

UNIV
SHELF

(WORKING PAPER SERIES - 127)

EVALUATION OF PERFORMANCE
AND PAY,

by

Daniel J.B. Mitchell* //

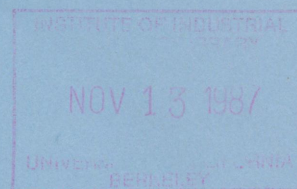
*Daniel J.B. Mitchell
Director
Institute of Industrial Relations
U.C.L.A.
Los Angeles, California 90024-1496
213/825-4339

and

Professor
Graduate School of Management
U.C.L.A.

DRAFT: July 1987

(2)



INSTITUTE OF INDUSTRIAL RELATIONS (Los Angeles)

UNIVERSITY OF CALIFORNIA,

LOS ANGELES

(1)

5

CHAPTER 3:

Evaluation of Performance

and

Pay

Draft chapter of a text on "Human Resource Management: An Economic Approach."

© Daniel J.B. Mitchell 1987

Chapter 3: Evaluation of Performance and Pay

I. Performance Appraisals.

- i. Links in the Chain.
- ii. Supervisors as Performance Raters.
- iii. Agents, Principals, and Performance Appraisal.
- iv. Performance Appraisal and the Rating of Raters.
- v. Employee Influence on Ratings.
- vi. Why Should a Well-Managed Firm Have Perverse Incentives?

II. Reducing the Perverse Incentives of Performance Appraisal.

- i. Documentation.
- ii. Rankings.
- iii. Reviewing the Review.
- iv. Alternative Raters.

III. Rewards and Performance Appraisal.

- i. Alternative Progression Systems.
- ii. Merit vs. Seniority: An Implicit Contracts Approach.
- iii. Rules and Employer Reputation.
- iv. Employer Stability and Corporate Restructuring.

IV. The Cost of Merit and Promotion Systems.

- i. A Numerical Example.
- ii. Gross vs. Net Confusion.
- iii. Merit Budgets and Misincentives.
- iv. Mixing of Merit and Other Pay Adjustments.

V. Conclusions.

Chapter 3: Evaluation of Performance and Pay

In the previous chapter, we noted that labor is not a homogenous factor of production. Even within narrowly defined occupations, labor can vary in productivity and value to the employer for two reasons. First, employees may have different "endowments" of skills, talents, innate traits, and learned behavioral characteristics. Second, employees may choose to vary their behavior in response to conditions at the workplace including incentives and disincentives that may be built into the reward system.

It is evident that the variability in employee quality-- which may not be detectable or predictable at the time of hiring -- will require some type of policy response from the employer. One possibility is simply to gear wages to productivity directly so that, for example, a worker who is 10% more productive than another will enjoy a 10% premium relative to his/her fellow employee. But such payment systems are only possible in the case of well-defined and measurable output.

Alternative pay systems which are geared to productivity will be discussed in a later chapter. But for many types of jobs, piece rates or other pay plans tied to measured production are not feasible. At this point let it simply be noted that historically, the long term trend has been away from such pay

formulas and toward time-based wages, e.g., hourly pay rates or weekly, monthly, or yearly salaries.

It was also indicated in the last chapter that the team element in production often would render output-linked pay differentials impractical in cases of large behavioral differences between employees. A poor worker, who spreads negative externalities to his or her fellow employees, may so adversely affect productivity that no positive wage, no matter how low, could make it worthwhile for the employer to continue to employment relationship.

I. Performance Appraisals.

Given these considerations, some mechanism in the workplace must be established to evaluate performance. When conducted on a formal basis, such systems are often described by HRM professionals as "performance appraisals." Once performance is evaluated, various employer responses are available. In cases of favorable reviews, rates of pay can be increased through a "merit plan." In addition, highly rated employees can be promoted or rewarded with bonuses. Where negative reviews occur, employees may be denied a merit pay increase, warned to correct their deficiencies, disciplined, or dismissed.

Table 1 presents the results of a survey taken in the early 1980s of employer practice with regard to performance appraisal systems. Ninety-two percent of the respondents reported that their firms had at least one performance appraisal system in place. As can be seen from the table, the use of formal performance appraisal is very common -- indeed, close to universal -- except in two cases: appraisal of top management and of unionized production workers. (In the former case, there is a question of who should do the appraisal; in the latter, unions may resist the subjective element in performance appraisals and may push for rewards based on non-subjective seniority).

Since performance appraisals will have desirable or undesirable consequences for workers, the performance appraisal system must be seen as part of the incentive arrangements at the workplace. If good reviews lead to rewards, employees will strive to receive such reviews. Similarly, they will attempt to avoid unfavorable reviews, if these lead to penalties.

Ideally, the performance appraisal system will lead to higher employee productivity, since that is supposed to be the behavioral response which is rewarded. But as with any rating system, problems of design and implementation inevitably arise. Any professor who has designed and administered an exam, or any student who has taken one, will have no difficulty in understanding these problems!

Table 1

**Use of Formal Performance Appraisal Systems
by Type of Employee**

Occupational group	Percent of Employers With Formal Plans	
Top management	67%	
Middle management	87	
First-level supervisors	91	
Professional/technical	88	
Office/clerical	88	
Production workers	63	
Union		42
Nonunion		84

Note: Table shows results of a survey of private and public employers based on the responses of 264 HRM executives.

Source: Bureau of National Affairs, Inc., Performance Appraisal Programs, PPF survey no. 135 (Washington: BNA, 1983), p. 4.

i. Links in the Chain.

Note first that in a performance appraisal system, the reward that the employee hopes to achieve is linked only indirectly to his/her productivity. In theory, high productivity leads to a superior rating which, in turn, leads to a reward. But there are two loose links in the chain between productivity performance and the reward. The productivity performance must in fact be recognized, and reflected in the employee's official rating, before a reward is possible. Even then, someone must examine the rating and decide to link it to the reward.

If either link in the chain is broken (productivity --> rating, or rating --> reward or penalty), the performance appraisal system will not be an incentive toward higher productivity nor will it be a deterrent to improper conduct. In the real world, there are reasons why the chain might be broken at one or both points. Indeed, there are incentives in the workplace which can damage or prevent the proper functioning of a performance appraisal system.

ii. Supervisors as Performance Raters.

Most employees have supervisors. And in most organizations, supervisors form part of a larger, hierarchical authority

structure. Supervisors have supervisors, who -- in turn -- have supervisors. One role of a supervisor is simply to provide instructions, i.e., to tell people what tasks to perform. However, the role of a supervisor is far more complex than simply being an order giver.

