



IRLE WORKING PAPER
#37-92
February 1992

Why Should Human Resources Managers Pay High Wages?

Lloyd Ulman

Cite as: Lloyd Ulman. (1992). "Why Should Human Resources Managers Pay High Wages?" IRLE Working Paper No. 37-92. <http://irle.berkeley.edu/workingpapers/37-92.pdf>



Title:

Why Should Human Resource Managers Pay High Wages?

Author:

[Ulman, Lloyd](#), University of California, Berkeley

Publication Date:

02-01-1992

Series:

[Working Paper Series](#)

Publication Info:

Working Paper Series, Institute for Research on Labor and Employment, UC Berkeley

Permalink:

<http://escholarship.org/uc/item/8378t1rz>

Keywords:

Ulman, human resource managers, wages, internal labor markets

Abstract:

This essay is about human resource management, internal labor markets, and assorted theories of wage behavior. A stereotype of managerial activities and policies is sketched out in the first section, which, although reflecting a parochial (U.S.) orientation, is intended to approximate a set of real-world conditions to which analyses of labor market behavior should presumably relate. In the second section, we consider "institutionalist" interpretations, which either supplement or challenge standard market analysis by appeal to the historical record, the behavioral sciences, or even the consensus of expert behavior. In particular, the importance of "conventional" forces, based heavily on perceptions of equity, interpersonal preferences, and custom and practice, is revealed by the scholarship and insights of Henry Phelps Brown. Next (section II) are assessed the claims and contributions of a sample of theories based on the assumption of individualistic utility maximization in competitive markets: the theories of equalizing wage differentials, human capital, deferred compensation, transaction costs, implicit contracts, and (least conventional) efficiency wages. None is found to be lacking in interpretive value or relevance to one or more attributes of the stereotype, and all help to relate it to a wider family of markets and to economic behavior. In general, however, this group is less satisfactory in explaining why wages should be high enough—in a present value sense and relative to market-clearing levels—to contribute to the relative insulation of internal labor markets and to restrict employment. An exception is provided by efficiency wage theory, in particular by one of its older variants which, because it is based on group (rather than individualistic) psychology and behavior, is discussed in the following section (V).

Copyright Information:

All rights reserved unless otherwise indicated. Contact the author or original publisher for any necessary permissions. eScholarship is not the copyright owner for deposited works. Learn more at http://www.escholarship.org/help_copyright.html#reuse



Why Should Human Resource Managers Pay High Wages?

This essay is about human resource management, internal labor markets, and assorted theories of wage behavior. A stereotype of managerial activities and policies is sketched out in the first section, which, although reflecting a parochial (U.S.) orientation, is intended to approximate a set of real-world conditions to which analyses of labor market behavior should presumably relate. In the second section, we consider “institutionalist” interpretations, which either supplement or challenge standard market analysis by appeal to the historical record, the behavioral sciences, or even the consensus of expert behavior. In particular, the importance of “conventional” forces, based heavily on perceptions of equity, interpersonal preferences, and custom and practice, is revealed by the scholarship and insights of Henry Phelps Brown. Next (section II) are assessed the claims and contributions of a sample of theories based on the assumption of individualistic utility maximization in competitive markets: the theories of equalizing wage differentials, human capital, deferred compensation, transaction costs, implicit contracts, and (least conventional) efficiency wages. None is found to be lacking in interpretive value or relevance to one or more attributes of the stereotype, and all help to relate it to a wider family of markets and to economic behavior. In general, however, this group is less satisfactory in explaining why wages should be high enough—in a present value sense and relative to market-clearing levels—to contribute to the relative insulation of internal labor markets and to restrict employment. An exception is provided by efficiency wage theory, in particular by one of its older variants which, because it is based on group (rather than individualistic) psychology and behavior, is discussed in the following section (V).

Lloyd Ulman
Professor (emeritus)
Institute of Industrial Relations
University of California
at Berkeley

This paper was written with the generous and valuable support of the Institute of Industrial Relations on the Berkeley campus of the University of California. I am especially indebted to my colleagues William Dickens, George Strauss, and Sheldon Zedeck for their kind helpfulness and interest. Caroline van Ryckeghem and Dennis Toseland provided expert and devoted research assistance.

I. THE STEREOTYPE

Over the years, a considerable variety of terms has emerged in an effort to define the role of labor managers in large American firms. Behind the variety and, in some cases, euphemism there has lurked a continuing effort to play down the adversarial aspects of industrial relations, to subordinate the firm's collective dealings with externally-based organizations to its individual relationships with its own employees, and to emphasize the professional nature and strategic importance of the latter functions. However, because of the dramatically accelerated decline in union membership and military in the past decade, combined with the continued growth in importance of the professional and technical content of labor management, "human resource management" (or HRM), the title under which the practitioners currently rejoice, has provided thus far to be less of a euphemism than such predecessors as "human relations" and even industrial or personnel relations.

Professionalization and specialization have grown apace with the systematic internalization of important labor market activities. These include, to begin with, the analysis of jobs and their description in terms of a common set of required performance characteristics—skill (in terms of education, training, and judgment and initiative), responsibility (for capital equipment and materials and for the safety and performance of others), effort (physical and mental), and working conditions (including job hazards). Such information is a prerequisite to the determination of performance standards, often via motion and time study. It is also prerequisite to the determination of the firm's desired internal wage structure, which requires some ranking of jobs that cuts across occupational and departmental boundaries. Such a ranking may be either ordinal or cardinal. In the latter case, the process of job evaluation assigns a quantitative score to each job based on the "degree" of each characteristic (or "factor") which it (or some key reference job) is adjusted by management (or some outside specialist) to possess and further, on the relative weights assigned to the factors themselves.

With information obtained from job analysis, the firm can also recruit actively in appropriate "external" labor markets and then select those individuals whose characteristics are best matched to job specifications. It can then proceed to domesticate its employees through various training and incentive programs, the establishment of "job ladders", a system of employee performance appraisal, and standards governing promotion, on the one hand, and discipline, on the other. HRM will also monitor and modify its programs and procedures in the light of information received from employee opinion surveys, grievance procedures, suggestion systems, or other feedback mechanisms.

These managerial activities are specialized and draw upon a variety of academic disciplines, such as industrial engineering, statistical analysis, psychology, and sociology, as well as economics. Technical expertise has helped human resource management in efforts designed to change the working environment as well as to adapt to it. Taken as a whole, both approaches have entailed the deployment of specifically targeted policy instruments; and they have been intended to respond to the diverse needs and

circumstances of a heterogeneous workforce. Jobs have been made less hazardous and unhealthy, and less transient. Work may also be made less monotonous and, hopefully, more rewarding.

As a result, line management may be restricted not only in its choice of technology or plant design but also in its administrative style. The latter occurred as the arbitrary exercise of supervisory authority gave way to procedures administered by specialized staff and based on explicit rules that are binding on management itself as well as on the operating employee. Requiring that employment opportunities be rationed by application of the principle of "equal division of work" or by layoff and recall in order of seniority (or by some stated sequence of the two methods) furnishes a specific example of government by law and not by foremen. Others are provided by rules and procedures governing promotion, transfer, and discipline, referred to above. They may not have eliminated what David Feller (1973, p. 737) has characterized as "the authoritarian nature of the employment relationship," but they have considerably reduced the element of arbitrariness.

Authoritarianism may be next in line. In recent years, there has been an upsurge of interest in "participatory management" in the advanced industrialized countries of Europe and North America, inspired mainly by the widely perceived effectiveness of the Japanese system of industrial relations. Such programs vary widely in scope and content, but they are generally designed to extend to most production workers a measure of autonomy on the job hitherto reserved to (and presumably enjoyed by) highly skilled manual workers. (Meanwhile many traditional manual skills have been rendered obsolete by computerization and other facets of the new information technologies). Under participatory arrangements, greater blue-collar versatility and autonomy have been generated both by the creation of broader and fewer job classifications (as earlier processes of job "dilution" give way to job "enlargement" or "enrichment") and by rotation of tasks among members of work teams. As a result, the authority and numbers of supervisory, front-line managers have been further reduced.

Human resource management has also been associated with the development within the form of an ever-increasing variety of dedicated compensatory instruments. Private pension plans and systems of health, accident, and life insurance fall into this category, as do vacations, holidays, leaves for specifically designated purposes, and various benefits in kind. (Such private "fringe benefits" have grown much more rapidly than wages in the United States in the postwar period, and they have also increased more rapidly than social security and other state-provided benefit payments. (Mitchell, 1989, p. 203). Introduction of so-called "cafeteria-style" benefit programs, whereby the individual employee may select a subset of his or her own choosing from an array of alternatives, can impart an additional element of flexibility: the individual's compensation package can omit one or more benefits which might be redundant or otherwise inappropriate to his or her circumstances (Lawler, 1971, pp. 253-5) — medical insurance for an employee already covered under a plan provided for a spouse. Thus, while benefit plans can be tailored to different categories of employee needs and preferences, they can also make allowance for individual variation in preferences and circumstances.

Determination of both the structure and the average level of wages within the HR-managed firm has also exhibited some distinctive characteristics. To begin with, the wage rate is attached to the job itself (Kerr, 1957, p. 175). Wage variation among individual employees with the same general occupational grouping can be obtained within defined rate ranges and also via promotional and transfer opportunities, which are afforded by relatively large numbers of narrowly defined job classifications. (This has been especially the case in large U.S. firms; it is definitely not characteristic of Japanese firms.)

