Choice of Employment Systems in Internal Labor Markets

Although many empirical studies of internal labor markets have appeared in the past decade, our theoretical understanding remains weak. Doeringer and Piore’s (1971) model is an uneasy amalgam of human capital theory, with its emphasis on specific training generating attachment between firms and workers, and a sociological argument about the development of group norms and rules that regulate worker behavior. Doeringer and Piore also implicitly emphasize the activities of unions in encouraging seniority-based promotion and pay systems (for more details on this point, see Jacoby [1984]). We characterize this combination of themes as “uneasy” because it rests on arguments with very different behavioral foundations. Human capital theory assumes maximizing individualistic behavior, but neither the sociological argument nor many of the union stories make this assumption. Doeringer and Piore never make clear how these different arguments come together, and as a consequence, their “theory” is more a list of considerations than a coherent explanation.

The arguments developed by Williamson et al. (1975) are more consistent. They contend that in large enterprises perpetual monitoring of employee activity is difficult and costly, and it is hard to specify in advance the full range of worker tasks and the desired responses to particular situations. In addition, workers who become proficient at specific tasks may be tempted to take advantage of their monopoly position to extract special rewards from the firm. Internal labor markets establish incentives which reduce both the need for monitoring and the gain from personally advantageous but globally inefficient behavior.

This efficiency-based theory clarifies the advantages of internal labor markets, but it does not provide a satisfactory general explanation of their emergence.

* Associate Professor of Economics, Boston University, and Visiting Associate Professor, Sloan School of Management, MIT.
1Kerr (1954) presents what is probably the earliest systematic statement of the ideas behind internal labor markets. More recent papers on the subject are collected in Berg (1981) and Osterman (1984).

nor does it explain how internal labor markets change over time. Taken as a whole, Williamson et al.’s (1975) approach borders on tautology. It is equally easy to establish efficiency-limiting aspects of internal labor markets: for example, seniority provisions and/or eligibility limitations may prevent the most able workers from attaining jobs best suited to their talents. These kinds of barriers can be costly. Proponents of efficiency-based explanations must assume as a matter of faith (since no data on costs and benefits are available) that the arrangement which prevails is by definition the most efficient. Furthermore, these arguments provide little insight into the historical record: Why did internal labor markets emerge when they did, what explains their differential rate of diffusion across firms, and how and why do they change today?  

The goal of this paper is to contribute to the theoretical literature on internal labor markets. We examine contemporaneous alterations in the structure of these markets, on the premise that this variation will help clarify the essential nature of internal labor markets. For analytical purposes, we categorize these alternative occupational arrangements as “employment subsystems” and identify four patterns. We assume that the types of subsystems in place result from the firms’ decision-making processes, and we therefore try to determine the factors underlying those decisions.

The origins of this paper lie in interviews with a large number of firms and examples are drawn from these interviews (as well as from other sources) to illustrate specific points. Despite this empirical foundation, however, the paper is primarily theoretical and hypothetical; the ideas presented here are intended to stimulate more analysis and additional systematic empirical research.

Employment Subsystems

Ten years ago, virtually all computer programmers learned their trade either in school or in private training institutions. Because their knowledge was general (i.e., applicable in many different settings) and because of the strong demand for computer personnel, programmers could change jobs easily. Programmers’ loyalties were to their profession, not to a specific firm. These arrangements had some advantages for firms, as well. There was no need to invest in training, and at least in principle, firms could easily lay off their computer professionals if demand fell. Over time, however, this pattern proved increasingly costly. High turnover disrupted key operations. Reliance on the external market in a high demand environment also led to sharp wage increases; these in turn threatened to disrupt internal wage patterns. Gradually,

---

\footnote{For comments on earlier versions of this paper, I am grateful to Robert McKersie; seminar participants at Boston University, Cornell University, MIT, and the Wharton School; and two referees.}
companies began to redesign the computer programmer career pattern. Tests were administered to employees elsewhere in the organization — often secretaries and clericals — and those with good aptitude were given programming training. This training was truncated compared to its counterpart in external institutions. It was also more company-specific in its applications. Nevertheless, for a wide range of activities, the new programmers were quite adequate and the narrow company-specific character of their training limited their ability to "jump ship." More complex programming tasks remained organized along earlier lines. A second (and more frequently noted) example of the transformation of employment arrangements is the growing use of temporary help services. This device transfers the personnel rules and internal labor market procedures, but not the work itself, to another firm.