Trying to define the supervisory role is difficult, since the role varies from employer to employer. However, a legal definition does exist which illustrates the nature and scope of the role. In 1947, Congress amended basic U.S. labor law by passing the Taft-Hartley Act in order to remove existing protections for supervisors who wished to engage in collective bargaining.¹ At that time, the management community was afraid that unionized supervisors would not adequately represent their employers' interests and pressed for legislation. To make this legal modification, Congress had to define a supervisor. Since that time, a supervisor has been defined as:

"...any individual having authority, in the interest of the employer, to hire, transfer, suspend, lay off, recall, promote, discharge, assign, reward, or discipline other employees, or responsibly to direct them, or to adjust their grievances, or effectively recommend such action..."

Supervisors may perform some or all of the above-listed tasks. But regardless of the scope of their duties, virtually all supervisory employees are capable of taking actions which can enhance or harm the welfare of their subordinates. Determining

the productivity of subordinates and playing some role in linking that determination to tangible rewards is an important element of supervision. Thus, if there is a potentially defective link in the connection between performance and rating, its roots are likely to be found in the incentives and disincentives facing supervisors.

iii. Agents, Principals, and Performance Appraisal.

If information were perfect and costless, supervisors would have little role to play in the firm. But because information is not perfect and costless, the firm must delegate to an individual (the supervisor) responsibility for local operations. Decentralization of authority (to some degree) is unavoidable. In this regard, the internal workings of the firm mirror the decentralized, external market.

Just as ordinary workers cannot be perfectly monitored, neither can supervisors. Supervisors can sometimes abuse their positions, i.e., take actions which benefit themselves at the expense of their employer. Since they have authority to take actions which can benefit or harm subordinates, supervisors could demand favors from subordinates. Such supervisory behavior, of course, is detrimental to the employer's interest (but may not be detected by the employer in a world of imperfect information).

It is a classic illustration of the difficulty principals have in controlling their agents.

Cases of overt monetary "kick-backs" from employees to their supervisors are not common, but are certainly not unknown, either. More common is the vague impression around many workplaces that employees who do favors for supervisors may receive rewards. This impression about the agent (supervisor) may or may not work in the interest of the principal (employer). A rule that to get ahead, you should "please your boss" is fine for the employer if what pleases the supervisor-boss is congruent with advancing the employer's welfare. But congruence of interests between employer and supervisor is not always perfect and certainly is not guaranteed.

In the early part of this century, for example, production workers often viewed their foremen as repressive, arbitrary, and exploitative figures.² This resentment, in fact, contributed to the growth of HRM as employers experienced high costs of turnover from dissatisfied workers during the "tight" labor markets of World War I.³ In addition, employers feared (and sometimes still do) that such resentments would provide footholds for union organizers among their workforces. More recently, the rash of litigation involving sexual harassment claims -- often from women workers claiming that male supervisors demanded sexual favors in exchange for good ratings or career advancements -- dramatically

illustrates that supervisors do not always act in their employer's best interests.⁴

Obviously, a performance appraisal system will be sabotaged if supervisors use their power to give good or bad ratings to extract personal "rents" from workers. However, such blatant supervisory misconduct need not be present for a performance appraisal system to fail in its mission of providing incentives for high employee productivity. Other, more subtle (mis)incentives can have that effect and are far more pervasive.

iv. Performance Appraisal and the Rating of Raters.

Supervisors are judged, in part, by the quality of their subordinates. Since supervisory workers are supposed to motivate subordinates and to correct or eliminate subordinate mistakes and misconduct, a supervisor must be concerned about the effects of reporting an "abnormally" high number of problem employees. The consequences of such a report could be adverse to the supervisor's own interests.

Even if a supervisor simply drew an unlucky hand in the workforce he or she must supervise, accurately reporting that fact poses a certain risk. In a world of imperfect information -- and imperfect appraisal of supervisors -- rules of thumb such as "where there's smoke, there's fire" could undermine a

supervisory career. A supervisor may feel it is best not to advertise workplace problems. Indeed, the worse the problems, the greater the incentive there may be to hide them. "Cover ups" are not just a practice of high government officials!

For the same reason, a department or production unit whose employees are highly rated is a positive reflection on that supervisor. Absent perfect information, those higher in the management structure may believe that a department or work unit with a highly-rated staff, must have a superior supervisor. There are, in short, incentives for supervisors to overstate the positive qualities of subordinates and to downplay subordinate deficiencies in any formal documentation.

Supervisors may feel that they can handle difficult or incompetent workers on an informal basis, bypassing the official performance appraisal mechanism (and the scrutiny of superiors). Those supervisors who elect the unofficial approach may find themselves embarked on a perilous course if the informal route fails, as we will see in a later chapter. But the temptation is there.

Finally, as will be discussed later, performance appraisal systems are often linked to pay; good ratings lead to merit increases for the employee. Generally, in organizational hierarchies, supervisors are expected (and themselves expect) to

be paid more than subordinates. In the absence of perfect information on marginal productivity, it is assumed -- not unreasonably -- that supervisors contribute more incremental value to the employer than those who they direct. Thus, a supervisor who succeeds in raising the average pay of subordinates may succeed in raising his/her own pay level.

v. Employee Influence on Ratings.

The tendency of higher management to consider a problem-infected workforce to be the consequence of an ineffective supervisor creates yet another misincentive. There is a tendency on the part of supervisors to boost ratings of subordinate performance to misleading levels. Employees may know that worker complaints could undermine management confidence in supervisors. Or, even if they do not recognize this possibility, their supervisor will surely be aware of it. Complaints stemming from dissatisfaction of poorly rated employees can produce a climate a supervisor would prefer to avoid. One way of averting such problems is not to give poor ratings, even when merited, except in extreme cases. In addition, to avoid inter-employee jealousies and tensions, supervisors may give relatively undifferentiated ratings.

Most medium to large sized firms have some kind of grievance mechanism, whereby workers can file complaints if they feel

mistreated by their supervisors. In unionized firms, these systems are usually highly formalized and typically provide for arbitration by a neutral, outside arbitrator if the grievance cannot be resolved internally. Larger nonunion firms, especially those which emphasize the HRM function, are likely to have formal grievance mechanisms, too (although only a few nonunion firms provide for an outside arbitrator as a final step of the process).⁹

Under a grievance system, the filing of a complaint by an employee inevitably will come to the attention of management. The supervisor involved will be questioned about the validity of the grievance and about his or her actions which precipitated the complaint will be scrutinized. If many grievances arise from within a particular supervisor's jurisdiction, questions about the quality of supervision may be raised.

Higher management may begin to wonder why the supervisor solve problems before they arise. Even if the supervisor is not the cause of the initial complaints, perhaps he/she is poor at resolving conflict. Perhaps there is an adverse impact on productivity and morale caused by the workplace friction which the grievances are indicating.

