

Strategic Behavior and the Impact of Unions on Wage Incentive Plans

W. KIP VISCUSI

*Duke University, Durham, NC 27706 and
National Bureau of Economic Research*

This paper analyzes optimal wage incentive plans in situations in which firms make capital investments and employees supply work effort. An efficient contract will take workers' response into account in its design, leading to a competitive contract that takes the form of a lump sum penalty and full marginal rewards to workers, or equivalently, a base wage and a piece rate above a quota. A union optimizing worker welfare will alter this lump sum amount rather than attempt to control worker effort or the marginal rewards for additional output. The greater utilization of piece rate systems in capital-intensive industries is correctly predicted by the analysis.

I. Introduction

Although wage incentive systems have attracted considerable attention in the labor economics literature, it has only been recently that their conceptual motivation has been formalized. Two approaches can be identified.¹ The first research approach, which extends the classic work of Alfred Marshall, pertains to sharecropping schemes. Landowners either rent land to workers who reap all of the value of the output or enter into sharecropping agreements in which the output is divided. Although there are two factors of production in these models, the form of the wage payments considered has been restricted primarily to a division of the output in some constant proportions.

The second approach is reflected in the more recent literature on monitoring worker effort. Since the worker's input to the firm cannot be monitored perfectly, the employer must structure a wage incentive scheme based on a

¹The first area of research is characterized by the work of Lucas (1979) and the references contained therein. The monitoring literature includes wage system analyses by Pencavel (1977) and Stiglitz (1975), the agency model of Ross (1973), and the optimal income transfer analysis of Zeckhauser (1971) whose work anticipates the optimal contract structure used in the monitoring literature and in this paper.

monitorable variable correlated with work effort — his output. The approach taken has been to treat the employer as, in effect, simply an insurer. The employer takes no actions other than designing a contract that is sufficiently attractive to acquire a work force. The time and piece rate provisions of the contract are manipulated to provide appropriate incentives for workers and to promote effective risk-spreading. These contract provisions parallel the use of coinsurance and deductible rates in standard insurance contexts in which there are adverse incentives (or moral hazard) problems.

Insurance aspects of wage incentive plans are clearly important. However, contract provisions of this type also have a more fundamental role to play within the context of the employment relationship. In this paper, I will focus primarily on the efficiency aspects of wage incentive schemes. Unlike modifications of principal-agent models, the employer will not be relegated to the role of an insurer. The firm will take substantive actions through its capital stock investments.

The presence of productivity-enhancing actions for both workers and firms introduces strategic complications not present in earlier analyses and alters the form of the optimal wage structure. If there is some economic rent produced by the employment relationship, the division of this surplus will depend on the relative bargaining power of the two parties. The most important determinant of the workers' power in the United States industrial relations system is representation by a union, which will be a major focus of this analysis.

The principal characteristics of worker behavior and market outcomes are analyzed in Sections II and III. Whereas in most market contexts it is optimal for each party to take the other's action as given when making their own decisions (i.e., a Nash equilibrium), the intrinsic externality involved in the employment relationship makes it optimal for the employer to be strategic in designing the contract. In particular, an efficient contract will take workers' responses into account and will take the form of a wage that consists of a lump sum penalty and full worker sharing in the value of output. Although contracts with this structure have been utilized in the monitoring literature, they have not appeared in analyses of two-factor employment models. The motivation for the optimal contract structure in this model is the presence of capital expenditures by the firm rather than use of deductible and coinsurance provisions in a wage insurance contract.

Sections IV and V address the role of unions and the relationship of the optimal wage structures to those observed in American industries. It has long been observed that unions' bargaining efforts focus on the wage guarantee rather than the piece rate and that utilization of piece rate systems is negatively related to the capital intensity of the industry. Despite the prominence of these and related relationships in empirical analyses of piecework systems, no conceptual basis has been provided for these effects. A principal purpose of this paper is to provide a framework for analyzing these phenomena.

II. *Worker Behavior*

The analysis below will consider individuals' supply of work effort in situations in which employers can monitor an individual's output with certainty, but cannot monitor the effort provided. In practice, other signals of work effort may also be available, and an individual's output may be monitored only imperfectly. I will neglect these problems to focus on more fundamental aspects of the employment relationship.