These two examples illustrate several points. They demonstrate that similar work can be performed under quite different hiring, training, promotion, and turnover patterns. And the mix of these patterns in the labor market can change, either when a specific occupation is transformed (as in the case of programmers) or when a given pattern gains at the expense of others (as in the case of temporary help). These examples also suggest that the evolution of work can be better understood by identifying some of these patterns and assessing the circumstances under which each becomes established. We denote the fact that there are several alternatives and that each may co-exist within a given firm by calling these patterns employment subsystems.

Most firms have numerous rules and procedures, and lengthy personnel textbooks and professional journals are devoted to developing and elaborating them. At the risk of over-simplification, it is useful to distinguish four categories which, taken together, define the internal labor market for a set of occupations.

1. **Job classification and job definition.** These are rules which determine whether jobs are defined broadly or narrowly (i.e., include many or few tasks) and whether job definitions are rigid or loose (i.e., whether a person doing a job does only those tasks or whether he or she will do other work).

2. **Deployment.** This refers to rules concerning how employees may be moved from job to job in the organization. For example, in some settings seniority determines bidding rights, while in other settings management retains complete discretion over how to deploy labor.

3. **Security.** Some firms operate with explicit or implicit promises of lifetime job security; other companies make no promises beyond payment for the current day's work. In some settings layoffs are determined by (reverse) seniority and elaborate job bumping rules, while elsewhere management is free to release whom it chooses.
4. Wage rules. The major distinction here is whether wages are attached to jobs or to individuals. In the former case, all individuals in a given job classification receive a given wage; in the latter circumstance, some combination of personal attributes (skill, education, performance) and seniority determines wages regardless of what set of tasks currently is assigned to the individual.

Within a given subsystem, these rules fit together with a coherent logic. For example, companies with lifetime job security guarantees are likely to have fewer and less restrictive job classifications than firms in which hire/fire is the rule. The logic is that rigid job classifications are another form of security or protection. In an environment in which security is provided through other mechanisms, strike classifications are less necessary. A comparable point might be made concerning the relationship between rules concerning deployment and security. Similarly, in circumstances in which wages are attached to individual jobs, classification will be less rigid than where wages are determined largely by job assignment.

The four employment subsystems which capture most of the choices available to firms may be characterized as industrial, salaried, craft, and secondary. As applied here, these familiar terms do not correspond exactly to their conventional usages. The characteristics of these categories and how they differ from each other are described in detail elsewhere (Osterman, 1982); they will only be sketched here.

The industrial model represents the manner of organizing blue-collar work which became the norm as a result of the unionization drives of the Great Depression and which was solidified in the era of postwar prosperity. In this model, work is organized into a series of tightly defined jobs with clear work rules and responsibilities attached to each classification. Wages are attached to jobs and hence an individual’s wage is determined by his or her classification. Management’s freedom to move individuals from one job to another can vary from situation to situation, but the typical case is that both promotions and lateral shifts are limited by seniority provisions and by requirements that workers agree to the shifts. Finally, there is no formal job security, and it is understood that management is free to vary the size of the labor force as it wishes. When layoffs do occur, they are generally organized according to reverse seniority.

Although the structure of this model emerged from the spread of unionism, it should not be construed as limited to such situations. Because of fear of unions, government pressures for uniformity, and imitation, the model spread throughout the economy. A recent survey of nonunion firms found that seniority-based promotion and layoff systems were extremely common even in the absence of formal contracts (Abraham and Medoff, 1984).
This model has a strong internal logic. Because wages are attached to jobs, the jobs must be carefully defined so that there is common understanding concerning who is doing what work, and thus who is entitled to what wage. Similarly, while the system provides no overall job security (management can vary the size of the workforce at will), individual security is based on a bumping system grounded in seniority; for that system to be effective, careful job classifications are necessary.

Most labor economics and industrial relations research has emphasized blue-collar work. Consequently, it is more difficult to describe the salaried internal labor market model. Understanding the model is important for three reasons: it describes the employment pattern of large numbers of workers; it extends beyond salaried work to a number of innovative blue-collar employment settings; and some of its characteristics represent the general direction in which management is trying to reorganize work.