Apart from monitoring formal grievances, larger firms may also conduct employee attitude studies to uncover areas of

employee dissatisfaction. These studies sometimes involve periodic polling of workers to determine the nature and source of their workplace concerns. Supervisors in firms which conduct attitude surveys are aware that such reviews might reflect dissatisfaction if many unfavorable performance appraisals are given. Unfavorably-rated subordinates are unlikely to have kind words to say about their supervisors. Again, an incentive exists for supervisors to give too-high, and too-uniform, ratings.

vi. Why Should a Well-Managed Firm Have Perverse Incentives?

That perverse incentives can create perverse behavior on the part of supervisors can hardly be a controversial proposition. But from the economic perspective, why does a profit-maximizing firm permit such misincentives to exist (or even create them in the first place)? Surely, by correcting the improper incentives (and obtaining accurate performance appraisals) firms could enhance productivity and reduce unit labor costs. Are not top managers aware of these potential profit-enhancing cost savings?

The answer is "yes," top managers (including HRM managers) are aware of the difficulties inherent in performance appraisal. They even have evolved (partial) remedies for the problem (discussed in the next section). But the ultimate answer to the question of why perverse incentives are allowed to exist and persist is one of trade offs. The difficulties associated with

performance appraisal are part of a general class of problems inherent in organizations. In general, these problems intensify as organizations grow larger. Larger organizations require that more and more delegations of authority must be made due to imperfect and costly information channels.

Balanced against the costs associated with organizational size and control are gains in coordination of operations which size brings. The firm effectively encloses a set of functions within its organizational structure and takes them out of the external market place. As was noted in the introductory chapter, in a perfect market of the type characterized in elementary economics textbooks, there would really be no firms. Rather, through a daily, costless, auction-like process, workers and owners of capital would organize themselves into temporary production units, based on prevailing prices and costs.

In the real world, however, forming such units on a daily basis would be prohibitively expensive. Firms (organizations) evolve as a result of these costs as the most efficient units of production. As organizations, the creation of firms entails costs associated with imperfect incentives, bureaucracy, etc. But firms that succeed show themselves to be more efficient than the outside market alternative. Within successful firms, HRM professionals attempt to minimize organizational costs (including

misincentives), but cannot entirely eliminate them. Trade offs must be made between competing objectives.⁴

As a simple example, top management might decide that by reducing the number of middle and first-line supervisors, and by replacing them with a smaller number of "trusted" management agents, more accurate performance appraisals would result. Perhaps such benefits would accrue, but even if they did, the firm would simultaneously lose the economies previously gained by delegation and decentralization. It is management's role to balance the two objectives: information accuracy and reduced misincentives vs. decentralized efficiency.

Or, as another example, management could decide to stop monitoring signs of employee dissatisfaction because such monitoring provides a perverse incentive to supervisors to avoid giving poor ratings to low-productivity workers. (Some firms have abandoned formal performance appraisals for just such reasons). But a decision of this type deprives the employer of a source of information that sometimes does indicate poor supervision. Alternatively, the employer could retain the monitoring mechanism, but expend more resources to distinguish between poor supervisors and poor employees. Unfortunately, resources are scarce and expenditures must be limited. Real world systems of performance appraisal will always be imperfect.

In any case, employers may want to retain performance appraisal systems -- even if they provide inaccurate information -- because of the signaling effect they provide. To not have a system, however imperfect, might signal to employees that the employer does not place much weight on quality performance. Having a system -- even if all involved understand its deficiencies -- at least communicates that performance is important to the employer.

II. Reducing the Perverse Incentives of Performance Appraisal.

In the previous section, two basic types of rating error were outlined, both related largely to incentives surrounding supervisors. First, supervisors might take advantage of their position of authority and discretion and award ratings in exchange for "favors" from employees. Second, they might find it advantageous to be overgenerous and undifferentiated in giving subordinates ratings in order to further personal career objectives. Various techniques have been tried to reduce these perverse behaviors which may otherwise undermine a performance appraisal system.

i. Documentation.

The more documentation that is required to back up ratings and the more that ratings are tied to tangible, i.e., verifiable,

criteria, the less leeway there is for false or inflated reports. For these reasons, performance appraisal forms are often accompanied with detailed instructions to supervisors, defining the various rating scales as precisely as possible, and -- in some cases -- providing examples of behaviors which should receive high or low ratings ("behaviorally anchored rating scales"). In some cases, supervisors may be required to cite specific instances ("critical incidents") of either superior or inferior performance on the part of the employee being rated.

Some rating systems rely partly or completely on essay-type responses, rather than on numerical scales. Proponents argue that use of the essay format will capture components of employee behavior that simple rating scales can miss. This argument is made especially for higher managerial and professional jobs, in which "check the box" answers are not informative. In addition, proponents insist that writing an essay requires more care than simply filling out a numerical form.

The reader will immediately see a trade off involved in such methods. Elaborate instructions, detailed requirements for documentation, and lengthy essays all consume substantial time. Moreover, the time involved is typically that of higher-paid employees who perform the rating function. And, of course, care and accuracy are not necessarily proportional to the time consumed. But again, signaling is involved. If the performance

appraisal system requires elaborate documentation, supervisors are given a sense that accurate appraisal is considered important by their superiors.

ii. Rankings.

Since there are incentives to rate most employees as above average, performance appraisals systems can be structured to provide constraints on such tendencies. Supervisors can be asked, for example, to rank employees rather than give them absolute ratings. Or limits can be placed on the proportion who can be top ranked. The result is similar to "grading on a curve" in the educational setting.

However, the pitfalls are the same as in systems of grading on a curve; superior (inferior) employees who are located in departments where there happen to be high concentrations of high (low) productivity workers will tend to be lower (higher) rated than in the average department. Thus, rankings and constraints create problems of "horizontal equity" across departments. They create incentives, moreover, for high productivity employees to seek to transfer out of departments where there are other good performers.

iii. Reviewing the Review.

Rather than rely on the design of the appraisal form to minimize misleading ratings, some firms prefer to subject the ratings to further review. Filled-out performance appraisals can be scrutinized by professionals employed in the HRM department. These reviewers may question results which appear out of line with past reports on the employee in question. Supervisors may also be questioned if their average ratings seem high relative to other departments. (Can everyone really be better than average?)

Such HRM reviews amount to a monitoring of the monitors by individuals who do not have an immediate, personal stake in the supervisor's relations with his/her employees. But there is an obvious expenditure involved in hiring monitors to monitor other monitors. And HRM professionals themselves are subject to potential misincentives. They do not want to appear to be constantly criticizing, and interfering with, the work of line managers. The goal of the organization is production, after all, not perfect appraisal.