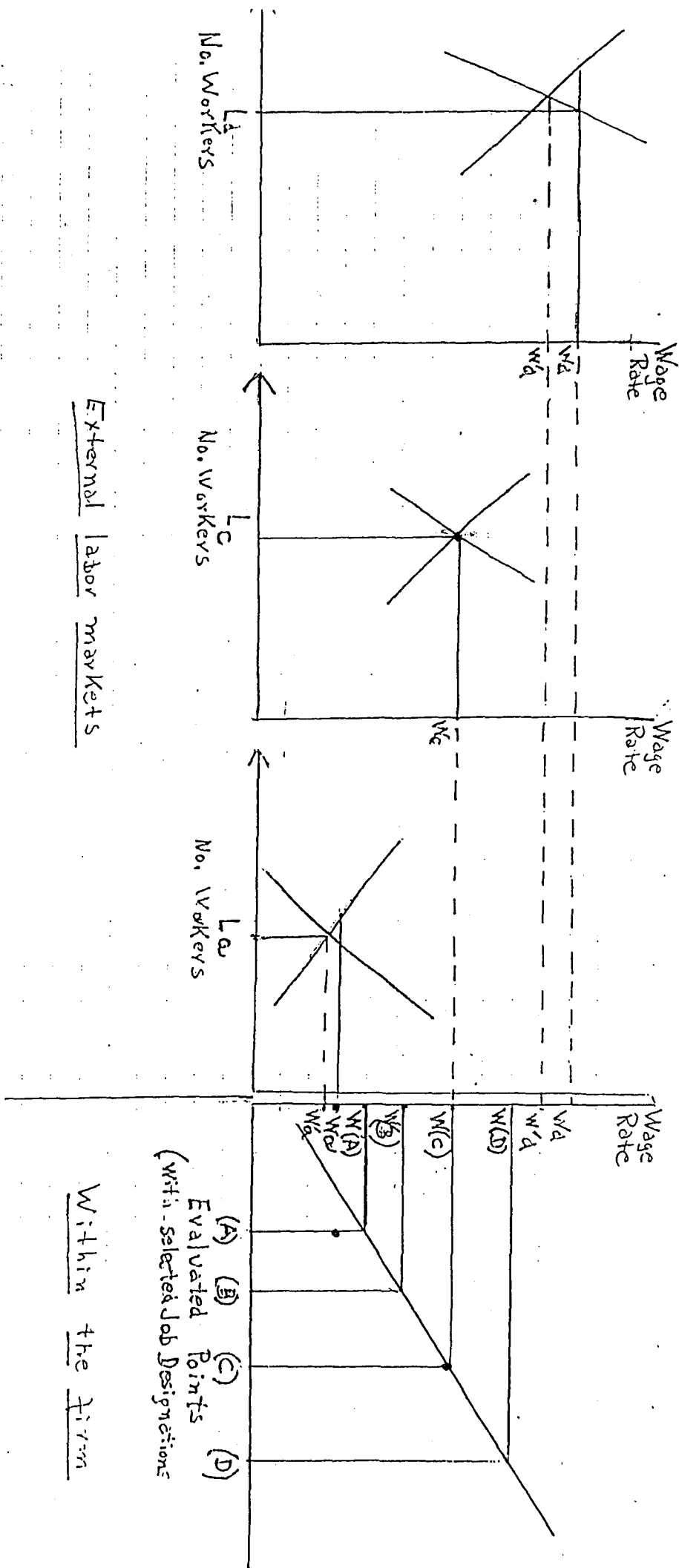
In the second place, a firm's wage structure, as determined by job evaluation, may yield wage rates which diverge significantly in either direction from median rates prevailing in relevant labor markets. This is not to say that wage rates in HR-managed firms are uninfluenced by external market conditions. Even where jobs are ranked and scored strictly on the basis of (nonwage) information derived from direct analysis of their personnel requirements, the conversion of "point" scores into money wage rates is based on wages paid elsewhere for work similar to certain "key" jobs within the firm in question. Key, or benchmark, jobs serve as "ports of entry" (Kerr, 1954) for new hires, nuclei in "job clusters" (Dunlop, 1957), or bottom rungs in promotional ladders. The wage rates and evaluated point totals attached to key jobs are determinants of the slope and the height of a wage line; the latter, in combination with point scores, determines the wage rates of the other jobs, which may be either highly specific to the firm in question or commonly found elsewhere. Point (B), in the right-hand sector of Figure 1, depicts an interpolated job which, unlike a key job, is not used to develop the job evaluation plan.

[Figure 1 about here]

The points (A), (C), and (D) in the Figure depict market-represented key jobs. It will be noted that the wage $W(C)$ is equal to the external market wage W_e , which is also a market-clearing wage. Neither condition need always obtain. Since key rates must reflect their own evaluated point scores as well as diverse market conditions, some of them may differ significantly from the median rates prevailing in their corresponding markets.

And when such disparities themselves are associated with differences between evaluated rates and market-clearing wage levels, management might be required to cope with divergence between the firm's demand for the type of labor in question and the supply available to it. Thus, with the evaluated wage $W(D)$ below both the prevailing market wage W_d and the market clearing level W'_d , the firm could experience a shortage of labor for job class D, although there might be an excess supply in the market ($W_D < W'_d < W_d$). (If, on the other hand, the firm's wage rate should equal or exceed the market clearing level, the firm need not suffer a shortage even if its wage were below the prevailing level ($W_d > W_D > W'_d$.) Management would have to adjust to its shortage of "d-type" workers. It might do so by drawing more work from the present force of "d-type" workers, as by increasing the number of overtime hours or the amount of incentive work; these are both devices which raise average hourly earnings at the evaluated rate and also increase the degree of capacity utilization, or the effective ratio of labor to

Figure 1: Stereotyped Wage Setting



External labor markets

Within the firm

capital. Alternatively, the firm could switch to more capital-intensive technologies, thereby substituting capital for labor (and obliging other branches of management to conform to the wage structure determined by HRM). The firm could also divert resources to the HRM functions of recruitment, selection, placement, training, and upgrading in the shortage categories. It also could reduce hiring standards. These methods of adjustment might be explored either in isolation or combination; all exemplify alternatives to raising the general level of wages paid to all categories of employees in the firm (raising the height of the wage line) in order to cope with specific shortages. (Lester (1955), Doeringer and Piore (1971), Osterman (1984)).

When the firm's evaluated rate exceeds the prevailing market rate, and when the latter in turn exceeds the market-clearing level ($W(A) > W_a \leq W'_a$), the firm finds itself paying more than is necessary to assure an adequate intake of appropriate labor.

Relatively less effort and fewer resources need be devoted to recruitment and training activities in these areas. On the other hand, the rationing task is greater and would presumably warrant a diversion of HRM resources to selection and placement. But reducing the firm's wage and moving it closer to the market median would be ruled out if that meant pushing it below the wage line: the firm is prevented from full exploitation of its market position by considerations of internal "equity". If, however, $W(A)$ should lie above the wage line—e.g., having been in place before the current wage line had been established—than it might be "red circled" and marked for reduction as individual incumbents leave the classification.

Obviously, the degree of either shortage or excess supply of labor to the firms associated with any given evaluated wage rate depends on the height of the wage line. In our Figure, the latter conforms to market conditions for labor hired to perform in the (C) category, which might, for example, comprise the major portion of this firm's work force, or be highly essential in its production process. But the determination of a firm's overall wage level might be dominated by considerations other than the availability or importance to the production process of specific types of labor. Thus, there is a general tendency for wage levels in HR-managed firms to exceed their industry or community averages. Human resource management tends to be elaborated most fully in larger enterprises, in which the economics of scale inherent in what Sanford Jacoby has aptly labeled the "employing bureaucracy" (Jacoby, 1985) and its professionalized and specialized functions can be most efficiently exploited. And the size of the firm or plant is invariably found to be a significant determinant of the wage level. (Knowles, 1983; Long and Link, 1983; Pugel, 1980). The market power and/or profitability of large firms tend to provide their human resource managers with some measure of discretion in determining pay levels; and while Fred Foulkes; (1980, pp. 158-167) study of a group of large nonunion companies in the United States reveals that not all of them chose to be wage leaders, their managers did determine their firms' relative pay standings as a matter of deliberate policy (and after systematic review of wage surveys).

At the same time, Foulkes reports considerable managerial preference for implementing pay level policies through individual "merit" adjustments rather than through general increases (even when the former are granted across the board and at

regular intervals). Moreover, human resource management has been typically reluctant to allow wage levels to respond sensitively to changes in market conditions which they expect to be temporary in nature, out of fear of adverse employee reaction to wage reductions. In the early postwar period, managers were initially disinclined to raise wages in response to labor shortages because they believed that it would be easier to remove candy from an infant than to reduce wage rates after the shortages disappeared (Duesenberry, 1958, p. 304). (This is not inconsistent with red-circling specific wage rates, which is a feasible expedient precisely because it avoids cutting wages of current members of the firm's work force.) And later on, human resource managers were most reluctant to reduce wage levels during general recessions (Foulkes, 1980), although this course of action has not infrequently been taken by firms in financial distress or severe competitive difficulty.

II. INTERPRETING THE STEREOTYPE: THE INSTITUTIONALISTS

The stylized facts laid out above have been subject to conflicting interpretations for over a quarter century. The issue was framed succinctly by Phelps Brown in *The Inequality of Pay* (1977, p. 130), when, in reference to Doeringer and Piore's seminal work on the internal labor market (1971), he cautioned:

to call this any kind of market is perhaps inappropriate, for it consists of those employments within any firm or organization that are largely insulated from market forces... There are certain 'ports of entry', through which the firm recruits outsiders, and what it pays here will depend on the state of the external market; but in other parts of its job structure it can vary pay within wide limits at its discretion, or by agreement with the employees concerned. Evidently, the scope for discretion at any point will be wider, the more closely the employees concerned are attached to the firm, and the more they themselves are concerned with their long-run prospects rather than their pay there and then.

The internal labor market has been "contested terrain" (to take Rick Edwards' expressive term out of context) for academics as well as managers and workers. The contestants have been the "institutionalists", who emphasize its insularity, and the neoclassical price theorists, who stress the penetrating power of competitive market forces. The institutionalists have long maintained that the influence of external market behavior is strong but limited, that it includes noncompetitive as well as competitive elements, and that it is subject to deflection by internal considerations. Not that the latter can be held responsible for all of the peculiar behavior observed in the internal market—peculiar, that is, when measured against the norms of competitive efficiency. Some anomalies can be laid at the door of legislated restraints, such as minimum wages, etc., the tax code (which, in the United States, subsidizes benefits at the expense of wages), and collective bargaining. But even in the absence of such interpositions,

anomalies could be produced by factors like “conventional forces” and the firm’s market position and profitability.

By convention [Phelps Brown wrote], we mean opinions about what is the right thing to do that are reinforced because most people agree in holding them. There is a general opinion that pay should be proportioned to deserts... When people consider how much a given man should be paid, they usually look not for the rate that will balance supply and demand in his part of the labor market but for the rate that will be fair, just, and equitable. One notion of the fair rate is that it shall be commensurate with the requirements of the work the man does: people try to arrive at it by comparing these requirements with those of other jobs whose rates they take as given ...
... job evaluation is only a painstaking application of the way in which people do continually think and argue about relative pay ... Firms find that it removes resented disparities and obviates disputes ... (Phelps Brown, 1962, pp. 125-6, 129.)¹

By constraining the firm’s internal wage structure and inhibiting its flexibility, the forces of convention have limited its effectiveness as an instrument for recruiting, selecting, training, promoting, and disciplining the firm’s work force. Thus, convention has contributed to HRM’s need to rely on and invest in the traditional personnel functions, which in this event can be viewed as substitutes for wage flexibility. It has also strengthened management’s incentive to substitute nonwage measures for wage adjustments. Even changes in techniques of production might be induced by shortages associated with persistent relative wage rigidities (as well as by changes in relative wages).

Some nonwage measures, on the other hand, appear not to have been prompted by inflexible or inefficient wage structures. They include specific-purpose benefit programs, which, as noted in the previous section, are designed to accommodate differences in workers’ needs and preferences. They also include the newer types of job redesign, which have been undertaken to reduce monotony and to enhance the autonomy of the worker. Yet monetary compensation has long been regarded as more efficient than nonwage instruments precisely because of its greater versatility and ability to accommodate differences and changes in preferences. Hence, recourse to these two important types of specifically targeted instruments reinforces the institutionalist view that the role of wages has been restricted and that, in the internal labor market as in Phelps Brown’s wartime army (1949), increased efficiency can be secured with nonwage incentives. In particular, current experience and experimentation validate his early judgment that the production process “which will suit human nature best” is that “under

¹In his celebrated study of “seven centuries of building wages” (which has unfailingly piqued the interest of Berkeley graduate students over the past tenth of a century), Phelps Brown, noting the lack of change in the daily money wages of Oxford carpenters and masons “in about 500 years out of 690,” commented, “It is unlikely that supply and demand exactly balanced at the ruling price; rather it must have been that their movements were not wide enough to overcome the inertia of convention.” (1981, p. 8.)

which the end product is within sight of those engaged on particular processes, who thus have more chance to feel that they are really making something instead of going through certain motions.” (Ibid., p. 54)

The wage structures and levels associated with internal labor markets and human resource managers owe some of their distinctive attributes both to the operation of conventional forms and/to the firm’s product market position and profitability, which determine the degree of discretion enjoyed by its policy makers. The first of these attributes is the job-based wage. “Institutionalized policy”, observed Clark Kerr (1957, p. 175) “regularizes, when it does not eliminate differentials among persons doing like work in the same plant” and prevails alike in “the union contract, the company job evaluation system, and government wage regulations”. “Personalized rates”, on the other hand, “reflecting the merit of the worker or the prejudice of the individual foreman or employer, are quite normal in the ‘natural market’.” (In fact, rates established under job evaluation move within a specified range above and below the computed wage line, and the range thus accommodates limited interpersonal variation in merit and seniority. A range is not included in Fig. 1.)