Consider a hypothetical worker with a utility function U on net wealth, where $U' > 0$ and $U'' < 0$. The worker's job choice involves a choice of a wage-work effort mix, where, for simplicity, effort is in terms of a monetary equivalent so that the net reward from the job is his wage w minus his work effort choice x . Let the production function be given by $f(x, y)$, where y is the enterprise's input (capital), with f having the usual concave shape. For simplicity, x and y are both scaled in units so that their price per unit equals one. Unlike earlier incentives models, both firms and workers make choices affecting f .

If the firm adopts a linear incentive system, workers receive a fixed wage α , which is possibly negative, plus some fraction β of output so that

$$w = \alpha + \beta f(x, y).$$

The efficiency of a linear incentive system will be analyzed in the following section. If nonlinear incentive systems are employed, one must also consider the problem of monitoring the output of each individual worker. Wages that are positively linked to the square of output, for example, would provide an incentive for workers to exchange their product, reap the highest net wage available, and then divide their gains. A nonlinear rewards schedule consequently would impose additional monitoring costs since the firm would have to monitor each individual's output to prevent trading of goods designed to exploit likely differences in marginal rewards for output.

Consider the situation in which the worker regards the reward system and the enterprise's choice of y as predetermined, as one would expect in a competitive situation involving a large number of individuals. The worker selects x to maximize $U(\alpha + \beta f(x, y) - x)$, leading to the first-order condition for work effort of

$$f_x = 1/\beta. \quad (1)$$

Upon total differentiation of equation 1, one obtains the result that effort increases with the piece rate factor β , or

$$\partial x / \partial \beta = -(f_x / f_{xx}) > 0. \quad (2)$$

The effect of additional input by the firm is ambiguous, since

$$\partial x / \partial y = -(f_{xy} / f_{xx}), \quad (3)$$

which is positive if x and y are complements ($f_{xy} > 0$, as one would expect) and negative if they are substitutes ($f_{xy} < 0$).

III. Market Outcomes

A. *The Nash Equilibrium.* Suppose the worker and firm take the reward structure and other's action as given. The worker's behavior is characterized by equation (1). The firm's profits π per worker consist of the total per capita output f minus both its per worker input y , which is scaled in monetary units, and the wage rate $\alpha + \beta f$, or

$$\pi = (1 - \beta)f(x, y) - y - \alpha.$$

Non-strategic selection of the input y by the firm will ignore its effect on x , leading to the requirement

$$f_y = 1/(1 - \beta). \quad (4)$$

Equations 1 and 4 characterize the Nash equilibrium actions.

B. *Joint Wealth Maximization.* The total net product W available to the worker and firm consists of the output minus each of their inputs, or

$$W = f(x, y) - x - y.$$

When x and y are selected to maximize their joint well-being, one has

$$f_x = f_y = 1. \quad (5)$$

Equation (5) can never be satisfied by non-strategic workers and firms in the Nash equilibrium situation since equations (1) and (4) are inconsistent with the requirement above. More specifically, there is no value of β such that

$$f_x = 1/\beta = f_y = 1/(1 - \beta) = 1.$$

The distinctive aspect of the employment relation is that there is an intrinsic externality involved in the production process. Work effort influences the productivity of the employer's capital stock, generating a positive effect in the case of complementary inputs and a negative effect in the case of substitutes. Similarly, the capital stock provided influences workers' productivity. Failure to account for this interdependence will not generally be optimal. This finding contrasts with the generally efficient aspects of non-strategic actions in other market contexts.

C. *The Competitive Firm.* A competitive firm typically will not ignore workers behavior. With a large number of workers, the firm's optimization problem will be to select the wage structure and input y to maximize its profits subject to two constraints. First, the reaction functions of workers given by equations (2) and (3) describe the effect of manipulating β and y on the worker's effort. Second, the marginal worker will be paid sufficiently so that his net wealth equals his reservation wage rate for the wage structure and the x and y values chosen. Consequently,

$$w_o = w - x = \alpha + \beta f(x, y) - x.$$

This requirement can be viewed as restricting the value of α to

$$\alpha = w_o - \beta f(x, y) + x, \quad (6)$$

where the worker does not take into account the effect of his actions on the wage structure.

Enterprise profits given by

$$\pi = (1 - \beta)f(x, y) - y - \alpha$$

can be rewritten after substituting for α as

$$\pi = f(x, y) - x - y - w_o. \quad (7)$$

The firm picks β and y to maximize π .

Differentiating equation 7 with respect to β , one obtains

$$\pi_\beta = 0 = f_x(\partial x / \partial \beta) - \partial x / \partial \beta,$$

or

$$f_x = 1. \quad (8)$$

From equation (1), we know that f_x equals $1/\beta$. Consequently, β equals one, or the marginal reward for the worker is linked to the *total product*, not his marginal product. Rewards based on marginal productivity do not provide efficient incentives since they neglect the externality to the firm.