Another commonly used option -- which does not involve hiring professional monitors -- is to have the rated employee read the completed form and add any comments he or she believes relevant. Indeed, in some systems, the employee is asked to help designate the relevant criteria to be rated in advance

("management by objectives"). Arrangements whereby the employee reviews the rating will tend to prevent false negative information from becoming part of the record without challenge. On the other hand, such steps are less likely to correct false positive information. And, since some systems require the supervisor to discuss the form personally with the employee, they may add to the incentive to inflate ratings and avoid distasteful confrontations. Even if a supervisor is skilled at giving constructive criticism, the subordinate may not be skilled at receiving it!

iv. Alternative Raters.

Although the vast majority of performance appraisal systems rely on supervisors to function as raters, other options for employers are available. A few companies ask employees to rate themselves, and provide a detailed list of questions for the employee to answer. On the surface, it might seem that employees would have strong incentives to overrate their performance. However, the incentives are more complex.

An employee who rates himself/herself uniformly high on all dimensions will be immediately suspect. No one is perfect. Hence, to appear honest, the employee has an incentive to identify some fault. Having done so, he/she is then under pressure to correct the self-identified problem. The exercise

bears a resemblance to confession in the religious setting or the "self criticism" practiced in communist countries.

Still another alternative is to have co-workers do the rating rather than supervisors. Such systems of "peer review" are extremely rare, but they have traditionally been used for faculty in institutions of higher learning. Where workers are part of teams, shirkers impose costs on others in the team. Thus, fellow team members may not take kindly to inferior performance. Co-workers are often less "understanding" of substandard behavior than supervisors. However, the incentives are again complex; cliques of workers can take advantage of their authority just as supervisors sometimes do. And fear of retaliation can undermine worker/rater systems.

III. Rewards and Performance Appraisal.

It has already been noted that performance appraisal systems will have little impact on behavior if they produce no tangible consequences (positive or negative) for employees. However, it is a common -- but not a universal -- practice, to separate the appraisal process from the reward system. Employers may insist that there is a process for performance appraisal and another for rewards. The reason for this (surprising) separation is linked to fear of exacerbating the misincentives discussed above.

As indicated, supervisors have an incentive to manage highly paid subordinates, since having high paid subordinates may tend to boost their own pay. In addition, supervisors will prefer that their subordinates be content with their pay level. Discontented workers could make achievement of the supervisor's production targets difficult, either through turnover, or through low productivity.

Thus, a supervisor might be tempted to pay a premium (out of his/her employer's pocket) to ensure workplace tranquility. This premium might be higher than the employer would find optimal. It is often felt, therefore, that separating the appraisal process from the reward decision will keep the latter more "honest."

Still, as Table 2 shows, simply keeping reference to rewards off the performance appraisal form does not necessarily sever the connection between ratings and rewards. Most firms which use appraisal systems report that they do use them as a guide for individual pay adjustments and for promotions. Although promotion opportunities may be limited and infrequent, many companies have merit pay systems -- effectively linked to performance appraisals -- which provide the opportunity for regular pay advances. An interesting question is how important performance appraisals are as a source of internal pay advancement within the firm.

Table 2
Uses of Performance Appraisals

Occupational group	Percent of employers using performance appraisal to:	
	Determine wage/salary levels	Make promotion decisions
Top management	80%	87%
Middle management	87	80
First-level supervisors	87	80
Professional/technical	87	79
Office/clerical	86	79
Production workers	78	75

Note: See note, Table 1 for survey details.

Source: Same as Table 1, p. 12.

i. Alternative Progression Systems.

There are various systems of "wage progression" in use. Under such systems, a rate range -- rather than a single wage rate -- is established for an occupation. Typically, employees enter the occupation at (or towards) the bottom of the range, and then have an opportunity to work their way to the top. Although the systems vary in detail, there are two basic options for determining the rate of advancement of the individual employee: time (seniority) and/or merit.

A mechanism of advancement by virtue of time or seniority is formula-driven. It is easy to verify whether an employee should receive a pay increase by applying the simple rules of the system; all that is needed is an accurate record of the date of entry into the job. In contrast, a system of pay advances on the basis of merit requires a subjective judgment (performance appraisal) by a management representative.

Table 3 shows that a merit-based approach is used by employers for most occupational groups. However, the plant/service category is substantially less likely to be under a merit system than the other groups, and is more likely than the others to be rewarded on a seniority basis. Further, Table 3 shows that much of the discrepancy regarding the plant/service occupations revolves around union status. Union-represented

workers are less likely than others to have a merit plan and are more likely to use seniority as a criterion for wage progression.

Unions and collective bargaining will be discussed in a later chapter. However, it is worth noting at this point that the use of seniority is stronger in the union sector than in the nonunion sector for two basic reasons. First, the internal union political mechanism tends to be dominated by more senior worker/members who naturally prefer to tilt workplace rules and benefits in their own favor. Second, unions prefer to limit management discretion on matters of pay.

The evidence of numerous empirical studies suggests that unions have the effect of boosting average pay of the workers they represent above levels employers would otherwise choose. This pay advantage for union workers could be eroded if management had the unilateral discretion to influence pay through subjective merit decisions. Average pay could be reduced by management simply through a finding that few employees were meritorious. To avert employer decisions that could lower pay back to "market" levels, unions will either push for objective seniority-based system or -- alternatively -- insist that whatever merit systems the employer operates be tightly controlled to avoid back-door pay reductions.

ii. Merit vs. Seniority: An Implicit Contracts Approach.

Although Table 3 reveals a strong employer preference for the idea of using merit rather than seniority as a guide to pay advancement, there can be a discrepancy between stated preference and actual result. For example, in the public sector, civil service procedures for pay progression are often nominally based on merit. Yet it is frequently the case that so-called merit decisions are routinely made after a designated time on the job has been served by the employee, and that almost all employees are found to be meritorious on a regular basis. Thus, a supposedly merit-based system can easily operate as a seniority system in practice.

Private employers are more likely to insist that their merit plans do, in fact, function on the basis of merit. However, some empirical studies based on internal company data suggest that seniority is often a critical factor, even in the nonunion sector. That is, time on the job shows up as an important variable in determining the pace of pay advancement for individual employees.

The economic theory of implicit contracting provides some insight into the seniority phenomenon. Under implicit contracting theory, the employee's pay profile and marginal revenue product profile are viewed as separated. That is,

Table 3
Use of Wage Progression Plans, Late 1980

Occupation Group	Percent of Respondents Reporting Use of a Progression Plan:	
	Automatic, Length-of-Service Increases	Merit Pay Plan
Plant/service	29%	44%
Union	39	16
Nonunion	22	60
Office/clerical	15	81
Professional/technical	10	85
Sales	4	83
First-level supervisors	8	88
Middle managers	6	88

Source: Bureau of National Affairs, Inc., Wage & Salary Administration, PPF survey no. 131 (Washington: BNA, 1981), pp. 10, 13.