Downward wage rigidity is another attribute that attests to the power of interpersonal—and intergroup—comparisons. Keynes (who for these purposes might be claimed as a stellar institutionalist) attributed general wage rigidity to the worker’s perception that a reduction in one’s money wage would imply a simultaneous reduction in one’s relative real wages (Keynes, 1936, p. 14). And while he arguably derived this insight from observation and analysis of British wage behavior under collective bargaining in the interwar period (Keynes, 1963 *ed.*, p. 247), he generalized it to apply to a nonunion environment as well.

Finally, institutionalist explanations of the relatively high levels of compensation—in terms of benefits, opportunity for advancement, employment security, and equitable treatment—prevailing among HR-managed firms have tended to posit elements of economic indeterminacy and managerial discretion. By this it seems to be meant that human resource policies and behavior can be significantly influenced by factors other than those emphasized in standard price theory. Thus, drawing inspiration from work in the field of organizational development, Thomas Kochan and Peter Cappelli (1984) and Kochan, Harry Katz, and Robert McKersie (1986) have stressed the independent role played by the values of top managers and the resulting “organizational cultures.” George Strauss (1988) finds a concrete example: he characterizes firms which depart from the “traditional internal market system” by adopting broad job classifications, employment guarantees, extensive training, and indoctrination programs and employee participation, as motivated by “high commitment policies.”

The coexistence of alternative managerial policies, strategies, or tactics need not signify a background of economic indeterminacy. It may characterize a temporary condition: some of the policies observed are economically superior to others and will ultimately drive them out. Or it could mean that the observed alternatives are of equal effectiveness (or ineffectiveness), in which case the institutional observer would have drawn a policy distinction without an economic difference.

Conversely, the existence of a genuine zone of economic indeterminacy—which might be caused by conditions of oligopoly or rapid growth in the product market or monopsony (static or dynamic) in the labor market—does not guarantee that management will exploit it with restraint. Why, for example, should a monopolistic employer offer his employees a wage in excess of their reservation wage any more readily than a competitive employer? (c.f., Dunlop, 1957). Some radical labor economists and historians have claimed that large-scale American employers have indeed acted in this spirit (whether in the interest of profit maximization or out of ideological bias). According to Harry Braverman (1974), they adopted technical changes in order to “deskill” and homogenize their work forces and they installed different job ladders to divide the workers and reduce their bargaining strength. These historical generalizations have been disputed, notably by Bernard Elbaum (1984), Paul Osterman (1989), and Sanford Jacoby (1985). For the purpose at hand it suffices to note that the divide-and-conquer hypothesis does not square with the observation of the relatively high levels of compensation in the sectors in which it was intended to apply.

In fact, it can be maintained from Jacoby’s historical account that the rise of modern personnel management marked a movement away from a wage-minimizing and effort-maximizing model in the direction of a high compensation model (in all the dimensions reviewed above). The history of labor management in the United States has been characterized by a long (and still continuing) tussle between line management (originally in the person of the “driving foreman”) on the shop floor and a professional staff, which was rooted in a social reform movement of the early twentieth century and which sought to adapt the job to the potentialities of a new, post-immigrant generation of workers. Their claim to reconcile the social objective of human development with the firm’s objective of profit maximization was (and remains) subject to challenge; but their policies have been of a discretionary nature, taking advantage of whatever leeway might be afforded by the firm’s market environment. Indeed, Jacoby seems to regard the personnel movement as a potential agent of market intervention, seeking to plant operatives within the firm, in a broader struggle between the principles of “economic liberation” and “social protection” which Polanyi held to characterize modern industrial society (*ibid*, p.4). In *The Growth of British Industrial Relations*, Phelps Brown had discussed a similar movement of “industrial betterment” among British employers in the second half of the nineteenth century “as an outcome of their respect for their working people as human beings.” (1959, p. 80).

Does this story have its dark side? Slichter (in his “Notes on the Structure of Wages”) had written that “a high average value added per worker tends to produce high wages (and high earnings) because it produces liberal wage policies.” (1950, p. 22). But Doeringer and Piore, who later set out to explore the origins and properties of the internal labor market on their own, were also concerned with its “dual”, the “secondary” market, which they found to be characterized by “low wages and fringe benefits, poor working conditions, high labor turnover, little chance of advancement, and often arbitrary and capricious supervision.” (Doeringer and Piore, 1972, p. 170.) These conditions they attributed not only to the economic inability of employers concerned to establish internal labor market conditions, but also to “attitudes and demographic traits of the secondary labor force”, whose members “place little value upon job security in particular

enterprises.” They noted, however, that such demoralization may itself be generated endogenously and be attributable to the poor job environment and the restricted opportunities afforded to these workers. One might further infer that, to the extent that “liberal wage (and other) policies” might have tended to restrict employment opportunities in the “primary” or internal labor markets, they have contributed to depressed conditions and poorer labor quality in the markets from which they were insulated. This, however, could prevail only to the extent that labor costs in the primary markets exceed the upper limits of the zones of indeterminacy imputed to them. (And it is presumably ruled out in the Slichter example which credits employer policies with generating high “earnings” along with high wage rates.

In short, observation of the structure of internal labor markets and their policies furthered by personnel (later human resource) managers shaped institutionalist evaluations of the underlying forces at work. Institutionalists have expected to find evidence of competitive influences in the market, but they have expected to find evidence of market failures as well. Changes in wage rates (and in either the level or slope of the firm’s wage line) as well as in other conditions of employment may be traceable to changes in conditions of product or labor demand (including technological change) or supply (including workers’ preferences and expectations). But conventional forces, on the supply side, in conjunction with discretionary managerial policies on the demand side, might be strong enough to perpetuate (or prolong) labor shortages and/or underemployment. Moreover, they might be responsible for wage differentials which, as Dunlop insisted, “are not transitory ... not to be dismissed as imperfections” (1957, p. 22) and are apparently in violation of the competitive Law of One Price. Institutionalists also observe that human resource managers call on the behavioral sciences for help in fashioning instruments designed to cope with these forces. They conclude that there is more in heaven and earth than is met with in the philosophy of the competitive price theorist. Slichter summed things up in his own conclusion that “the models used in accepted wage theory are too simple and need to be supplemented. This is not exactly news.” (1961 ed., po. 378.)

III. INTERPRETING THE STEREOTYPE: THE NEW UTILITARIANS

The eclectic approach adopted by the Institutionalists has been rejected emphatically by American price theorists, especially of the “Chicago School”, in the postwar period. They hold strongly that the forces of individual maximizing behavior and competition are as invincible and pervasive as they are socially virtuous. They deny the importance of convention, market imperfection, and managerial discretion, and they presume the operation of the Law of One Price. The presumption may be rebuttable, but testimony offered by observers or practitioners or other sensory evidence would be dismissed as superficial (“anecdotal”) by economists who would find nothing amusing in a question put by Groucho Marx: “Who do you believe—me or your own two eyes?” Nor have these New Utilitarians been fazed by evidence on taste formation, etc., gleaned from the findings of the behavioral sciences, for, armed with their own behavioral

postulates, they have not hesitated to invade these other academic jurisdictions in an ambitious exercise of what Melvin Reder (1987) has referred to as “disciplinary imperialism.” If there is more in heaven or earth than is met with in their philosophy, it would be news to the New Utilitarians—but not for long, if they can help it.

But they would not deny that the theoretical models referred to by Slichter needed supplementation, because only with augmentation and elaboration was competitive wage theory able to address awkward problems that were posed by apparent wage anomalies and other symptoms of market “imperfection” or failure. For the Law of One Price to hold in the presence of significantly great and persistent wage differences, the latter must be clearly related to differences among jobs or to differences among workers. However, the theory of equalizing (or compensating) differentials, which relates wage differences to differences in job content or other nonwage conditions of employment, has failed to receive consistently strong empirical verification. This theory has long appealed to intuition and common sense; and it is certainly applied by job evaluators in awarding extra value to jobs that are generally regarded as undesirable because their normal working conditions result in discomfort, fatigue, hazard, etc. On the other hand, Foulkes reported that those companies in his sample which had pursued policies aimed at stabilizing employment (his “full-employment” firms) did not pay correspondingly lower wages; on the contrary, they were “also among the leaders in pay and benefits.” (1980, p. 153.) Nor did he observe any trade-off between wages and benefits; his high-wage companies were also characterized by high proportions of benefits to payrolls. (1980; pp. 223-4.) In general (as noted above), firms in which human resource management has been more fully developed and in which considerable resources have been invested in creating a good working environment (as well as in employee search, selection, and training) have also tended to pay relatively high wages.

In his survey of econometric studies which have sought to relate wage measures to various hedonic estimates of nonwage job characteristics and forms of compensation, Charles Brown (1980) found some support but “inconsistent support for the theory of equalizing wage differences” (p. 131). These and later studies typically include various measures of worker quality and control variables, which have usually proved statistically significant. Failure to find evidence of compensating differentials has nevertheless been attributed to the existence of unmeasured worker abilities, on the grounds that the most capable workers are found in jobs with the best working conditions. Yet Brown's own analysis of longitudinal data does not find improvement in the explanatory value of his job-characteristics variables after inclusion of a set of individual-specific characteristics. (Nor, incidentally, does this study reveal a wage-supplement tradeoff, which is consistent with Foulkes' findings.)

The theory's failure to secure stronger and more consistent econometric support has also been traced to the assumption of perfect information on which it is based. According to Sherwin Rosen (1986), this assumption cannot be satisfied on a current basis in markets in which workers with differing preferences seek to match up with employers with different job attributes or offers. Nevertheless, “the theory must be considered as one of longer run tendencies and of equilibrium behavior in the steady state of a more complex dynamic process.” (p. 643) But William Dickens (1990) relying

on psychological studies, maintains not only that job information is imperfect but that, even where workers do have information about the probability of occurrence of job hazards, they will ignore the odds where the odds are low. As a result, “employers have only limited incentives to pay compensating differences and measured differences will not reflect willingness to pay to avoid those risks—even, presumably in the long run.”

Failure to find stronger evidence of an equalizing element in wage differences may be due to departures from strongly competitive market conditions on which the theory is also based. In this connection, repeated findings of a negative relationship between wage rates and quit rates is significant. (Ulman, 1965; Pencavel, 1970; Freeman and Medoff, 1984; Dickens and Katz, 1985; Krueger and Summers, 1987.) If wage rates do in fact serve as strong and pervasive equalizers, they should not be correlated with quit rates: a worker should be no more inclined to leave a low-wage but otherwise desirable job than a high-wage job which she finds otherwise undesirable. Therefore, the existence of a negative quit-wage relationship suggests that workers are more reluctant to leave high-wage jobs because they are regarded as better jobs all around. This relationship is consistent with the existence of high-wage “primary” markets (with a high proportion of internal labor markets) which are relatively insulated from low-wage “secondary” markets. If (whether out of necessity or “policy”) human resource managers pay wages that are generally in excess of market-clearing levels, then they tend to pay “rents” on both good jobs and bad jobs and have no particular trouble in recruiting workers to the latter from outside the firm. And if wages and other conditions are generally poor in the secondary markets, workers therein might quit relatively good (or less bad) and worse jobs alike.