The value of y is selected so that

$$\pi_y = 0 = f_x(\partial x / \partial y) + f_y - \partial x / \partial y - 1$$

or

$$0 = (f_x - 1)(\partial x / \partial y) + f_y - 1$$

Since f_x equals one, this requirement is that

$$f_y = 1. \quad (9)$$

Equations (8) and (9) jointly determine the optimal actions x and y , while β equals one and a is determined by equation (6) for these (x, y) values. These conditions are identical to that for the maximization of the joint wealth of the parties as described by equation (5).

Strategic enterprise behavior that optimizes with respect to worker reaction functions in the choice of the wage structure and y will be socially optimal, whereas the non-strategic behavior led to sub-optimal inputs by both workers and firms in the case where these inputs are complements. It should also be emphasized that the linear wage incentive scheme generated efficient outcomes in this situation. Nonlinear wage structures offer no efficiency gains.

IV. *Union Influence and Incentives*

The view of workers being nonstrategic is plausible in situations in which there is no formal worker organization, but if a union is present there may be a

substantial effort to influence the wage incentive scheme. Unions have generally opposed straight piecework systems and, in practice, most incentive schemes now observed provide a base guaranteed pay level coupled with the marginal reward structure.² Though often quite vocal in their opposition to piece rate schemes, the extent of union influence has been rather limited.³ The level of the guaranteed wage before meeting the quota at which piece rates begin is a primary concern, as is the level of the quota. As shown in the following section, these actions are akin to manipulations of α . The actual choice of the marginal wage incentive β is typically a prerogative of management.

The reason for this emphasis will be illuminated by the subsequent analysis of union action. In particular, unless the union can dictate both the incentive scheme and worker actions irrespective of the enterprise's choice of y , the union can potentially reap greater gains by focusing on the base pay and the production standard rather than altering marginal incentive rates.

Although the balance of power between firms and unions spans a continuum of possibilities, the extreme situations are readily amenable to analysis. The case of dominance by the firm has already been considered. Suppose instead that the union assumes the strategic position. Consider two possible scenarios. First, suppose the union could determine both the worker effort level x and the wage structure, taking into account the firm's reaction in its input choice.

Enterprises that select y with the wage structure and x predetermined by the union will be affected by the method for dividing output in much the same way as are workers confronted with wage incentive systems. As before, let the workers receive a lump sum amount α plus some fraction β of the output. Then enterprise profits will be given by

$$\pi = (1 - \beta)f(x, y) - y - \alpha.$$

The firm will set y such that

$$f_y = 1/(1 - \beta). \quad (10)$$

The effect of the predetermined variables on the choice of y are similar to the effect of the environment on worker actions since

$$\partial y / \partial \beta = [f_y / (1 - \beta) f_{yy}] < 0 \quad (11)$$

and

$$\partial y / \partial x = -f_{yx} / f_{yy} \quad (12)$$

which takes on the same sign as f_{yx} .

²See Rees (1979) for additional discussion along these lines.

³Dunlop (1958), especially on p. 258, discusses the shortcomings of union efforts.

The net wealth of the hypothetical worker is $\beta f - x - \alpha$. If α is set at the lowest level acceptable to the firm, the enterprise will be able to make zero profits, or

$$\alpha = (1 - \beta)f - y. \quad (13)$$

The worker's utility function consequently is $U(f - x - y)$. The optimal values for β and x must satisfy

$$0 = [f_y - 1] (\partial y / \partial \beta) \quad (14)$$

and

$$0 = f_x - 1 + (\partial y / \partial x)[f_y - 1]. \quad (15)$$

The first condition reduces to $f_y = 1$, and, upon substituting for this value of f_y , equation (15) reduces to $f_x = 1$. As in the cases of the competitive firm and joint wealth maximization, both x and y are set at levels where the marginal products equal 1.

The principal difference is that this condition implies that β must equal zero, from equation (10). Whereas previously the worker's marginal incentive was proportional to the total product, it is now the marginal reward of the company that increases with the output level. The modification of the wage structure is made so that the final economic actor making an input decision will behave in an efficient manner.

Although the total product, x , and y are the same whether the union or the firm is the dominant force, the relative power of the two parties is not irrelevant.

The net rent to the employment relationship is $f - x - y$. Unless workers are paid their reservation wage rates and firms earn zero profits, this amount will be positive. The ability to control the wage structure as well as one's productive input generates the capability for appropriating this surplus.