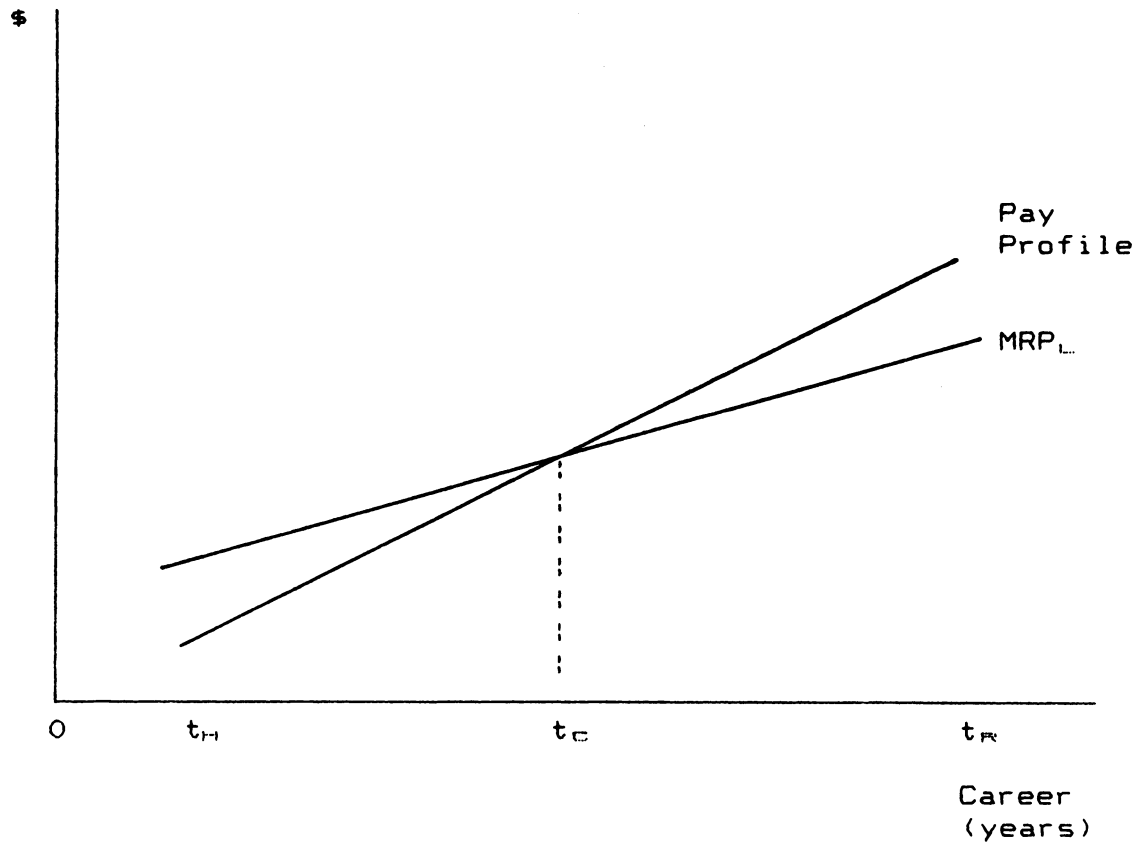
implicit contracting departs from the simple economic theory that $\text{wage} = \text{marginal revenue product of labor (MRP}_L\text{)}$ at every moment of time. Instead, under implicit contracting, the employee is paid less than marginal revenue product in the initial phase of his/her career with the employer and more than marginal revenue product towards the end. Thus, the equality of wage and MRP_L is maintained on average (with appropriate adjustment for discounting) over expected career life, but not instantaneously.

Figure 1 provides an illustration of typical pay and MRP_L profiles over a career. At time t_H , the employee enters employment and is paid less than his/her incremental value to the firm (although at a high enough level for the firm to be able to recruit and retain workers). The firm effectively makes an implicit contract with the employee that if his/her performance is satisfactory, pay will gradually be raised so that eventually -- by the time of retirement t_R -- pay will exceed the level "justified" by productivity. In short, the theory of implicit contracts puts a kind face on the so-called "Peter Principle" (which states that "in a hierarchy every employee tends to rise to his level of incompetence").

Various explanations have been forward for such implicit contracts. On the supply side, the pay profile (low at first; high at career's end) is said to match employee "needs" for income, i.e., young workers have few dependents while older ones

Figure 1

Hypothetical Profile of Pay and Marginal Revenue
Productivity over Career Life



support families and must make provision for retirement. This argument must be accompanied by a presumed constraint on the ability of employees to save when they are young for later expenses.⁷

On the demand side, employers will experience lower turnover costs since employees -- once hired -- must remain with the firm to gain the eventual rewards of seniority. Young employees will also have an incentive to perform at least at a satisfactory level to avoid being terminated before collecting the reward. And older employees will have a strong incentive to maintain their performance level, since such workers will know that -- if terminated -- they would have to find work in the outside market where their productivity-based wages would be significantly lower. For this implicit contract system to work, however, there are two requirements: an "honestly" run wage progression (and possibly promotion) system and an employer who remains in business long enough to "honor" the implicit contract.

iii. Rules and Employer Reputation.

At time t_c on Figure 1, the employee reaches a crossover point in his/her career. After that time, the employee is "overpaid"; before that time, the employee is "underpaid" relative to productivity. Employers would therefore have the temptation to entice job seekers into underpaid, entry-level

service with promises of eventual pay progression. But when the employees actually reached time t_0 , they would be terminated, preventing worker recoupment of the promised return for loyal service.

Given this problem of moral hazard on the part of the employer, potential employees might be reticent to accept an implicit contract of the type represented on Figure 1 without some assurance that the contract would not be broken. Having formal progression plans -- and establishing a reputation for not abusing their discretion -- is one way that employers can provide such assurance. If an employer had a merit system in which pay advancements were commonly denied, potential employees might be reluctant to accept employment, upon learning of the firm's poor reputation as a keeper of commitments. And morale problems might develop among current employees who feared that the deals they thought they had made with the firm were being, or were likely to be, dishonored.

The theory of implicit contracting, therefore, suggests that seniority is likely to play an important role in determining pay advancement, whether or not a merit system is used. It also suggests that when layoffs are made, firms may prefer to let junior workers go before seniors, since terminating seniors before retirement age would tend to violate implicit agreements.

Or seniors might be offered special retirement incentives to compensate them for early separation from the firm.