While the explanatory power of Adam Smith’s theory of equalizing differences, based on “the agreeableness or disagreeableness of the employments themselves” was revealed by postwar econometricians to be more limited than generally supposed, the explanatory range of a companion theory, based on “the difficulty and expence of learning” became greatly extended with the development of the theory of human capital. Smith himself had attributed a premium wage for skill to the worker’s need to recoup “the whole expense of his education, with at least the ordinary profits of an equally valuable capital.” (p. 101). In the postwar period, this insight was exploited and enriched by T.W. Schultz (1961, vol. 51) and especially by Gary Becker (1964), who pointed out that the marginal rates of return on alternative human capital investments would tend to equality (since all investments would proceed until their respective internal rates approximated the opportunity costs of funds.) (Rosen 1987, vol.2). In addition, he included on-the-job training in a unified analysis of “education.” And then came the distinction between “general” and “firm-specific” human capital; the cost of acquiring the former is necessarily incurred wholly by the wage-earner, in part by accepting wages below opportunity levels, whereas the cost of the latter is shared between the employer and the employee.

Many elements of our HRM stereotype can be interpreted in terms of human capital theory. In addition to the firm’s outlays on training, other HRM staples—recruitment, selection, placement, transfer, promotions and discipline—can all be interpreted as investments in specific human capital, since, to follow Walter Oi’s

(1962, p. 57) definition, they “increase... a worker’s productivity to a specific firm, without affecting his productivity in alternative employments.” Resources devoted to holding down labor turnover, promoting long-term employment relationships, and enhancing morale—by providing pensions and insurances and soliciting employee opinions and suggestions—also help to protect the firm’s investments in specific human capital. Evaluating jobs in terms of their skill and “responsibility” requirements can obviously be squared with human capital theory.

No self-respecting maximizing theory cares to assign an important role to the elements of managerial discretion and market indeterminacy, by which institutionalists have set considerable store. Human capital theory sought to minimize the importance of market indeterminacy by offering an alternative to labor market monopsony and product market oligopoly as an explanation of how positive inequality between the marginal product of labor and the wage rate might arise. When the firm’s marginal cost of labor includes a fixed component attributable to its investment in specific human capital, marginal productivity will normally exceed the wage rate, even under perfectly competitive conditions. Thus, the existence of market “imperfections” of either variety cannot be inferred solely on the basis of evidence pointing to an excess of marginal productivity over the wage rate.

Neither of course, can the existence of a specific human capital investment be similarly inferred. The two sources of discontinuity are not mutually exclusive and, indeed, may be complementary: Becker wrote that “monopsony power as a whole would appear to increase the importance of specific training and the incentive for firms to invest in specific human capital.” (Becker 1964, p. 2).

The fixed costs of the firm’s specific human capital investment and the associated inequality between the marginal product and the wage rate can also be exploited in an attempt to explain downward wage rigidity plus employment security. Indeed, it has been suggested that employers would protect their investments in human capital by retaining the employees in question and maintaining their wage rates during temporary downswings in product demand, not only as long as resources attributable to their work cover their wages, but even after they fail to do so. (Becker 1964, p. 125). However, human capital explanations of the persistence of wages in excess of marginal revenue productivity are incomplete. According to the theory, the wages of (mainly) generally trained workers should decline during downswings (Oi, 1962). Therefore, the wages of more specifically trained workers should also fall as long as they maintain their premiums over the former. Moreover, if the premia are maintained, employers can lay off their specifically trained workers, secure in the expectation that they will be able to recall them after demand has picked up again. Thus, the high degree of cyclical wage rigidity associated with HRM is not convincingly explained by reference to human capital theory. (Ulman, 1990, pp. 282-3).

Human capital theory would deny that wages in HR-managed firms would normally be set above market-clearing levels, whether in response to conventional force of habit or equity or as a matter of company policy. It does predict, of course, that marketable skills will fetch premia over the wages of unskilled labor, but this may not be

very useful in explaining high wages in firms with heavy concentrations of relatively low-skilled production workers and narrow classifications, which are approximated by our stereotype. On the other hand, human capital theory also predicts that the wages of even low-skilled workers will include a portion of the returns to whatever specific training they have received in their firms; and payment of relatively high wages to semiskilled workers is strongly characteristic of the stereotypical firm in the primary sector.

Two reasons have been advanced to explain this phenomenon. The first is to avoid quits and the costs of replacement and retraining. (Becker, 1964, pp. 21-3.) This reason for splitting both the costs and the returns associated with specific human capital investment has been routinely accepted. Yet an employee whose wage is equal to whatever he or she could earn outside the firm has no wage-related reason for quitting, especially since quitting entails costs of search, relocation, and/or retraining to the worker. Nor should the payment of a wage in excess of the worker's opportunity marginal product reduce a worker's desire to quit for nonwage-related reasons.

The wage premium for specific human capital has also been viewed as essentially a form of tribute, or blackmail, that is extorted from the employer by experienced workers who are in a position to refrain from imparting their own specialized know-how to their juniors. (Doeringer and Piore, pp. 32-3). Similarly, the awarding of credit, under job evaluation schemes, for the exercise of employee "responsibility" and the consequent minimization of "downtime" and other forms of underutilization of capital equipment could help to account for the payment of relatively high wages in large-scale firms. Increased compensation may well take the form of job security arrangements as well as higher pay. However, the payment to experienced employees of part of the return to investments in specific human capital—in whatever form and for whatever reason—should tend to generate competition among job applicants who would bid entry level wages down until net present values of career earnings were equalized everywhere. (Ulman, 1990; Groshen, 1991). Thus it is an implication of this theory that specific human capital *per se* cannot cause market segmentation, which is associated with higher present values as well as high current rates of compensation.

Empirical investigation has tended to lend stronger support to the theory in its broad, original form than to the more refined and elaborated versions. Wage rates have been consistently correlated with years or levels of education and hence with "the difficulty and expence of learning". Wage rates have also been correlated with measures of work experience, some of which are supposed to proxy on-the-job training and productivity. Yet a study of James Medoff and Katherine Abraham (1980) failed to reveal any positive relationship between experience and rated performance of managerial and professional employees in two large corporations, although experience and earnings were strongly correlated.

Moreover, two of the functions which wage rates are supposed to perform in equalizing present values have not received satisfactory empirical verification. These are, first, that wages should inversely "vary with the constancy or inconstancy of employment" and, second, that entry-level wage rates in firms should vary inversely with their average wage levels. Adam Smith pointed to the seasonal nature of the building trades and their

relatively high wages as evidence of the first relationship; but Dickens and Katz (1988) in an extensive (and more recent) analysis of interindustry wage differences found no evidence of an inverse relationship between wages and weekly hours of work. (Neither did they uncover an inverse relationship between wages and nonwage compensation, which, as noted above, should also be expected on theoretical grounds.)

Nor has empirical evidence of the expected inverse relationship between a firm's average and entry-level wages been forthcoming. An early study by George Hildebrand and George Delehanty (1966, p. 278) found that "effective minimum rates" in most high-wage industry groups rose relative to average wages during a period (1957-64) when "excess supplies of unskilled labor" were increasing. They attributed this in good part to union policy under collective bargaining; but Foulkes (1980, pp. 60, 154, 156) reported that starting rates in the large nonunion establishments in his survey tended to be as high as those in union plants even when average nonunion wages were kept above union levels. Recently, Erica Groshen, in the first econometric wage study to evaluate starting rates, found them to bear a positive relationship to average enterprise wages. This relationship is what might be expected to emerge from job evaluation under human resource management, but not from human capital theory.

Finally, the limited explanatory capacity of both the human capital and wage-equalizing theories is reflected in repeated quantitative analyses of interindustry wage differences, beginning with Slichter's "Notes" in 1950. These studies (while varying greatly with respect to coverage, data, and methodological sophistication) have uniformly found evidence of persistent "industry efforts" or wage differences within occupational categories, such that, as Dickens and Katz (1988) put it, "all workers in some industries are highly paid relative to similar work in other industries." Industry effects have survived increasingly extensive and careful efforts to approximate or control for unmeasured worker characteristics and nonwage conditions of employment; in the former case, most studies employing longitudinal data have found that when workers move from low-wage industries to high-wage industries, their own wages are increased (Vroman, 1978; Krueger and Summers, 1988). On the other hand, these industry effects themselves have been related to differences in such characteristics as firm size, capital-labor ratios, and profitability. These relationships, which have persisted over time, suggest the influence of habit, equitable comparison, and other "conventional" forces.

However, other utilitarian theories of wage determination emerged to suggest alternative interpretations of various facets of the HRM stereotype. We shall touch briefly on four of them: the theories of deferred compensation, transaction costs, implicit contracts, and efficiency wages. They are all interrelated but can be divided into two groups, on the basis of academic-ideological lines (i.e., with respect to their efficiency implications). The first two claim that (despite appearances to the contrary, as per Fig. 1), the institutions and policies of human resource management in large bureaucratic enterprises do not result in significant wage anomalies and market failures, but instead generate the types of outcome which conform to the requirements of competitive efficiency. The second two claim that, even under competitive conditions, wage-setting can result in market failure and unemployment. Deferred compensation and transactions

costs are in the spirit of human capital theory; implicit contracts and efficiency wages are more in the spirit of Keynes.

Edward Lazear's (1979, 1981) theory of deferred compensation (our label, not the author's) tracks human capital theory quite closely in that it features a rising age-earnings profile, but it offers an alternative (although not mutually exclusive) neoclassical explanation. As an application of agency theory, it argues that, because the employee can increase his or her effort wage by loafing on the job, or "shirking", the employer can increase productivity and profit by reducing the employee's incentive to loaf. This can be done by installing an upward-sloping age-earnings profile whereby the employee's wage rises continually and predictably over the course of his or her career of the firm. Initially, the employee's current wages fall below the value of their current marginal products in the firm, but later on, wages exceed current marginal products; hence the present value of both income and productivity streams are equal. The younger worker has an incentive to accept a relatively low wage, work diligently, and win promotions; and the older worker has an incentive to remain with the firm and work hard which varies directly with the excess of his or her current wage over his or her marginal productivity.

This theory, which features the wage rate as an incentive rather than an investment or an equalizer, readily offers an alternative utilitarian explanation for stereotypical long-term employment relationships, with provision for promotion, security, review, and discipline—in short, for rewarding effort and detecting and penalizing shirking. It can obviously accommodate the existence of pensions as a highly visible form of deferred compensation—and one which not only constitutes a disincentive for getting fired for shirking but can also induce retirement by older workers (when they would no longer desire to work if paid no more than the value of their current marginal productivity. (Lazear, 1990). And deferred compensation theory offers an explanation of divergence between marginal products and wage rates which in one respect is superior to the human capital story. For unlike the latter, deferred compensation theory provides for an upward sloping age-earnings profile "even in the absence of on-the-job or skill acquisition" (Lazear, 1981, p.606). Hence the empirical finding that, while wage rates have varied directly with expenses, employee productivity has not, poses no problem for compensation theory, as it does for human capital theory.

