffered and bonuses paid from layoffs minus the wages paid when there is no layoff. The incentive constraints (IC) state that it must be in the employee's interest to tell the truth about g . There are two kinds of participation constraints. The ex ante participation constraint (XA-PC) requires that the employee receive at least the minimum expected utility level; otherwise, he would not agree to the contract. The ex post participation constraints (XP-PC) require that the employee's wage be no worse than his outside opportunity when he is employed; otherwise, he would quit. All the wages and bonuses are constrained to be non-negative: There is no bonding.

Substituting (A2) into (A1), the objective may be more usefully expressed as the total expected revenues from inside (productive) and outside (employment) opportunities minus the employee's share:

$$(A3) \quad \sum_i p_i \left[\pi_i + \int_0^{\bar{g}} \{q_i(g)[g - \Delta_i] - U^i(g; g)\} f(g) dg \right]$$

Theorem 3. (a) There exist numbers w_i^* and b^* such that an optimal solution to the firm's problem has $w_i(g) = w_i^*$ and $b_i(g) = b_{iQ}(g) = b^*$ for all i and g and has $q_i(g)$ equal to zero or one according as $g + b^*$ is less or more than w_i^* . Also, $w_i^* > w_j^*$ if and only if $\Delta_i > \Delta_j$.

(b) Let λ be the Lagrange multiplier on (XA-PC).

Always, $\lambda \leq 1$. If $\lambda < 1$, then $b^* = 0$. If $\lambda = 1$,⁶ then and $w_i^* = \Delta_i + b^*$.

Proof. Given any feasible solution, increasing both $w_{iL}(g)$ and $w_i(g)$ by any constant k increases the employee's ex ante expected utility by kp_i and reduces the employer's profit by an equal amount; hence λ can never exceed unity.

Let subscripts on U^i denote partial derivatives. In view of (IC), we must have $U_1^i(g; g) = 0$ for all g . Then, using (A2), $\frac{d}{dg} U^i(g; g) = U_2^i(g; g) = q_i(g)$. So,

⁶The case with $\lambda = 1$ is relatively uninteresting, since it has wages in excess of the employee's marginal product Δ_i for all i . In such a case, one would not expect the firm to hire the employee at all.

$$(A4) \quad U^i(g; g) = U^i(\bar{g}; \bar{g}) - \int_g^{\bar{g}} q_i(s) ds$$

Notice that the constraint (XP-PC), the non-negativity constraint on $b_{iQ}(g)$, and (A2) imply that:

$$(A5) \quad U^i(\bar{g}; \bar{g}) \geq \bar{g}.$$

Now, we consider a "relaxed" problem: We maximize the objective (A3) subject only to the constraints (XA-PC), (A4) and (A5). To accomplish that, let $K = \sum p_i \pi_i$, let λ be the multiplier on (XA-PC), and form the Lagrangian:

$$(A6) \quad K + \sum_i p_i \left[\int_0^{\bar{g}} q_i(g) [g - \Delta_i] f(g) dg + (\lambda - 1) \int_0^{\bar{g}} U^i(g; g) f(g) dg \right]$$

which is to be maximized by choice of $q_i(\cdot)$, $w_i(\cdot)$, $b_i(\cdot)$, and $b_{iQ}(\cdot)$. Substituting from (A4) and reversing the order of integration for the resulting double integral, the Lagrangian becomes:

$$(A7) \quad K + \sum_i p_i \left[(\lambda - 1) U^i(\bar{g}; \bar{g}) + \int_0^{\bar{g}} q_i(g) [g - \Delta_i + (1 - \lambda) F(g)/f(g)] f(g) dg \right]$$

Let $H_i(g) = g - \Delta_i + (1 - \lambda) F(g)/f(g)$. Since $\lambda \leq 1$, the hypothesis that $g + F(g)/f(g)$ is increasing implies that H_i is increasing. Then, given our hypothesis about the support of F , the equation $H_i(w) = 0$ has a unique solution \bar{w}_i lying in $(0, \bar{g})$. Let $q_i(g) = 0$ for $g \leq \bar{w}_i$ and $q_i(g) = 1$ for $g > \bar{w}_i$.

Consider the case $\lambda < 1$. Letting $w_i(g) = w_i^* = \bar{w}_i$ and $b_i(g) = b_{iQ}(g) = 0$ for all i and g , the constraints (IC) and

(XP-PC) are satisfied by inspection. By proper choice of λ , (XA-PC) is satisfied. These choices lead to $U^i(\bar{g}; \bar{g}) = \bar{g}$ for all i , so together with the specified $q_i(\cdot)$ functions they maximize (A6) by inspection. Since this optimal solution of the relaxed problem satisfies all the constraints of the original problem, it is an optimal solution of the original problem.

Finally, consider the case $\lambda = 1$. Set $w_i(g) = w_i^* = \Delta_i + k$ and $b_i(g) = k$, where:

$$(A8) \quad k = \bar{u} - \sum_i p_i \int_0^{\bar{g}} \max(\Delta_i, g) f(g) dg .$$

By inspection, the choice of q specified in the Theorem is the unconstrained expected surplus maximizing choice, and all the constraints are satisfied.

In each case, w_i^* increases strictly with \bar{w}_i which in turn increases strictly with Δ_i . This establishes the last statement of (a) in the Theorem. □

APPENDIX B

Proof of Part of Theorem 4

What must still be shown is that if discretion is optimally permitted for the parameter pair (I, k) and if $\hat{k} < k$, then discretion is optimally permitted for the pair (I, \hat{k}) . For this, it suffices to show that for all (s, t) and k , $V(s, t | \hat{k}) \geq V(s, t | k)$, where the notation now notes explicitly the dependence of the optimal value of (8) on the parameter k .

Suppose $\{w_i\}$ solves (8) for parameter value k and let $\bar{u} = \sum p_i u(w_i)$. For $\hat{k} < k$, define \hat{w}_i by $u(\hat{w}_i) = (1 - \hat{k}/k)\bar{u} + (\hat{k}/k)u(w_i)$. Let $U(s, t | k, w) = \sum p_i(s, t)u(w_i) + q(s)k$ and define $U(s, t | \hat{k}, \hat{w})$ similarly. Then for all (s, t) ,

$$(B1) \quad U(s, t | \hat{k}, \hat{w}) = (1 - \hat{k}/k)\bar{u} + (\hat{k}/k)U(s, t | k, w)$$

so that if (s, t) maximizes $U(s, t | k, w)$ then it also maximizes $U(s, t | \hat{k}, \hat{w})$.

Since u^{-1} is convex, we have by Jensen's inequality:

$$(B2) \quad u^{-1}(\bar{u}) = u^{-1}\left[\sum p_i u(w_i)\right] \leq \sum p_i w_i.$$

Applying Jensen's inequality and substituting from (B2),

$$(B3) \quad \begin{aligned} \sum p_i \hat{w}_i &= \sum p_i u^{-1}[(1 - \hat{k}/k)\bar{u} + (\hat{k}/k)u(w_i)] \\ &\leq \sum p_i [(1 - \hat{k}/k)u^{-1}(\bar{u}) + (\hat{k}/k)w_i] \\ &\leq \sum p_i w_i. \end{aligned}$$

It follows from (B1), (B3) and the definition of $V(\cdot)$ that $V(s, t | \hat{k}) \geq V(s, t | k)$, as we required. \square