Finally, implicit contracting suggests that firms -- if unconstrained by legal barriers -- would want to establish some kind of mandatory retirement age (say, at t_R). Such an age would avoid having employees "over-recoup" their initial investments made as junior workers. However, for most employees, Congress forbade mandatory retirement after 1986. Hence, firms might instead have an incentive to offer financial retirement inducements.⁶ Or, they might have an incentive to offer less of an upward sloping pay profile initially. A later chapter will return to some of these points.⁷

iv. Employer Stability and Corporate Restructuring.

Implicit contracting theory was partly inspired by case studies of actual HRM practices. Observations included the "paternal" employer policies such as "lifetime" employment at major Japanese companies, "full" employment policies at certain "progressive" American employers such as IBM, the general social ethos which condemns employers who fire long-service employees, and studies which suggest that older workers have a hard time starting over when they are terminated (due to the existence of career ladders at other potential employers). It is clear, however, that regardless of initial intent, career oriented

commitments of any character are unlikely to be met if the employer ceases to exist before an employee's working life has ended.

Employers may cease to exist because market conditions no longer can maintain them profitably in business. Even before actual shutdown occurs, such employers may effectively cease to defer to seniority since maintaining a reputation as a good employer no longer matters. Companies may also cease to exist as independent entities if they are swallowed by, or merged with, other firms. The new owner/operators may not feel constrained by the implicit commitments of their predecessors.

During the 1980s, there was a considerable acceleration of merger activity and corporate restructuring. This instability in the business sector was due to changes in antitrust policy, increased international competition, deregulation in some industries, and the severity of the recession in the early years of the decade. Such volatility must inevitably have an effect on the credence employees can put in corporate HRM policies (including those related to merit systems, progression, and promotion opportunities).

Rapid change in corporate ownership may have the effect of prodding management to greater efficiency; this is often seen as the social function of the "market for corporate control."¹⁰

Presumably, takeovers, mergers, spin-offs, and the like could generate wealth for those undertaking these activities by improving corporate management and efficiency. But it is also possible that such activity can benefit its undertakers by transferring wealth, rather than generating it.

Usually, the issue of wealth re-distribution is debated when so-called "greenmail" is paid, effectively by existing shareowners to potential raiders. But, since employees are stakeholders in the firm, it is possible that a wealth transfer from employees to new owners could also occur. Instances in which wages or benefits are reduced, or other conditions of work deteriorate, may fall into this category. The critical issue is whether the enterprise was viable at the old pay rates and conditions.

For example, suppose a newly de-regulated business finds itself burdened with pay and benefit obligations established during the prior period of regulation and protection. If it faces new-entrant competition at lower rates of pay which would make the firm uncompetitive, new owners may enter the picture and proceed to lower labor standards. But there is not a transfer of wealth from employees to the new owners in this case; rather de-regulation has transferred wealth from employees to consumers. Although employee anger may focus on the new owners, they were

simply the instruments and reflection of a change in government policy (de-regulation).

Alternatively, if new owners simply renege on past implicit commitment to employees in the absence of a tangible deterioration in the economic conditions facing the firm, higher short run profits may be generated. In such cases, there may be a wealth transfer from employees to new owners. Much depends on the specific circumstances.

Consider first a circumstance in which previous owners-- through their managers -- adopted HRM policies on compensation and other matters that turned out to be excessively generous from the firm's viewpoint. That is, the extra productivity, loyalty, etc. generated by the policies did not outweigh their costs at the margin even though the firm could nevertheless operate at a reasonable profit. If a change in ownership results in a take-away of the policies, there is a simple wealth transfer from employees to the new owners.

If, however, the previous HRM policies -- though seemingly generous -- were optimal, a shift toward reduced pay and conditions might still generate short run accounting profits. In the long term, it might be argued, the external valuation of the firm should be lower -- even if short run accounting profits increase -- because of the departure from optimal HRM practice.

The situation is analogous to a firm which generates short term accounting profits by neglecting optimal maintenance of equipment.

Economic theory suggests that external evaluators -- those buying and selling the company's shares in the case of a publicly-traded firm -- would see through the "veil" of accounting data. In the case of the maintenance deferral, they would realize that the firm was accelerating actual physical capital depreciation in a non-optimal fashion. And in the case of the deterioration in pay and conditions for employees, they would realize the firm was non-optimally depreciating its stock of human capital. Given information costs -- and the difficulty of determining precisely the optimal level of maintenance expenditures or HRM generosity -- the external market may not always pierce the accounting veil.

Where collective bargaining is involved in setting the firm's HRM policies, shifts in bargaining power can lead to transfers of wealth back and forth between employees and owners. For a variety of reasons, discussed in a later chapter, employer bargaining power increased in the 1980s. Sometimes changes in management, which accompanied restructuring, were the vehicle through which the increase in employer bargaining power was expressed. New management came in and demanded -- and often

received -- wage, benefit, and workrule concessions from the union or unions involved.¹¹

In any case, if the volatility of corporate ownership and control which developed in the 1980s continues indefinitely, changes in the implicit contracts offered to (or accepted by) workers in the future are likely to occur. Lack of trust in, or ability to make, forward employment contracts would result in the pay profile of Figure 1 rotating clockwise towards the MRP_L line. In effect, employees would be paid what they are worth, when they are worth it. Performance appraisal and merit rewards would continue, but the rewards might be weighted towards current year bonuses rather than "permanent" upward shifts in an individual employee's pay rate.

Some symptoms of a shift in this direction developed in the 1980s as the use of temporary employees increased. Temporary employees -- whether hired through a personnel supply agency or directly by the employer -- typically are employed under "spot" contracts. That is, they are paid for the work they do, when they do it, and with no commitment by either party to a continuing employment relationship. Employers who wish to maintain forward implicit contracts with a core group of employees may find it desirable, in a volatile world, to rely more heavily on such spot arrangements with a second tier of temporary employees than has been the practice in the past. The

temporaries can absorb the peaks and valleys of production, thereby insulating the core group.¹²

IV. The Cost of Merit and Promotion Systems.

Given the importance of labor costs to most firms, it is essential that HRM professionals develop the ability to project such expenses. Unfortunately, the existence of merit pay (and, to some extent, promotion) systems seems to be a source of endless confusion for those charged with such costing. The costs of merit are often overstated by the very professionals who should know better.

It is important to start with the observation that in a steady state, i.e., a situation in which the firm maintains its employment level over a long period of time in the face of normal turnover, a merit system should not cost anything. This surprising conclusion holds despite the fact that individual employees under the system are receiving regular merit pay increases. Although the case of promotions is not explicitly elaborated below, the reader will quickly see that the issues of merit awards and promotions are closely related.

i. A Numerical Example.