But deferred compensation is no more successful than human capital in accounting for the high entry rates which prevail in large-scale HR-managed firms. Under deferred compensation, as under human capital, "steeper age-earnings profiles are coupled with low starting wages"; indeed, the starting wage may be lower not only than the value of the worker's marginal product but even than his or her reservation wage. (Lazear, 1981, pp. 618 and 607 n. 2). Hence the prevalence of relatively high starting rates would imply relatively high present values of age-earnings profiles unless those profiles are correspondingly flat. In the latter case, their effectiveness as a work incentive is relatively weak. The former case—high starting rates associated with higher present values—is inconsistent with the assumption of perfect competition. Glenn Cain (1987, p. 286) put the point well: "The flat profile is not a challenge to neoclassical theory, as some segmentationists imply, but Mill's challenge of 140 years ago remains: reconciling

neoclassical models of competitive markets with persistent differences in the present values of career earnings among worker of comparable abilities and preferences.”

Another utilitarian approach which seeks to reconcile apparent inequalities between marginal products and wages and other apparent anomalies with the properties of economic efficiency is the theory of transaction costs. According to this theory, where impediments to free competitive exchange exist in product or factor markets, they induce the emergence of private institutions—notably firms—to correct them efficiently by internal means, thereby incurring transaction costs. Oliver Williamson (1985) has identified “asset specificity”, “uncertainty”, and “nonseparability” as important obstacles to free competitive exchange in industrial labor markets. Nonseparability occurs when individual productive contributions to work performed by a group are difficult to identify. The first two conditions according to an earlier contribution by Williamson, together with Michael Wachter and Jeffrey Harris (1975) are conducive to “opportunistic behavior” by a small number of experienced workers exploiting their mini-monopolies of firm-specific skills by (e.g.) threatening not to teach them to their juniors. To cope with such problems, to minimize turnover and costs of monitoring, and to encourage teamwork, long-term employment relationships with provision for job-based wage rates, procedures governing performance review and promotion, and pensions and other deferred compensation features are judged superior to “sequential spot contracts”. Hence human resource managers have emerged and have incurred transactions costs in performing all of the personnel activities which the firm has internalized and in acquiring the information necessary to perform them. And hence the human capital and deferred compensation theories can be subsumed under transaction cost analysis.

It is also claimed that the costs of collective bargaining to the firm, or at least some parts of them, can be classified as transaction costs. That is because the union allegedly contributes to economic efficiency in various ways. Thus by preempting wage bargaining, it minimizes opportunistic behavior by individuals, thereby increasing the profitability of the firm's investment in specific skills. In addition, as Richard Freeman and James Medoff have argued (1979), collective bargaining reveals to the firm the needs and preferences of its employees, mainly through the institution of the grievance procedure. The grievance procedure has also permitted management to minimize costs of uncertainty, change, and “bounded rationality” by permitting it to make changes promptly, subject to appeal. Establishment of promotion ladders and adoption of the principle of seniority has reduced the “authority relation” (to which we referred in the first section).

At this point, four questions arise which test the limits of the theory's explanatory range in the fields of institutional development and human resource management. First, why are unions necessary, in the transaction cost scheme of things, if the efficiency-promotion functions which they perform are also performed within the human resource departments of firms? Second, if unions do perform efficiency-promoting functions, why have they been on the decline (most markedly in the United States but in Britain and elsewhere as well)? Third, if on the other hand, unions are not cost-efficient institutions, how could their emergence as economic institutions be accounted for?

And fourth, why have human resource managements in nonunion enterprises instituted some of the same internal labor market policies and practices as those favored by unions?

The premise of our first question has been implicitly rejected by analyses which maintain that unions play a unique role in maximizing (or optimizing) productivity. This is the thrust of the “union voice” theory of Freeman and Medoff (1979) which in effect argues that unions, largely through their contractual grievance procedures, lower the cost of acquiring information about the concerns and preferences of the firm’s own (inframarginal) employees and thereby reduce turnover and improve employee morale and efficiency. It is not the case, however, that unions have provided management with their only efficient source of such information, since nonunion firms have utilized grievance procedures (as these authors explicitly recognize) and both union and nonunion firms have also had recourse to other channels, both informally and systematically (through employee attitude surveys, suggestion systems, etc.). Moreover, the costs of acquiring union-provided information are the strike and settlement costs of collective bargaining, which are not only presumably greater than the costs of using other channels of information but also greater than any net gains in productivity properly attributable to the use of union grievance procedures. In addition, econometric studies have consistently demonstrated that wages in union establishments or of union members have been higher than corresponding nonunion wages. Hence it can no longer be maintained that “unions pay for themselves;” so at least in this respect their emergence is not traceable to the requirements of economic efficiency, in the manner envisioned by transaction cost analysis.

The same evidence on relative labor costs provides an obvious reply to the second rhetorical question. If unions were cost-effective, profit-maximizing employers would, *cet. par*, presumably demand to be organized. If, however, unions were not cost-effective—as suggested by the evidence on union—nonunion wage differences (and assuming that the latter exceed any differences in productivity), profit-maximizing employers would *cet. par* avoid them, if the costs of avoidance did not exceed the value of the difference in relative costs. And in terms of the theory, unionism presumably would not have been induced in the first place, so there would be nothing to avoid, and no unionized firms. But, given the existence of unions in the real world, a widening of union wage differentials, such as occurred in the United States in the 1970s, would have been expected to contribute to the ensuing decline in unionized sectors of the economy.

The third question, however, asks, how could the unions, as bargaining institutions, have gotten there in the first place? Of course, unions, even as primarily bargaining institutions, have arisen as part of political and social movements; it is well recognized that, as Williamson (1985, p. 250) puts it, “Unions are complex organizations, have many facets, and serve many purposes.” These might be regarded provisionally as extra-theoretical origins and purposes of unions. It is also well-recognized that unions have been formed and exist primarily to exploit monopoly positions by skilled groups or otherwise strategically situated workers. These, however, might be regarded as counter-theoretical origins: when unions are called on to exploit or exacerbate impediments to free competition rather than to compensate for them, they furnish

examples of theoretically perverse endogeneity. The impediment does not generate an institutional corrective; instead, it generates an institutional exacerbation. Thus a consideration of the origins of these labor institutions suggests that the explanatory range of transaction cost theory has its limits, and that generality ought not be assumed.

Indeed, even competitive situations, which require no institutional interaction on the grounds of efficiency, may generate institutional reaction; and union political activity as well as collective bargaining organizations may become endogenous creations. According to John R. Commons (1913),

... Instead of "exploitation", growing out of the nature of production, our industrial evolution showed certain evils of competition. Instead, therefore, of an idealistic remedy sought for in common ownership the practical remedy always actually sought out has been the elimination of the competitive menace through a protective organization or protective legislation (p. 259)

The conflict is ultimately one between the interests of the consumer and the interests of the producer. (262)

The fourth question (which inverts the first) asks why managements have adopted some of the same policies and practices favored by unions when (we can now add) their relationships with the latter have been adversarial rather than cooperative. As Paul Osterman (1984, p. 9) noted, "... in many instances employers exhibited strenuous, even bloody, resistance to the establishment of seniority systems, formalized wage-setting procedures, and the like. The extent and nature of this resistance is difficult to rationalize in the language of the model." We shall postpone spelling out the obvious answer until the next section.

The theory of implicit contracts offers yet another explanation of (stipulated) divergences between marginal productivity and wage rates under competitive conditions. As first propounded by Martin Baily (1974) and Costas Azariadis (1975), it identifies greater risk aversion by employees than by employers as the causal element: relatively risk-neutral employers "insure" their more risk-averse employees against wage fluctuations (specifically wage cuts in bad times), and the latter are willing to pay for that insurance by accepting lower wage rates on average. Therefore, the (competitively determined) wage rate is less than marginal productivity at full employment, which is associated with excess demand or labor shortages. But when the demand for labor declines, wage rates stick at their previous levels, and exceed full-employment marginal products, so that unemployment ensues.

Unlike the other utilitarian theories, implicit contracts explain the outcome in terms of unemployment, instead of efficient market performance. It does not try to plug the productivity-wage gap with unmeasurable costs of human investment or unmeasurable transaction costs, but instead attributes it to a posited peculiarity of wage earner preferences. Implicit contract theory is more in the spirit of Keynes, in seeking to

explain market failures in a competitive environment. Yet its view of the proletarian psyche is more traditional than the Keynesian view; relative risk aversion claims an impeccable neoclassical provenance, whereas relative wage preoccupation, with its connotation of interpersonal comparison and money illusion and its sociological overtones, is decidedly suspect. Moreover, a sticky wage in bad times is bitter medicine for the Keynesian employer to swallow because it implies that he would cut wages if he were free to do so, whereas the (implicitly) contracting employer happily accepts sticky wages as part of a long-term deal which permits him to pay lower money wages over the course of the cycle.

Thus, implicit contracts can account more directly than the other utilitarian analyses which we have encountered thus far for the characteristic reluctance of human resource managers to reduce wages during bad states of business. It also provides a rationale for labor shortages (such as those depicted in Fig. 1) that are associated with wages when held below their current market levels. However, the specific shortages that occur in the stereotype are not cyclical phenomena nor are they attributed to risk aversion on the part of particular occupational groups. More importantly, average wages in the stereotype firms and in the primary sector tend to exceed market levels, whereas the competitive risk-aversion theory explains why equilibrium wage levels should be relatively low, not high. Firms offering implicit contracts to their employees enjoy lower wages than their noncontracting competitors, obliging the latter to offer contracts of their own or lose out.

Nor does this version of implicit contract theory provide a rationale for pensions and other devices deployed by human resource management to provide employees with long-term employment security. This is curious: why would not implicit contracts offer risk-averse workers employment stability as well as—or instead of—wage stability? As George Akerlof and Hajime Miazaki (1980) demonstrated, a contract guaranteeing stable employment and a stable wage rate could be more profitable to a firm than a contract which guarantees only the wage rate, if workers prefer to accept a still lower equilibrium wage rate for the firm's employment guarantee than for a wage-only contract. Therefore, risk aversion theory should predict sticky wages with full employment, rather than with unemployment (which is what it set out to explain).

In reply, it has been claimed that double-barrelled agreements would shift all the variability in the firm's income to profits and that the risk aversion of employers, while greater than that of their employees, is too limited for such agreements to be acceptable. Moreover, if the guaranteed wage rate is sufficiently low to make the wage-employment contract profitable, it could be too low to be "feasible"—i.e., to allow employers to lure enough workers away from a life of leisure on the dole. An employment-only guarantee, on the other hand, might be feasible. But if such a contract permitted the employer to pay a lower wage in bad times than in good times—and if it is assumed that workers are more ignorant of business conditions than their employers—the latter would have an incentive to cheat by pleading poverty and reducing wages when conditions were really favorable. (Azariadis and Stiglitz, 1983; Grossman and Hart, 1988; Green and Kahn, 1983). However, if the employer is obligated to reduce employment (and output) as much as wages under these conditions, he would have nothing to gain by

misrepresentation, and an employment-only guarantee would be “implementable” (as well as “feasible”). In this case, the global upshot would be a combination of full employment and flexible wages, so that once again the prize of explaining rigid wages and unemployment under perfectly competitive conditions eludes the group of the risk-aversion variant of implicit contract theory.