The salaried model combines more flexible and personalistic administrative procedures with greater commitment to employment security. Although individuals have job descriptions, much as industrial employees have work rules, these descriptions are subject to revision by superiors and the employees are prepared to take on new activities as demanded. By the same token, the clearly defined job ladders and promotion sequences which characterize industrial settings are generally absent (Kanter [1978] notes that salaried workers rarely have a clear sense of the paths into and out of their jobs). The flexible career lines and job descriptions are consistent with another aspect of this employment system — the greater role of personalistic considerations in wage setting. There is a considerably greater scope for merit considerations in pay setting and the wages of two individuals in the same job can vary considerably. Put differently, the pay system of industrial settings, in which the dominant consideration is job assignment, is far less prevalent in the salaried model.

If rigid job classifications and reliance on nonpersonalistic procedures are the key to job security and worker acquiescence in the industrial system, what plays a comparable role in the salaried model? One long-time theme in the literature, first expressed by Selig Perlman, is that white- and blue-collar employees view work differently. Whereas a sense of limited opportunity and collective consciousness is central to the blue-collar mind-set, the white-collar worker is an individualist who believes that his or her hard work will pay off and that no artificial limits prevent individual success. In this view, then, the salaried employment system works because its flexibility and fluidity is not viewed as a threat.

Surely many employees in the salaried system share this view and are therefore willing to forego the protections inherent in the industrial model.
This cannot be the complete explanation, however, for the important reason that the salaried system is not limited to white-collar work. A number of highly visible employers apply the salaried model to blue- as well as white-collar work. Moreover, many white-collar workers are interested in security and protection from arbitrary administrative actions.

What closes the salaried model is employment security. In the classic salaried model, individuals, once they pass a probationary period, can expect lifetime employment with the firm. Unlike the industrial model, in which it is explicitly understood that the firm will adjust the size of the labor force in response to product market conditions or technological change, the implicit promise in the salaried system is that layoffs either will not occur or that the firm will make every reasonable effort to avoid them. Note that absolute promises are not necessary; if they were, the scope of the salaried model would be extremely limited. What is crucial is that employees are sufficiently convinced of the sincerity of the firm's commitment to employment stabilization that they are willing to provide a great degree of flexibility to the firm.

The salaried model characterizes much white-collar work. The career patterns of most managers and many professionals who work in bureaucracies are accurately captured by the model. However, the salaried model is not limited to describing white-collar work. A few American firms have used it as well for their blue-collar workers. This was true in the twenties (the exception was termed "welfare capitalism") and remains true for a small but important number of innovative firms today. In return for flexible work rules and a willingness to accept managerial prerogatives with respect to deployment, these firms offer their workers employment security.

The industrial and the salaried models represent the dominant models for organizing work in the core firms. The careers of a small number of employees take a different form. Under both the industrial and salaried models, employees make their careers within the firm. The firm is responsible for training, and external entry to job ladders generally occurs at only a few points. In contrast, the distinctive characteristic of both the craft and the secondary models is the presumption of considerable interfirm movement.

Craft subsystems. These are characterized by considerably greater mobility and more loyalty to the skill or profession than to the firm. Craft training typically occurs outside the firm, in schools, formal training programs, or apprenticeship programs. The skills acquired are not firm-specific. Workers thus have more market power than under industrial and salaried arrangements. As a result, these jobs are not embedded in lengthy job ladders, and mobility, far from being penalized, is often rewarded. Several white-collar jobs within firms operate under craft subsystems. In addition to the programmer example
cited earlier, in some firms senior sales people, some of whom are on payroll and others an incentive system, also may be characterized as craft workers.

Secondary subsystems. This final pattern contains jobs which lack career prospects, either within the firm or via interfirm movement. They tend to be (but are not always) low skilled and poorly paid. These jobs are more accurately conceptualized as lacking clear linkages to future jobs. Examples in so-called primary firms include many clerical workers, mailroom staff, and messengers.

This fourfold classification is far from the first effort to understand subdivisions or categories of employment relationships. It does, however, differ in important respects from previous approaches.