Figure 2 presents a hypothetical example of a job with a rate range spread between \$10/hour and \$14/hour in five equal steps. Employees enter the occupation in step 1 (\$10) and progress annually to the next step. (A seniority-based pay system is being assumed for pedagogical purposes, but there will be no difference between merit and time advancement if, under the former, employees progress "on average" one step per year). In year 1, employees A, B, C, D, and E are spread over the five steps. The following year, employee E retires and a new replacement -- employee F -- is hired as a replacement. Also, A, B, C, and D are advanced one step.

It can be easily seen that the total hourly payroll will be the same in year 1 and in year 2 ($\$60 = \$10 + \$11 + \$12 + \$13 + \14) as will the hourly wage ($\$12 = \$60/5$). Although the four employees who remain with the firm in year 2 (A, B, C, and D) each get \$1 raises (\$4 in total), the replacement of E at \$14 by F at \$10 saves the firm an equivalent amount ($-\$4 = \$10 - \$14$). The net cost of the merit system is, therefore, zero in a steady state situation.

Of course, at any point in time, an employer will probably not be in precisely the steady state. However, the basic principle holds. What determines the net costs of merit is the

Figure 2

Hypothetical Example of Gross and Net Costs of Merit Pay

	Hourly Wage	Employee Distribution		Incremental Cost of Merit Pay Adjustments, Year 1 to Year 2	
		Year 1	Year 2	Gross Cost Basis	Net Cost Basis
Entry Step 1	\$10	A	F		-\$4 ¹
Step 2	\$11	B	A	\$1	\$1
Step 3	\$12	C	B	\$1	\$1
Step 4	\$13	D	C	\$1	\$1
Top Step 5	\$14	E	D	\$1	\$1
<hr/>					
Total hourly payroll	\$60				
Mean hourly payroll	\$12				
Increment to payroll				\$4	\$0
Increment/payroll				6.7%	0%

¹Difference between wage of newly-hired employee F and departing employee E.

change (if any) in the proportions of workers at each step. In the steady state, the proportions remain unaltered from year to year. But in the real world, there could be modest shifts in the proportions depending on surges in new hiring (which increase the "weight" of the lower step), hiring freezes (which decrease the entry weight), the age distribution of employees (which will determine how long they stay in the top step before retiring), etc. As long as there is a clear distinction made between net and gross costs, none of these departures from the steady state pose analytical problems.

ii. Gross vs. Net Confusion.

Figure 2's arithmetic is so simple that the reader may have difficulty understanding why confusion should ever exist. Nevertheless, there is ample evidence that merit pay systems are a source of confusion. As an example, during the anti-inflation wage controls program of the Nixon administration, the rules proposed for costing merit plans provoked a substantial controversy. Yet the various participants in the dispute, including the rule makers themselves, seemed to have great difficulty understanding the difference between gross and net costs.¹³

A second bit of evidence comes from the responses of HRM specialists to various private surveys concerning the wage

increases they have awarded or are planning to award. Compared with other data on wage trends, the responses appear consistently too high. This upward bias suggests that respondents have trouble differentiating the net and gross costs of merit, and report the latter when they should use the former.¹⁴

On Figure 2, the gross cost of merit is \$4, the equivalent of 6.7% of payroll. Thus, unless the "savings" of replacing expensive employee E with cheap employee F are recognized, it might (erroneously) appear that having a merit system raises average pay by 6.7%. This arbitrary gross amount, of course, simply results from the assumptions of the example. A more typical annual gross cost of merit pay in real world plans is 1-2% of payroll. However, regardless of the amount, the gross cost is not appropriate for judging the impact on employer expenditures. Only the net amount (zero in the steady state example of Figure 1) is appropriate.

There appear to be two reasons for the common confusion between net and gross merit costs. The first has to do with the problem of supervisor misincentives discussed earlier in this chapter. And the second is related to the propensity of firms to mix merit pay with general, across-the-board pay adjustments.

iii. Merit Budgets and Misincentives.

The merit plan of Figure 2 is tightly controlled with "normal" annual steps carefully delineated. However, even in such a system considerable supervisory discretion might be allowed. For example, a supervisor might be allowed to decide that a particularly meritorious employee could jump a step, e.g., move from step 3 to step 5. Since there are pressures on supervisors -- in the absence of other constraints -- to give out merit increases, such discretion might cause the supervisor to raise the average pay level from step 3 in year 1 to, say, step 4 in year 2.

For example, suppose employee A were jumped to step 4 (\$13) and B and C were boosted to step 5 (\$14). Suppose, as in the previous example, that D advanced normally to step 5, E retired, and F entered employment at step 1 (\$13). The average wage in the unit would rise from \$12, the equivalent of step 3, to \$13, the equivalent of step 4. Such an increase would represent a rise in net costs of \$5 or 8.3% of payroll as shown on Figure 3 (including the \$4 savings of replacing E with F). Gross costs would rise by \$9 or 15% of payroll.

To guard against profligate awards of merit increases, firms will often ration supervisors through the imposition of "merit budgets." A supervisor in the example just cited might be

Figure 3

**Hypothetical Example of Gross and Net Costs of Merit Pay
with Supervisory Discretion**

	Hourly Wage	Employee Distribution		Incremental Cost of Merit Pay Adjustments, Year 1 to Year 2	
		Year 1	Year 2	Gross Cost Basis	Net Cost Basis
Entry Step 1	\$10	A	F		-\$4 ¹
Step 2	\$11	B			
Step 3	\$12	C			
Step 4	\$13	D	A	\$3	\$3
Top Step 5	\$14	E	B,C,D	\$3+\$2+\$1	\$3+\$2+\$1
<hr/>					
Total hourly payroll	\$60				
Mean hourly payroll	\$12				
Increment to payroll				\$9	\$5
Increment/ payroll				15%	8.3%

¹Difference between wage of newly-hired employee F and departing employee E.

constrained by assignment of a merit budget of \$4. Within that budget, the supervisor might exercise discretion by awarding more to one candidate than another. For example, employee A might receive a \$2 increase, B and C a \$1 increase each, and D might be bypassed. Or the awards could be spread evenly as in Figure 1. Either way, the \$4 merit budget constraint holds the net cost of the merit system to zero, given the \$4 turnover savings.

However, the accounting artifact of \$4 can easily become a source of confusion. A supervisor who "spends" (allocates) a \$4 merit budget might believe that he or she had raised pay by 6.7%, since the merit budget is presented on a gross basis. And, as noted, this confusion is surprisingly common.

iv. Mixing of Merit and Other Pay Adjustments.

Compounding the confusion created by merit budgets is the temptation to mix merit pay and across-the-board pay adjustments. In the example just cited, the rate range for the job (\$10 to \$14) was unchanged from year 1 to year 2. However, external market wages typically rise from year to year, especially during periods of general inflation. Thus, the firm might feel the need to raise the rate range by, say, 40¢, thus increasing the steps to \$10.40, \$11.40, etc. Spread over the five steps, the net cost would be \$2, i.e., $40¢ \times 5$, or 3.3% of payroll.