From our viewpoint, however, to the extent that the theory can contemplate implicit employment as well as wage commitments, it does more complete justice to managerial reality. Jacoby (1985) relates that the American reform movement out of which the profession of personnel management emerged in the late nineteenth century fostered the belief that unemployment was a problem for industry to solve and that it was up to management to promote employment stabilization within the firm. Contemporary human resource managers believe that some measures to increase employment security in the firm are consistent with profit maximization, and they have been willing to incur the extra training, hiring, transfer, overtime, and inventory costs associated with such measures. Their gains result mainly from improved employee morale and productivity (due in part to greater employee willingness to accept new methods of work) and also from lower costs of unemployment insurance and, as Foulkes, 1980, p. 188) put it, “favorable image in the community.”²

Gains, however, did not include lower wages, so that these employment contracts, as well as the reluctance of the firms in question to reduce wage rates, have been associated with the existence of relatively high wages. By the same token, “full employment” policies are seldom found in low wage firms and sectors. The combination of high wages and employment-protective policies is, as noted above, a characteristic of internal labor markets in large-scale, more profitable firms and therefore does not accord with the theory’s assumption of perfect competition.

Members of the wages-only school of implicit contracts might claim some support from the fact that, even within firms in primary markets, employment protection has been partial and restricted. When work is scarce, it may be rationed among all employees by reduced work schedules or it may be rationed by layoff and recall seniority. Hence, as a very rough generalization, it may be fair to say that implicit employment contracts are more restricted in coverage than implicit wage contracts. Moreover, implicit employment contracts have been broken, although more under *force majeure*—notably during the 1930s and again in the Eighties—than through mean-spirited exploitation of opportunities afforded by asymmetric information. But during those periods of severely depressed activity and in cases of financial stringency, wage promises were also abandoned.

²It should be added that contracts can be enforced, up to a point, by fear of loss by the firm of its reputation as a “good employer.” This flows directly from risk aversion theory, which implies that the penalty for laying off an employee in violation of an implicit employment contract consists of a wage increase (or “marginal employment premium”) for all remaining employees. The employer would be deterred from violating the commitment—provided that the wage “saved” from firing an employee is less than the premium plus the marginal product foregone. (Azariadis, 1975)

Finally, efficiency wage theory provides a utility-maximizing and competitive model (by Carl Shapiro and Joseph Stiglitz, 1984) which seeks to account for a combination of downwardly rigid wages, involuntary unemployment, and high real wages (relative to marginal productivity at full employment). This combination can result from employer attempts to minimize “shirking” by their employees. Shirking is allegedly a decreasing function of the cost of employee monitoring to the firm, the firm’s relative wage, and the level of unemployment. The relative wage and unemployment together determine the value of the alternative income which an employee might expect to receive if he or she is caught and fired for shirking; the firm will seek to prevent or reduce employee loafing (or increase effort and efficiency) by trying to raise its relative wage—together with, or even instead of, its outlays for detection. As all competitors raise their own wages, this becomes a self-defeating exercise, but the resulting increase in the overall level of wages tends to reduce the total demand for labor; and so, as the relative-wage disincentive to shirk is weakened, the unemployment disincentive becomes stronger. By the same token, wages will remain sticky or sluggish in response to a decline in aggregate demand, declining only to the extent that they are replaced by increased unemployment as a guarantor of employee efficiency.

Thus the wage that is sufficiently high to underwrite profit-maximizing levels of worker effort and efficiency may also be high enough to result in involuntary unemployment; but this wage-employment combination is not one of long-run competitive equilibrium. Unemployed workers will compete for scarce jobs by bidding down entry level wages, which will both increase the slopes of age-earnings profiles and reduce their present values. (Yellen, 1984). Unemployment plus high real wages now gives way to deferred compensation as a more efficient anti-shirking device. An analogy with the superiority enjoyed by a system of payment by results over the straight time wage as an incentive device would be in point, since a rising age-earnings profile is analytically identical to a system of increasing piece rates. A perfectly discriminating system of increasing piece rates, as Leontief (1946, p. 79) pointed out, can be approximated by a “broken exchange path... clinging as closely as possible to the (seller’s) marginal indifference line” (and hence off her supply curve of effort).

In response it has been argued that, while deferred compensation—like the posting of a bond “up front”—provides the worker with a strong incentive not to cheat the boss, it also provides the boss with an incentive to cheat the worker by firing him after he has begun to collect wages in excess of current marginal productivity and replacing him with a young bond-poster whose wage will be less than the value of his current marginal product. Lazear (1981) had argued that a cheating employer would forfeit his reputation as an honorable one, subsequently compel him to pay higher wages—especially high starting wages—and ultimately face competitive extinction. Still, if we return to our analogy with incentive pay systems, we find (A) that (at least in the United States) employers did not resist the temptation to cut piece rates in response to increased productivity and higher employee earnings, and (b) that, partly in response to opposition by employees to what they regarded as cheating, the pay systems themselves, rather than the cheating employers, were ultimately forced out of business. We might also recall that starting wages tend to be high, rather than low, in large-scale, high-wage firms, which, as Dickens, Katz, Lang and Summers (1989) reasonably surmise, have

generally been regarded as reputable employers. "Perhaps," they conclude, "there are limits on the size of bonds that bind before reputation constraints are reached for many firms." (p. 339) These writers also claim that monitoring activities of American firms are extensive, as indicated by hefty supervisory and security budgets.

If, however, marginal costs of monitoring are sufficiently low, direct administrative activity could be substituted partially for efficiency wages as well as for deferred compensation. Three characteristics of human resource management tend to hold down costs of monitoring. The first is the fact that the functions of employee recruitment and selection are designed in part to screen out bad risks at the outset. The second is that much monitoring is performed as a joint product, by the same individuals who perform other functions. The most familiar joint producer is the foreman, whose personnel functions—especially in connection with performance evaluation and promotion as well as disciplinary proceedings—are in addition to his production duties. Furthermore, team work and other forms of participatory management provide opportunities for peer monitoring along with social pressure on the individual not to let his teammates down. In this context, we must also note that absenteeism and tardiness, the most extensive forms of shirking, are readily detectable.

Another potential source of cost containment consists in management's ability to vary—to tighten or relax—standards of discipline. The strength of the theoretical deterrent to shirking is measured by the probable loss of wage income which it entails. The probable loss of wage income is the product of the probabilities of detection and of income loss if detected. Since our theory assumes that every act of shirking is punishable only by dismissal, the latter probability is constrained to unity. Hence the strength of the theoretical deterrent to shirking is measured solely by the probability—and hence the cost—of detection. In fact, discipline is carefully calibrated by modern human resource management. It is "graded" in accordance with the seriousness of the offense and it is applied "progressively", reaching maximum levels only after a specified number of repetitions. Steps along the way include "an informal, friendly talk," an oral warning, a written warning, disciplinary layoff, and demotion. (Pigors and Myers, 1977). Dismissal is undertaken only with reluctance, both because of the added cost of turnover and replacement to the firm and because discipline is regarded by HRM primarily as an instrument of correction. Now this appeal to real life does not invalidate the theory's conclusions in a purely formal and qualitative sense: it is theoretically irrelevant, for example, whether a dismissable act of shirking is defined as a first, second, or fifth offense. But it has some bearing on the empirical significance of this theory of efficiency wages. If management can react to a decline in demand by "tightening up" its disciplinary standards—increasing the severity of punishment at earlier stages—it can increase the probability of income loss at a given probability of detection and, therefore, at trivial marginal cost. This should make for greater downward wage flexibility; or, rather, it should reduce the capacity of this version of efficiency wage theory to account for the downward wage rigidity which actually exists.

The draconian disciplinary policy to which the employer is constrained in this neoclassical model might have been adapted for analytic convenience, but it happens to reflect a philosophy and a view of employee motivation which has long been rejected by

human resource management. The history of labor management in American enterprise reveals movement away from the early “drive system” under the control of the shop foreman and, later, away from the ideological sway of “Taylorism”, with its emphasis on “de-skilling” and, it might be noted, incentive wage systems. It also records a concomitant long-term movement in the direction of “human relations”, and the motivation of employees through a mix of pecuniary and nonpecuniary inducements. It marks a movement away from what Paul Pigors and Charles Myers (1977) called a “negative view” of discipline to a “positive view.” The negative view had been implemented by authoritarian managers who proceeded on the premise that “compliance depends on fear of penalties” (p. 300). Their recourse to dismissal tended to be uninhibited and was designed to serve as exemplary discipline—a “deterrent”—as well as “retributive justice.” The positive view, in contrast, adopted the premise that “most employees want to do what needs to be done, at work, and in their other conduct, to meet organizational requirements and accepted standards of behavior.” If rules of conduct are “fair” and “if the employment relationship is good in other respects, most employees can be counted on to exercise a considerable degree of self-discipline” (p. 302). “Inefficiency”, defined as “failing to do the amount and quality of work that were expected of the employee when first hired” (p. 307) is a well-recognized cause for discipline; but discipline is necessary primarily to prod the rogue elephants (those who had escaped detection and rejection prior to employment) back into the herd.

It would be painfully obvious to a human resource manager that the neoclassically oriented efficiency wage theory exemplifies a “negative” view, not only of the role of discipline but also of the psyche of the worker, who is seen at all times to be a potential shirker. (Neoclassical efficiency wage theory, of course, bears a strong family resemblance to other utility-maximizing theories in this respect: when, for example, our worker is not shirking he is very likely to be opportunistically shaking down his employer who, for his part, may be about to welsh on an implicit contract or two with the worker.)

In contrast, George Akerlof’s model of “partial gift exchange” (1980), which reflects the view that employees will respond to superior treatment and terms and conditions of employment with superior performance, is much more in accord with the prevailing HRM approach to personnel relations in the areas of discipline and pay. This theory also features the payment of high “efficiency” wages but more as a positive inducement to above-average levels of productivity by people predisposed to honor their implicit commitments than as a disincentive required to avoid shirking and loafing by those who are economic criminals at heart. Another distinguishing characteristic of this theory—and other sociology-based theories— which would probably be regarded as analytically more significant than its positive orientation, is that the worker is held to be motivated by norms of fairness which are formed and held in the workplace group. It cannot be classified among theories of individual utility maximization; in this respect, it is more in the institutionalist tradition of Phelps Brown and others, which emphasizes the operation of conventional forces. It invites explicit consideration of the influence of employee group behavior on wage determination and the development of human resource management in the United States.

IV HUMAN RESOURCE MANAGEMENT AND CONCERTED BEHAVIOR; LESSONS FROM U.S. EXPERIENCE

Group norms, when threatened or actually violated, bring group reaction—emotional, strategic, and tactical. The emotion is a shared sense of injustice and betrayal, and it drives strategy and tactics. The strategy is to prevent or reverse what is perceived as a disequilibrium, rather than to adjust, as uncoordinated individuals are assumed to do, to a new and less favorable equilibrium by a lower expenditure of effort and diligence. And the tactics are collective action, whether taking the form of instrumental slowdown or other types of restriction of output or of work stoppages.

Management's defenses against concerted withdrawals of efficiency are likely to be weaker than its ability to punish and prevent individualistic shirking. The costs of plant slowdowns or shutdowns are lumpy and (by definition) can't be reduced by reassignment or increased overtime for nonshirkers within the plant. The costs of implementing more comprehensive strategies of resistance or deterrence to concerted employee behavior are likely to be high. The costs of identifying "ringleaders" when everyone is shirking, under a strategy of exemplary punishment, is presumably greater than the costs of detecting isolated acts of shirking. Moreover, the cost-effectiveness of exemplary discipline is presumably lower in the former case, where it would vary inversely with the cohesiveness of the employee group—which, in turn, would depend on the magnitude of the prospective or actual wage cut (or other departure from custom and practice) and with the intensity of employee emotion thereby provoked. Alternatively, if the strategy should be one of prevention from the outset, management must confront marginal costs of recruitment and selection in blacklisting and otherwise weeding out reported "troublemakers"—an activity which has been illegal in the United States since the passage of the Wagner Act in 1935. Or, in extreme (but by no means purely hypothetical) cases, management might have to contemplate the costs of abandoning the site and investing in new (or expanded) operations in a "safe" part of the country (notably the southern states in the United States) where workers as well as fields are green.