In dual labor market models, it is conventional to speak of "primary firms" and "secondary firms," with the implication that the bulk of a firm's employment falls into one or the other category. In fact, the most widely cited explanation for the emergence of dualism — Piore's argument that the product market is divided into stable and unstable shares, with primary firms capturing the stable portion and the flux and uncertainty of the remainder passed on to secondary firms (Piore, 1980) — rests on the view that primary and secondary firms are distinguishable on dimensions beyond their internal labor systems. The subsystem perspective developed here differs from dual labor market theory because firms which by any standard would be considered "primary" — large banks and insurance companies for example — contain substantial numbers of employees who labor under secondary subsystems.

In a variant of dual labor market theory, Gordon, Edwards, and Reich (1982) argue that the primary segment of the labor market can be subdivided into "independent" and "subordinate" segments. However, the axis upon which they make these distinctions is control: worker versus employer control over skills and mobility. The assumption is that this is the major consideration for both parties. This paper recognizes control as one factor which shapes internal labor markets, but it also argues that other important costs and benefits pertain, and that certain constraints limit the possible range of choices.

One key constraining factor is technology. William Form (1976) argues that operative workers enjoy less physical and social freedom than craft workers, who move about the factory and maintain equipment. Consequently, craft workers are more likely to form work groups and to become politically active. Piore and Sabel (1985) also emphasize the importance of technology. In their argument, multi-skilled craft workers are required for flexible-specialization, while unskilled operators dominate mass production. Such consequences of alternative technologies are incorporated into the model developed below.
The focus here, however, is on career patterns and the location of training rather than on the nature of the skill. For example, multi-dimensional, craft-like skills can be developed by the firm and employed in stable internal employment arrangements (as in the German "dual" training model); or, instead, these skills can be generated externally and involve considerable interfirm mobility (as in the computer programmer example). How are these choices made?

Finally, the argument developed here draws on the organizational behavior literature concerned with the internal structure of organizations; specifically, how internal span of control and the size and number of hierarchies are influenced by variables such as technology, market structure, and environmental uncertainty (see Caves, 1980; Perrow, 1979). However, the organizational behavior perspective is more macro than ours; and it typically characterizes organizations broadly (e.g., number of divisions), although several different employment subsystems may exist (or coexist) in varying proportions within a division. Moreover, the organizational behavior literature pays relatively little attention to labor market variables as determinants of organizational form; nor does it sufficiently assess the impact of alternative forms upon careers.

Firms' Goals and Their Choice of Employment Systems

What does it mean to argue that firms choose in a conscious manner among different alternatives for organizing work? The notion of choice does not imply that firms continually review and up-date the internal labor market structure for the full range of relevant occupations. Such constant activity is clearly beyond the resources of any company and would almost certainly be counter-productive. What we have in mind, instead, is a process similar to that described by Nelson and Winter (1982): A given internal employment and training system is not examined or reconsidered until some event — technological change, changing product markets, reduced supply of an appropriate labor force, sharply rising wages, etc. — forces management to attend to the issue. At this point, the options are examined and a choice is made.

In recent years, firms have enlarged the resources needed to make these kinds of choices. In particular, the size and power of the personnel staff has grown (Jaeger, 1977; Kochan and Cappelli, 1984). An examination of the personnel planning process undertaken by specific firms supports the notion of planned employment systems. One major Massachusetts manufacturing

---

3This example is drawn from interviews conducted by the author.
company (not in the high technology field) requires business and personnel managers to conduct an annual human resources planning process. Each of the firm's divisions is classified into one of three product-market categories (high growth, mature, and "cash-cows"), and different expectations are set for each group. For example, high-growth divisions are less concerned with cost minimization, while this is a major concern for the cash-cows. The managers complete an extensive questionnaire drawing out the personnel implications of new products, new work technologies (e.g., word processing), the demographics of the labor force (e.g., the age distribution of employees), and of specific personnel issues (e.g., rising wages or high turnover). On the basis of this analysis and the appropriate product market category, managers are expected to formulate an annual human resources plan. Planning as sophisticated as this is not universal, but firms do indeed consciously structure their labor forces.

Goals

Our analysis of how firms choose employment systems will be structured around a discussion of the goals of firms and the constraints they face in achieving those goals. Firms have three aims concerning their employment systems. These are: cost effectiveness, flexibility, and predictability. Each is discussed separately below.