While complete dominance by either party assumes efficient outcomes, intermediate strategic possibilities may not simply involve a division of the surplus but may also affect its size. Suppose, for example, that the union selects the wage structure on the first move, the firm picks y on move 2, and the worker picks x on move 3. Following a procedure similar to that employed above, it can be shown that if each party anticipates subsequent reactions to its input choice, then no wage structure selected by the union will provide efficient incentives for both the firm and workers. In particular, the worker in move 3 picks the value of x satisfying

$$f_x = 1/\beta,$$

while the firm selects y such that

$$f_y = 1/(1 - \beta) - f_x(\partial x / \partial y) = 1/(1 - \beta) - (1/\beta)(\partial x / \partial y).$$

There is no value of β for which the efficient inputs will be generated, that is, where

$$f_x = f_y = 1.$$

There is an inefficiency involved in union manipulation of the wage structure when unions cannot pre-determine both worker actions and the wage incentive scheme. Except in the case of unusually demanding incentive systems, worker inputs are typically not dictated by central command but instead are selected in response to the work environment. If unions focus on simply raising the lump sum wage payment α by, in effect, boosting the workers' reservation wage w_0 , they will not affect the total net value of the employment relationship. The share of the rent accruing to management will be diminished as α increases. The union will be able to appropriate more of this rent for workers without diminishing the size of the pie to be divided.

The efficiency losses associated with union manipulation of the incentive scheme may explain in part why unions have not emphasized wage incentive choices in their collective bargaining efforts. Certainly the major practical concern in situations with incentive pay is the production level at which piece rate remuneration will begin. Mixed systems with an initial quota require a modification of the stylized incentives systems considered thus far.

V. *Implications for Wage System Design*

The optimal wage incentive plan will link worker incentives proportionally to the total product with a lump sum wage amount α that is necessarily negative if the firm incurs any real costs (i.e., if $y > 0$).

While this format is attractive on a conceptual level, one does not observe workers making lump sum payments to their employers. In the usual piece rate system, workers must attain some minimum quota f before piece rate remuneration begins. Below that amount, they are given a wage w . Above that level, the wage is $\bar{w} + [f - \bar{f}]$ in the case where the marginal incentive is proportional to output.

The incentive scheme described above accords with the practices advocated in the personnel administration literature. A good wage incentive system will be associated with a guaranteed rate for production up to a standard, coupled with marginal rewards proportional to output or "100 percent sharing."⁴

This incentive scheme offers the same work incentives as does the (α, β) linear incentives scheme for values of output not below f . Worker wages will increase apace with additional output — a result that characterizes generally employed incentives schemes. Below the quota, workers are offered a flat wage payment. The absence of effective incentives in that region may be a major

⁴See p. 393 of Pigors and Myers (1973).

explanation for why companies usually set relatively low quotas. If workers consistently fail to meet the quota, they can be transferred or fired. Moreover, the firm can employ other monitoring and motivational devices, such as offering promotions to productive workers and utilizing an assembly line to set a pace for work. These mechanisms will not be considered here.

A piece rate system with a minimum quota will impose the same costs as the (α, β) system, where both offer marginal incentives proportional to output, if

$$\alpha + f = \bar{w} + \max[0, f - \bar{f}].$$

or

$$\bar{w} = \alpha + f - \max[0, f - \bar{f}].$$

If f is sufficiently low that it will always be met by workers,

$$\bar{w} = \alpha + \bar{f}.$$

The value of α was given by equation (6), where unions can manipulate w_0 to raise worker rents, so that

$$\bar{w} = w_0 + x - (f - \bar{f}).$$

Under the base pay system with production levels in excess of the quota, enterprise profits will be given by

$$\pi = \bar{f} - y - \bar{w}.$$

Consider all possible (\bar{w}, \bar{f}) combinations yielding a particular level of π . More specifically, let us analyze the nature of competitive equilibrium in which profits equal zero. Then \bar{f} and \bar{w} are related by the condition

$$\bar{f} = y + \bar{w}.$$

Higher rates of base pay \bar{w} must be associated with a higher quota \bar{f} . The role of y is of particular interest. As the cost of the enterprise's input increases, the gap between the quota and the base pay must widen.

This finding illuminates a principal determinant of piece rate incidence. More specifically, the industries for which piece rate utilization is high tend to be labor cost intensive, as the theoretical analysis would suggest.⁵ Industries such as the shoe and apparel industry have traditionally emphasized piece rates, in large part because such incentive schemes are effective when output can be monitored and capital costs are low.