If the costs of employer resistance are high, however, so are the "gains" from prevention. These gains of course include resumption of normal operations and levels of productivity if the group activity subsides. But the employer might have to assess the probability that the informal group would jell or merge into a full-fledged trade union which, as a "continuous association," is better able "to exact higher wages or more favorable working conditions" (Slichter, 1920, p. 38, n.1)—especially if it can establish or coordinate collective bargaining over a wide competitive area. This means that unionism and collective bargaining might be endogenously generated in response to adverse market conditions in a form essentially similar to that traced out by Commons. The process is not irresistible; but the employer must decide whether (a) the cost of preventing unionism is less than the cost of operating under collective bargaining into the indefinite future, and if so, (b) whether the costs of prevention are greater or less than the cost of preemption by payment of an efficiency wage.

Hicks, in 1932, wrote that “even in a market where labour is still unorganised, the principal check ... on the action of employers is generally their fear that reductions will stimulate combined resistance.” (1964 ed., p. 137) In fact, a major share of the responsibility for the development of human resource management in the United States can be traced to that fear. It was not the only source, nor was preemptive policy the only strategy. Labor turnover and labor investment, which were greatly increased by labor shortages during the period of the First World War, became problems with which the old foreman-directed drive system was unable to cope and increased employer interest in gaining “good will” and loyalty from their employees through establishment of shop committees for the airing and resolution of grievances and the provision of employee life, health, and disability insurances. However, the effectiveness of the pecuniary benefit programs in reducing turnover became subject to question, especially since turnover was heavily concentrated among young and junior employees (Jacoby, 1985, pp. 115-126, 197). And labor shortage (which was the cause of the great increase in turnover) ended abruptly with the severe depression of 1920-21. During that depression, firms came under great pressure to reduce wage rates as wholesale prices declined sharply—by 37 per cent between 1920 and 1921, or 26 per cent over the longer period 1920-22. Yet hourly earnings declined by only 12 per cent and 16 per cent, respectively, so that real earnings rose by 25 per cent in 1920-21 and 10 per cent in 1920-22.³ Slichter (1929) interpreted this relatively sluggish wage behavior as evidence of continued effort by employers to secure employee good will due to a “dread of labor trouble” brought on by the great gains in union membership that had been registered during the war years.

These gains, together with dramatic postwar strike activity and fear of labor radicalization, provoked successful policies of all-out resistance on the part of many recently organized employers, who participated in “open shop” campaigns to root unions out. At the same time, employers developed positively oriented (nondisciplinary) policies, which Slichter interpreted as complementary to their efficiency wage policies: “Having failed to reduce wages in proportion to the fall in prices, employers were compelled to make their men more efficient.” (Slichter, 1929, p. 189; also 1928, pp. 634-5). Efficiency was to be increased by the implementation of a bundle of policies which were designed primarily to enhance “morale” by providing opportunity for “the best men” to gain advancement on the basis of “merit”, and, by providing “security—steady work, protection against arbitrary discharge, a pension in old age, and, in some cases, insurance against sickness”, for “the average and subaverage man, the wage earner who cannot expect to advance by unusual skill, knowledge, or exertion, and who is not likely to be interested in group action.” (*Ibid*, p. 211).

It is interesting to note that the direction of causation in this historical interpretation is the reverse of that followed in contemporary human capital theory. According to the latter, prior decisions to invest in firm-specific human capital “cause” wages to be high enough to restrict quit rates to investment-protective levels. According

³In contrast, real earnings had risen by only 2 per cent in the severe depression of 1892-94 (when the declines in hourly earnings and wholesale prices were only 2 per cent and 4 per cent, respectively. The source on hourly earnings in both downswings is A.O'Brien (1983).

to Slichter's efficiency approach, high wages result instead from employer fear of concerted employee behavior and can therefore "cause" human capital investment by the firm. They can serve as an incentive to invest in order to offset the high wages with higher efficiency; and they can provide an opportunity to invest by their effect on quit rates. And finally, quit rates that are lowered in this manner could make it profitable for the firm to invest in general as well as specific human capital.

In the 1920s, union membership and strike activity in the United States fell to very low levels in an economic climate characterized by stable employment and low quit rates and also by high rates of increase in real wages and productivity. It was agreed that levels of unemployment were sufficiently high to account for the decline in organization, although it does not appear to have been suggested that the unemployment resulted (*int. al.*) from efficiency wage policies (as a shirking model might suggest). Moreover, as Jacoby (1985) pointed out, the growth of personnel departments slowed down in the Twenties, many "employee representation plans" were eliminated after the union scare had subsided, and, in many companies personnel policies were decentralized and returned to control by line management. Thus, this period suggests that, absent a serious threat of concerted behavior, interest in "enlightened" personnel policies ebbed, which might be inferred from the old efficiency wage theory.

The Thirties revealed that even unprecedentedly severe unemployment was not proof against labor unrest and, ultimately, a surge of unionism which resulted in extensive organization of the manufacturing sector. In fact, the unrest was a reaction to the severity of the Great Depression. The union organization was in part a response to the breaking of many implicit contracts, after early attempts by employers to hold the line on wages, to stabilize employment, and ration work by seniority or work-sharing and after pension plans and other benefit programs suffered a similar fate. The breaking of implicit contracts (albeit under duress) generated worker demands for the explicit variety—and for organizations which could negotiate and enforce them.

Union organization and the establishment of collective bargaining arrangements were made possible by supportive legislation, especially the Wagner National Labor Relations Act of 1935. That landmark piece of legislation greatly reduced the cost of organization to unionists by requiring only that they win government-held elections to be "certified," that employers refrain from using time-honored methods to frustrate organizing efforts, and that they bargain with duly certified unions in "good faith." The new legal restraints were aimed at activities in which the autonomous line foreman had long excelled; and the obligation to bargain was conducive to the reestablishment of centralized, professional, and nonideological labor management. The latter was required for effective collective bargaining in unionized firms; but it also served as an instrument for the revival of preemptive policies by many nonunion firms in the latter half of the Thirties and the wartime Forties. For much of corporate management, the lessons to be learned—or rather relearned—from the Thirties were, first, that unionism was a function of (or a reflection on) management and, second, that broken implicit contracts should be replaced by new ones -- and taken more seriously.

Passage of the Taft-Hartley amendments to the National Labor Relations Act in 1947 facilitated direct employer resistance to organization by imposing various restraints on union organizing behavior, explicitly guaranteeing management freedom to express its views to its employees (provided that such expression contains “no threat of reprisal or force or promise of benefit”). Yet, the period between the early Fifties and the early Seventies, when union membership density underwent renewed decline, also witnessed a strong development and extension of the HRM stereotype. Postwar preemption abandoned its heavily paternalistic heritage in favor of a “human relations” approach which featured recognition of the dignity and autonomy of the individual worker as a critical determinant of the “quality of life”. The new approach also carried forward the older immunizing strategy of imitating accomplishments in the collective bargaining domain; and these included not only such nonpecuniary benefits as grievance procedures (but without provision for outside arbitration) but also a wide array of pecuniary benefits. Prior to the early Seventies, moreover, nonunion wages, on average, gained on union wage levels.

However, during the ensuing period (in the 1970s and 1980s) of dramatically accelerating decline in union organization and collective bargaining, managerial strategies of direct (i.e., administrative and other noncompensatory) strategies prevailed—not only among nonunion employers but also among many unionized employers seeking either to disestablish or weaken their bargaining partners. “Ability to pay” efficiency wages (and other conditions) was reduced by adverse movements in the terms of trade (which exerted downward pressures on domestic costs and profits), intensified international competition and deregulation in transport and telecommunications, and a general slowdown in productivity growth. Hence, monopoly rents, from which preemptive wage payments are “extracted”, according to William Dickens (1986), were squeezed, as were what might be regarded as rents of growth.

At the same time, two factors contributed to an effective reduction in direct costs of opposition. First, while the Wagner Act continued to proscribe recourse to old-fashioned forms of exemplary discipline, administrative delay in holding various selections and disposing of “unfair practice” charges brought by unions, together with increased recognition of the derisory penalties for commission of unfair practices, reduced employer costs of direct opposition. And second, a secular decline in proletarian militancy (one possible inference from a declining trend in union victories in representation elections) lowered the probability that failures to match negotiated increases in union firms would spur off concerted reaction in nonunion establishments. Instead, union-nonunion differentials widened during this period, frequently placing unionized firms at a competitive disadvantage and inducing them to adopt tougher policies vis-a-vis their unions. Hence the Commons effect was obviously much weaker in the Seventies and the Eighties than it had been in the Thirties, when it prevailed in a much harsher economic and legal environment.

CONCLUSION: THE FADING STEREOTYPE?

Thus far, HRM survived both serious weakening of the market environments from which many internal labor markets have drawn support and (as noted at the outset) major diminution of the threat of collective action. In part, its survival has been required by the firm's need to cope with new and increased governmental regulation—e.g., in the areas of occupational safety and health and employment protection, discrimination, and “affirmative” action (Strauss, 1982). Moreover, the survival of traditional functions of HRM would be predicted by most of the utilitarian models of labor-market behavior which have been discussed above. To the extent that markets become more competitive, theories which posit relatively low wages at some point, whether to offset superior nonwage conditions of employment or employee risk aversion (as the basis for implicit contracts) or to account for employee training or rising age-earnings profiles, gain in explanatory power.

In fact, even high-wage features of HRM have thus far survived, along with the nonwage attributes. Union-nonunion wage differentials have declined generally, but evidence of a “threat effect” in the 1980s was found by Dickens and Katz (1988) in the cross-industry correlation between nonunion wages and union density. Indeed, preemptive wage strategy has been strongly regarded as a complement to, rather than a substitute for, measures of direct opposition to unionism (especially by consultants who advise nonunion employers on maintaining a “union-free environment”). High wages have also been complementary to other preemptive policies. Thus, Foulkes (1980) noted that nonunion firms whose wages exceeded union levels also established more generous welfare payments as a union-preventive strategy. It has been commonly believed that the chief advantage conferred by nonunion operation consists in higher levels of productivity, due primarily to the absence of union restrictive practices. Even so, Foulkes observed nonunion firms assigning as great weight to seniority, relative to individual ability, in promotion cases, as might be required under collective bargaining. Grievance procedures (as noted above) were also copied from the union play books. Reluctance to impose harsh discipline (also noted above in discussing efficiency wage theory) in some firms was ascribed to fear of unionization). “The practices of many companies are such”, concluded Foulkes (*ibid.*, p. 61) “that they refrain from using much of their cherished flexibility”.