Often, nonunion employers prefer to insist that all their pay increases are based on merit. Sometimes this assertion is maintained even during periods of inflation when there is clearly an across-the-board market factor included in wage decisions. In a period of generally rising wages, such an employer might raise the top and bottom step in Figure 1 by 40¢ and give the supervisor a \$6 budget to award increases. Implicitly, \$2 of the \$6 is for inflation and \$4 is for merit. However, the \$2 is part of the supervisor's discretionary adjustment fund and is not nominally being awarded to all workers.

Unfortunately, by throwing both types of monies into the same pot, and calling the entire sum a "merit budget," ambiguity is inevitably created. The average wage is raised by 3.3%. But the supervisor has a (gross) budget for increases of \$6, creating the misleading appearance of a 10% average increase.

The impact of inflation and other factors on internal wage policies will be discussed more fully in a later chapter. However, the examples just reviewed indicate that periods of inflation have a cost beyond the nominal dollars expended for labor. Inflationary periods have the effect of distorting merit systems which do not clearly separate the amount being awarded for merit from that which reflects the upward trend in market wages. Of course, these costs could be minimized by a more

careful (and realistic) segregation of merit pay decisions from other forms of pay adjustment.¹⁵

V. Conclusions.

The evaluation of performance and pay turns out to be a complex matter. While the general notion that better performers should be rewarded is not controversial, the actual workplace implementation of that principle faces many obstacles. Good performance must be evaluated by someone, typically a supervisor, who acts as an agent for the employer. Yet, it is difficult in practice for the principal (employer) to create the incentives necessary for the agent to act in the best interests of the firm.

Linking the evaluations produced by performance appraisals to merit pay and other forms of advancement is also a complicated process. While employers (absent union pressure) often eschew seniority as a criterion for wage progression and advancement, there are reasons to suspect that seniority is actually an important consideration. Implicit contract theory suggests that pay and productivity will not be tied tightly together, except over an extended horizon. However, changes in the economy in the 1980s may force a closer current tie between individual performance and pay.

Finally, the use of merit pay systems seems to create confusion about the trend in labor costs. The gross costs of merit exceed the net costs, because the former excludes turnover savings. But control systems, designed to limit supervisory discretion, and external inflation conspire to blur the important gross/net distinction.

FOOTNOTES

1. Collective bargaining arrangements are discussed in later chapters. Private sector supervisors after 1947 could form unions if they wish, but they had (and still have) little or no legal protection for such activity. That is, they can be fired for being union members or for attempting to form unions. However, some statutes covering public sector employees give supervisors some protected rights to unionize.

2. Sanford M. Jacoby, Employing Bureaucracy: Managers, Unions and the Transformation of Work in American Industry: 1900-1945 (New York: Columbia University Press, 1985), chapter 1.

3. A "tight" labor market is one in which unemployment is very low and vacancies are plentiful, i.e., it is one which benefits the employee. The opposite situation -- when unemployment is high and jobs are scarce -- is known as a "loose" labor market. These terms are often confused in the media and in common parlance.

4. The topic of sexual harassment is part of the general subject of equal employment opportunity (EEO) which will be discussed in a later chapter.

5. Fred K. Foulkes, Personnel Policies in Large Nonunion Companies (Englewood Cliffs, N.J.: Prentice-Hall, 1980), chapter 15; Industrial Relations Department, National Association of Manufacturers, Settling Complaints in the Union-Free Environment (Washington: NAM, 1982); Ronald Berenbeim, Nonunion Complaint Systems: A Corporate Appraisal (New York: The Conference Board, 1980); Bureau of National Affairs, Inc., Policies for Unorganized Employees, PPF Survey No. 125 (Washington: BNA, 1979). Employee complaints and systems for handling them are discussed in a later chapter.

6. The question of the optimal size of firms -- which we cannot discuss further here -- is obviously linked to the trade off between efficiencies of centralized control and diseconomies of hierarchical delegation of authority.

7. Such arguments are sometimes made in connection with public policies which provide for retirement incomes, such as Social Security, or tax-favored treatment for pension plans. It is argued that individuals are shortsighted and will not on their own provide for later income needs. Of course, if young employees are shortsighted, they might not look kindly on an employer which cut pay now in exchange for more pay later.

8. In principle, the worker would not retire unless the firm compensated him/her for the discounted value of all potential future "overpayment." But, if that is the price for retirement, the firm would not gain any cost saving. Nevertheless, since workers cannot be sure of their health or life expectancy, and since they are likely to be risk-averse, an offer of a financial inducement of less expected value than the stream of future overpayments might still produce a retirement.

9. Although some studies suggest that pay rises with seniority (controlling for other factors), there is conflicting evidence. See Katherine G. Abraham and Henry S. Farber, "Job Duration, Seniority, and Earnings," American Economic Review, vol. 77 (June 1987), pp. 278-297.

10. U.S. President, Economic Report of the President, February 1985 (Washington: GPO, 1985), pp. 187-216, presents a standard economic appraisal of this activity.

11. Critics of corporate restructuring have argued that it is a device to "discipline" labor. See, for example, Barry Bluestone, "Deindustrialization and Unemployment in America" in Paul D. Staudohar and Holly E. Brown, eds., Deindustrialization and Plant Closure (Lexington, Mass.: Lexington Books, 1986), pp. 3-15, especially pp. 12-14.

12. Bureau of National Affairs, Inc., The Changing Workplace: New Directions in Staffing and Scheduling (Washington: BNA, 1986); Max L. Carey and Kim L. Hazelbaker, "Employment Growth in the Temporary Help Industry," Monthly Labor Review, vol. 109 (April 1986), pp. 37-44.

13. Arnold R. Weber and Daniel J.B. Mitchell, The Pay Board's Progress: Wage Controls in Phase II (Washington: Brookings Institution, 1978), pp. 89-93.

14. Sanford M. Jacoby and Daniel J.B. Mitchell, "Alternative Sources of Labor Market Data" in Barbara D. Dennis, ed., Proceedings of the Thirty-Eighth Annual Meeting, Industrial Relations Research Association, December 28-30, 1985 (Madison, Wisc.: IRRR, 1986), pp. 42-49, especially pp. 46-48.

15. Economists often have difficulty explaining, in terms of standard theory, why inflation is a problem? In principle, if all prices rise at the same rate, and the rate of inflation is recognized, no one should be made any better or worse off. However, it is sometimes argued that inflation causes confusion among actors in the economic system. General price (and wage) increases are misperceived as relative price (wage) increases. The merit problem regarding inflation is an example of such confusion.