_Cost effectiveness_. This is the traditional objective which economists assign to profit-maximizing firms. If the capital or operating technology is taken to be fixed, then the firm seeks to minimize the wage (and benefits) bill for any given level of output. When the firm considers new investment, it chooses the optimal mix of capital and labor by comparing the costs of incremental units of each factor with their marginal contribution to output.

In the context of the problem we have posed — choosing among alternative employment systems — the firm can be characterized as estimating a potential wage bill for each arrangement (either in the short run with technology fixed or in the long run with technology variable) and then choosing the system with the lowest bill for a given set of technological choices and labor market conditions. For example, in a labor market with a limited supply of secretaries (and hence high wages) a firm may choose to create career ladders for this position and capitalize each secretary with distributed work stations. In a labor market with a cheap and plentiful supply of people willing to perform secretarial tasks, the work is likely to be organized along the lines of a pool, which in turn encourages higher turnover and less internal mobility.
This stylized description is reasonable for many purposes. Doubtless firms wish to spend as little on wages and benefits as is consistent with acceptable levels of output and quality, but it must also be admitted that standard theory provides little guidance concerning the nuances of implementing the goal of cost minimization. In particular, this goal is not always best achieved by choosing the employment system that minimizes the wage (and benefits) bill.

The concept of cost must be broadened to include potential as well as actual costs. Employees can impose costs on the firm through errors of various kinds. For example, a particular kind of capital equipment may be simple to operate and require little skill and yet be very expensive and subject to damage through employee error. Many firms will choose to employ higherskill labor and create stable employment systems, not because the day-to-day task demands skill, but because of potential downside costs.

An additional complication for the cost-minimization goal emerges from recent models of firms’ wage-setting behavior. Several models, collectively known as “efficiency wage theories” (Yellen, 1984), have formalized long-held views that wages are positively correlated with productivity. Holding ability constant, well-paid workers are more productive because they stand to lose more if they leave the firm and because their morale and commitment is higher. Similar arguments can be made about subsystems which promise job security and stability. From this perspective, cost-minimization must be weighed against the gains from more expensive but more productive employment systems.

Maximizing predictability. This goal is less often recognized, yet it too constitutes one of the three central objectives firms consider when designing employment systems. As used here, the term “predictability” means that firms are able to plan confidently upon the availability of a qualified labor supply at foreseeable prices. In the textbook world of neoclassical economics, firms are small and are always confident of their ability to purchase all of the inputs they need at the going market price. In reality, the modern firm operates as a planning bureaucracy with a fairly long time horizon. In order to plan output and set prices in advance, it must be confident of assured supplies and knowable prices. In this context, it matters less if the price of labor is high or low as long as it is known with some assurance. In recent years, both shortages and unpredictable wage swings for computer programmers, certain engineering fields, and some blue-collar skills (e.g., welding) have disrupted production and planning in many companies.

Achieving predictability places a premium upon internal ladders and inhouse training schemes, for supplies and prices are then more subject to the
firm’s control. James Thompson (1967) makes a similar argument. He contends that minimizing uncertainty is a key goal of organizations in general and that one strategy for achieving this goal involves a “seal-off” of the central productive functions of the enterprise.

A firm is interested in more than simply the predictable presence of a certain number of bodies with certain characteristics. It is important that the work force produce or deliver a reliable and predictable amount of labor. In addition to the efficiency wage argument that higher wages produce more work, there is also an enormous literature on the topic of worker satisfaction, the quality of working life, and productivity (see Schein, 1978). Another way of making this point is to return to the Gordon-Edwards-Reich argument concerning control and to note that the industrial and salaried patterns provide more control than does the craft subsystem. Our model’s emphasis on the firm’s other objectives (cost minimization and flexibility), however, makes it significantly different from Gordon, Edwards, and Reich’s formulation.

A sophisticated variant of this argument rests on Herbert Simon’s (1947) assertion that one of the central goals of organizations is to shape people’s habits of thinking in ways that lead them to “intuitively” reach decisions in the organization’s interest. According to Simon, organizations supply employees with “premises,” a term which denotes their information, frame of reference, identification with the goals of the organization, and motivation.