Historical trends in the use of piece rate systems also may be explained in part by this relationship. Dunlop (1958) observed that in the mining industry increased mechanization generated a decreased reliance on piecework. Several

⁵See, in particular, McKersie, Miller, and Quaterman (1964), and Slichter, Healy, and Livernash (1960). The relative stability of the incidence of wage payment schemes by industry is documented by Seiler (1979) and Mangum (1962).

studies suggest that this effect is not due to the mode of technology but simply the level of capital costs.⁶

The theoretical analysis suggests the source of this influence. As the enterprise's input value increases, the quota must be successively raised, thus diminishing the range of output over which piece rates can be both effective and economically viable.

VI. *Conclusion*

Linear incentive schemes, such as a guaranteed wage coupled with a constant piece rate above a production quota, can provide effective incentives if either the enterprise or union can pre-determine both the wage structure and its own input choice. Because of the intrinsic externalities involved in the employment relationship, an efficient incentive scheme requires that the party designing the wage structure take into account its effect on the incentives that are generated. Moreover, the optimal wage structure should be linked to the worker's total product rather than his marginal product or else the correct incentives will not be provided.

Although similarly efficient outcomes can result if unions dictate the wage structure, they must also be able to control worker effort. If individual incentives are responsive to any actions taken by the firm, the union could potentially reap greater gains by focusing not on the marginal incentive structure but on the level of base pay and the production quota. This emphasis is reflected in actual union actions.

As the enterprise's capital costs rise, the production quota must be raised or the base pay lowered in order for the wage structure to be viable. The importance of incorporating an analysis of the firm's inputs into a model of wage incentive schemes is borne out by the most firmly established empirical determinant of piece rate incidence — the negative relationship between the share of capital costs and the industrial incidence of piece rate systems.

⁶See the discussion in McKersie, Miller, and Quarterman (1964) and Slichter, Healy, and Livernash (1960).

REFERENCES

- Bush, Joseph, "Incentive Pay Patterns in the Steel Industry," *Monthly Labor Review*, (Aug. 1974), pp. 75-77.
- Cox, John, "Time and Incentive Pay Practices in Urban Areas," *Monthly Labor Review*, (Dec. 1971), pp. 53-56.
- Dunlop, John, *Industrial Relations Systems* (Carbondale: Southern Illinois University Press, 1958).
- Lucas, Robert E. B., "Sharing, Monitoring, and Incentives: Marshallian Misallocation Reassessed," *Journal of Political Economy*, Vol. 87, No. 3 (1979), pp. 501-521.
- Mangum, Garth, "Are Wage Incentives Becoming Obsolete," *Industrial Relations*, Vol. 2, No. 1 (1962), pp. 73-96.
- Marshall, Alfred, *Principles of Economics* (London: Macmillan, 1949).
- McKersie, Robert, Carroll Miller, and William Quarterman, "Some Indicators of Incentive Plan Prevalence," *Monthly Labor Review*, (March 1964), pp. 271-276.
- Pencavel, John, "Work Effort, on-the-Job Screening, and Alternative Methods of Reimbursement," in R. Ehrenberg, ed., *Research in Labor Economics*, Vol. 1 (Greenwich: JAI Press, 1977), pp. 225-258.
- Pigors, Paul, and Charles Myers, *Personnel Administration* (New York: McGraw-Hill, 1973).
- Rees, Albert, *The Economics of Work and Pay* (New York: Harper and Row, 1979).
- Ross, Stephen, "The Economic Theory of Agency: The Principal's Problem," *American Economic Review*, Vol. LXIII, No. 2 (1973), pp. 134-139.
- Seiler, Eric, "Piece Rates vs. Time Rates: The Effect of Incentives on Earnings," Harvard University (1979).
- Slichter, Sumner, James Healy, and Robert Livernash, *The Impact of Collective Bargaining on Management* (Washington: Brookings Institution, 1960).
- Stelluto, George, "Report on Incentive Pay in Manufacturing Industries," *Monthly Labor Review*, (July 1969), pp. 49-53.
- Stiglitz, Joseph, "Incentives, Risk, and Information: Notes Towards a Theory of Hierarchy," *Bell Journal of Economics*, Vol. 6, No. 2 (1975), pp. 552-579.
- Zeckhauser, Richard, "Optimal Mechanisms for Income Transfer," *American Economic Review*, Vol. LXI, No. 3 (1971), pp. 324-334.