It is to be expected that nonunion HRM should lower its guard as union membership recedes to extremely low levels and as bargaining institutions become fragmented and weaker. Yet it is unlikely that the lessons of the Thirties will be lost on the managers of the Nineties, as the lessons of the early post-WWI years had been lost on the managers of the Twenties. Even if organization should asymptotically approach zero, today's managers of human resources would have in the Wagner Act and its continued protection of the right to organize a reminder of the continuing potential of concerted behavior as a response to serious frustration of worker aspirations. And the latter, in the United States as in virtually every other place, have been characterized by a persistent secular rise in demand for “industrial democracy” and improved status—which sooner or later entails concomitant improvement in pay. The depth and universality of this source of concerted behavior find expression in the concluding sentences of

A Century of Pay (Phelps Brown and Browne, 1968):

Industrialisation has created many jobs which do not engage the whole man, and hold his interest, nor do they give him the satisfaction of achievement, and sustain his self-respect and his standing with his fellows. Whatever the job, it usually requires the man who does it to adopt the status of employee, with its loss of independence and its need for self-defence in the wage bargain and in daily relations with the authority of management... The achievement of a better life for the worker will continue to depend, now as before, on the advance of productivity; but more than before this needs to be accompanied by improvements in the design of jobs and the human relations that go with them,...



REFERENCES

- Akerlof, G. and H. Miayaki, "The Implicit Contract Theory of Unemployment Meets the Wage Bill Argument," *Review of Economic Studies*, vol. 1 xlvi, 1980.
- Azariadis, C., "Implicit Contracts and Underemployment Equilibria," *Journal of Political Economy*, vol. lxxxiii, 1975.
- _____, and J.E. Stiglitz, "Implicit Contracts and Fixed-Price Equilibria," *Quarterly Journal of Economics*, vol. xcvi, 1983 Supplement.
- Baily, M.H., "Wages and Employment under Uncertain Demand," *Review of Economic Studies*, vol. xli, 1974.
- Becker, G.S., *Human Capital* (New York: National Bureau of Economic Research, 1964).
- Braverman, H., *Labor and Monopoly Capital* (New York: Monthly Review Press, 1974).
- Brown, Ch., "Equalizing Differences in the Labor Market," *Quarterly Journal of Economics*, vol. xciv, No. 1, February 1980.
- H. Phelps Brown and S.V. Hopkins, *A Perspective of Wages and Prices* (London: Methuen), 1981.
- _____, *The Inequality of Pay* (Oxford: Oxford University Press, 1977).
- _____, *The Economics of Labor* (New Haven: Yale University Press, 1962).
- _____, *The Growth of British Industrial Relations* (London: Macmillan, 1959).
- _____, "Labour Morale," *Economic Journal*, March 1949, pp. 40-55.
- Cain, G., "Segmented Labour Markets," in J. Eatwell, M. Milgate, and P. Newman, eds, *The New Palgrave: A Dictionary of Economics* (London: Macmillan, 1987).
- Commons, "The American Shoemakers, 1648-1895," reprinted from *Quarterly Journal of Economics*, Nov. 1909, in *Labor and Administration*. (New York: Macmillan, 1913), pp. 219-66.
- _____, "American Shoemakers, 1648-1895," Reprinted in J.R. Commons, *Labor and Administration* (New York: Macmillan, 1913).
- Cooter, R.D., "Coase Theorem," in J. Eatwell, M. Hilgate, and P. Newman, eds., *The New Palgrave: A Dictionary of Economics* (London: Macmillan, 1987).
- Dickens, W.T., L.F. Katz, K. Lang, and L.H. Summers, "Employee Crime and the Monitoring Puzzle," *Journal of Labor Economics*, vol. 7, No. 3, July 1988.
- _____, L.F. Katz, "Inter-Industry Wage Differences and Theories of Wage Determination," National Bureau of Economic Research, Working Paper No. 2271, August 1988 (revised).

- _____, "Wages, Employment, and the Threat of Collective Action," National Bureau of Economic Research, Working Paper #1856, March 1986.
- _____, "Assuming the Can Opener: Hedonic Wage Estimates and the Value of Life," National Bureau of Economic Research, Working Paper #3446, September 1990.
- Doeringer, P.B. and M.J. Piore, *Internal Labor Markets and Manpower Analysis* (Lexington, Mass: D.C. Heath, 1971).
- Duesenberry, J.S., *Business Cycles and Economic Growth* (New York: McGraw-Hill, 1958).
- Dunlop, J.T., "The Task of Contemporary Wage Theory," in J.T. Dunlop, ed., *The Theory of Wage Determination* (London: Macmillan, 1957).
- Elbaum, B., "The Making and Shaping of Job and Pay Structures in the Iron and Steel Industry," in P. Osterman, ed., *Internal Labor Markets* (Cambridge, Mass., MIT Press, 1984).
- Feller, D.E., "A General Theory of the Collective Bargaining Agreement," *California Law Review*, 61, May 1973.
- Foulkes, F.K., *Personnel Policies in Large Nonunion Companies* (Engelwood Cliffs, New Jersey: Prentice-Hall, 1980).
- Freeman, R.B. and J.L. Medoff, *What Do Unions Do?* (New York: Basic Books, 1984).
- Green, J. and C.M. Kahn, "Wage-Employment Contracts," *Quarterly Journal of Economics*, vol. xcvi, 1983, Supplement.
- Groshen, E.L., "Source of Intra-Industry Wage Dispersion: How Much Do Employment Matter?" *Quarterly Journal of Economics*, vol. 106, August 1991.
- _____, "Five Reasons Why Wages Vary Among Employers," *Industrial Relations*, vol. 30, no. 3, Fall 1991.
- Grossman, S.J. and O.D. Hart, "Implicit Contracts under Asymmetric Information," *Quarterly Journal of Economics*, vol. xcvi, 1983 Supplement.
- Hicks, J.R., *The Theory of Wages* (London: Macmillan, 1964, 2nd ed.).
- Hildebrand, G.H. and G.E. Delehantry, "Wage Levels and Differentials," in R.A. Gordon and M.S. Gordon, eds., *Prosperity and Unemployment* (New York: Wiley, 1966).
- Jacoby, S.M., *Employing Bureaucracy* (New York: Columbia University Press), 1985.
- Kerr, C., "The Balkanization of Labor Markets," in E.W. Bakke, et. al, *Labor Mobility and Economic Opportunity* (Cambridge, Mass: MIT, 1954), pp. 92-110.
- _____, "Wage Relationships-The Comparative Impact of Market and Power Forces," in J.T. Dunlop (ed.), *The Theory of Wage Determination* (London: Macmillan, 1957).
- Keynes, J.M., "The Economic Consequences of Mr. Churchill," (1925). Reprinted in *Essays in Persuasion* (New York: Norton, 1963)
- _____, *The General Theory of Employment, Interest, and Money* (New York: Harcourt Brace, 1936)

- Kochan, T.A. and P. Cappelli, "The Transformation of the Industrial Relations and Personnel Functions," in P. Osterman, ed., *International Labor Markets* (Cambridge Mass., MIT Press, 1984).
- _____, H.C. Katz, and R.B. McKersie, *The Transformation of American Industrial Relations* (New York: Basic Books, 1986).
- Krueger, A.B. and L.H. Summers, "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, vol. 56, March 1988, pp. 259-94.
- Kwoka, J.E. Jr., "Monopoly, Plant, and Union Effects on Worker Wages," *Industrial and Labor Relations Review*, vol. 36, No. 2, January 1983.
- Lawler, E., *Pay and Organizational Effectiveness: A Psychological View* (New York: McGraw Hill, 1971).
- Lazear, E.P., "Pensions and Deferred Benefits as Strategic Compensation," *Industrial Relations*, vol. 29, No. 2, Spring 1990.
- _____, "Why Is There Mandatory Retirement?" *Journal of Political Economy*, vol. 87, No. 61, December 1979.
- _____, "Agency, Earnings Profiles, Productivity, and Hours Restrictions," *Journal of Political Economy*, vol. 71, No. 4, September 1981.
- Leontief, W., "The Pure Theory of the Guaranteed Annual Wage Contract," *Journal of Political Economy*, vol. liv, February 1946.
- Lester, R.A., *Adjustments to Labor Shortages* (Princeton: Industrial Relations Section, 1955).
- Long, J.E. and A.N. Link, "The Impact of Market Structure on Wages, Fringe Benefits, and Turnover," *Industrial and Labor Relations Review*, vol. 36, No. 2, Jan. 1983.
- Medoff, J.L. and K.G. Abraham, "Experience, Performance, and Earnings," *Quarterly Journal of Economics*, vol. xcv, December 1980.
- Mitchell, D.J.B., *Human Resource Management: An Economic Approach* (Boston: PWS-Kent, 1989).
- O'Brien, A., "The Cyclical Sensitivity of Wages," *American Economic Review*, vol. 75, No. 5 (December 1985).
- _____, "Labor as a Quasi-Fixed Factor," *Journal of Political Economy*, vol. 70, No. 6, December 1962.
- Osterman, P., ed., *Internal Labor Markets* (Cambridge, Mass., MIT Press, 1984).
- Parsons, D.O., "Specific Human Capital: An Application to Quit Rates and Layoff Rates," *Journal of Political Economy*, vol. 80, no. 6, Nov-Dec 1972, pp.1120-43.
- Pencavel, J., *An Analysis of the Quit Rate in American Manufacturing Industry* (Princeton, N.J.: Industrial Relations Section Princeton, 1970).
- Figors, P. and C.A. Myers, *Personnel Administration* (McGraw Hill, 1977 ed.).
- Pugel, T.A., "Profitability, Concentration, and the Interindustry Variation in Wages," *Review of Economics and Statistics*, May 1980, pg. 308-314.

- Reder, M.W., "Chicago School" in J. Eatwell, et al., eds., *The New Palgrave: A Dictionary of Economics* (London: Macmillan, 1987, vol. 1, pp. 413-418).
- Rosen, S., "Human Capital," in J. Eatwell et al., eds., *The New Palgrave: A Dictionary of Economics* (London: Macmillan, 1987).
- _____, "The Theory of Equalizing Differences," in O. Ashenfelter and R. Layard, *Handbook of Labor Economics*, vol. 1 (Elsevier Science, 1986).
- Schultz, T., "Investment in Human Capital," *American Economic Review*, vol. 51, March 1961.
- Slichter, S.H., "Notes on the Structure of Wages," *Review of Economics and Statistics*, February 1950, Reprinted in *Potentials of the American Economy* (Cambridge, Mass: Harvard University Press, 1091).
- _____, "The Current Labor Policies of American Industries," *Quarterly Journal of Economics*, May 1929. Reprinted in *Potential of the American Economy* (Cambridge, Mass: Howard University Press, 1961).
- Smith, A., *The Wealth of Nations* (New York: Modern Library ed., 1937).
- Strauss, G., "The Transformation of Industrial Relations: Toward the Study of Human Resources Policy," Business and Public Policy Working Paper no. OBIR-29, September 1988.
- _____, "Personnel Management: Prospect for the Future," in K. Rowland and G. Ferris, eds., *Personnel Management* (Boston: Allyn and Brown, 1982).
- Ulman, L., "Labor Mobility and the Industrial Wage Structure in the Postwar United States," *Quarterly Journal of Economics*, lxxxix, no.1, Feb. 1965, pp.98-105.
- _____, "Labor Market Analysis and Concerted Behavior," *Industrial Relations*, vol. 29, no. 2, Spring 1990.
- Vroman, W., "Cyclical Earnings Changes of Low Wage Workers," in R. Ehrenberg, ed., *Research in Labor Economics*, vol. 2 (Greenwich, Conn. JAI Press, 1978).
- Williamson, O.E., *The Economic Institutions of Capitalism* (New York: The Free Press, 1985).
- Yellen, J.L., "Efficiency Wage Models of Unemployment," *American Economic Review: Papers and Proceedings*, vol. 74, no. 2, 1984.