Organizations employ a variety of techniques to establish the proper “premises” in the minds of their employees. Among these are training, education, sanctions, advice, and information. The trade-offs among these techniques depend upon the nature of the job under discussion (e.g., its centrality to the organization or the amount of discretion involved), the character of the labor supply, and the relative effectiveness and costs of the different techniques. Thus, for example, a firm is more likely to be concerned with inducing the proper “premises” among managers than among equally skilled and educated but less organizationally central computer analysts. As Simon (1947, p. 228) has noted:

... the problems of organization cannot be considered apart from the specifications of the employees who are to fill the positions established by the organization. The whole subject of job classification needs to be brought into much closer coordination with the theory of organization. The optimum organizational structure is a variable, depending upon the form of staffing of the agency. Conversely, the classification of a position is a variable, depending upon the degree of centralization or decentralization which is desired or anticipated in the operation of the organizational form.

Returning to the example of the employment system of managers compared to computer professionals, this logic provides some insight into the long managerial ladders compared to the short, high turnover computer employment
systems. Long career ladders are more likely for managers because this arrangement permits extended socialization and the creation of the proper “premises.” These premises are less important for the equally skilled but less central computer staff. Similarly, we can understand why in-house training is more common for managers, while firms are more willing to rely upon external training agencies for technical staff. In-house training includes an important element of premise-building or socialization and this is much harder to control in outside training agencies.

**Flexibility.** The third broad goal of the firm is flexibility. Achieving a flexible product and employment policy has become a rallying cry of the eighties. A common diagnosis of the loss of American markets to other nations, particularly to Japan, has been that American firms are less able to respond to changing circumstances than are our competitors. This general notion has been given a serious intellectual foundation in recent research on “flexible-specialization” (Piore and Sabel, 1985) and labor relations (Kochan, McKersie, and Katz, 1986).

The firm desires flexibility with respect to staffing levels, with respect to the deployment of labor, and with respect to the abilities of the labor force. In principle, labor is a variable cost which firms can adjust in response to changes in product market conditions. However, a consistent goal of unions, of informal organizations of workers in nonunion settings, and of the welfare state has been to limit the risks borne by workers. A wide range of provisions in union contracts and legislative protections limit the ability of firms to reduce staffing.

Rigidity in staffing levels is generally costly to firms. If revenue falls but labor costs cannot be reduced commensurately, cost-price margins are reduced and the firm faces losses in profits and markets. Of course, complete flexibility in hiring and layoffs is not necessarily desirable. When firms have made costly training investments in particular individuals, it may pay to engage in some labor hoarding over the business cycle, rather than risk these workers finding jobs elsewhere (Fay and Medoff, 1985). This particular trade-off is important to the model because it conditions choices regarding employment systems which differ in their flexibility.

The firm's concern with flexibility extends also to its ability to deploy labor in what it regards as the most productive manner. In industrial subsystems, in both union and nonunion settings (Abraham and Medoff, 1984), seniority plays an important role in limiting the ability of management to assign workers to new tasks. By contrast, in many European nations and in Japan, management
faces few limits on its ability to assign particular workers to new jobs. Achieving flexibility of this sort has been a major management goal in recent rounds of "concession bargaining" in the U.S. and, as we shall see, is also reflected in efforts to transform industrial subsystems in the direction of the salaried model.

Flexibility in the deployment of labor is contingent upon the labor force either having a broad range of skills or being capable of learning quickly. Hence, a third aspect of flexibility relates to the hiring and training standards of the firm. If flexibility is strongly valued, the firm may hire more broadly trained workers than are required for the particular job and may also invest in apparently "superfluous" training, which in reality is aimed at preparing workers for subsequent mobility. This flexibility-enhancing aspect of training is often cited in cases ranging from the ability of a broad liberal arts education to prepare managers for new situations to the role of deeply grounded vocational education in preparing craft workers for new technologies.

To summarize the argument thus far, firms have several options regarding the organization of work. In making choices among alternative employment systems, firms weigh three objectives: cost minimization, predictability, and flexibility. Before turning to the constraints which shape these choices, it is important to note that each of these three objectives can have different, and conflicting, implications for the particular employment system that is chosen. There are trade-offs among the goals, and it is these trade-offs that introduce variety and dynamism into the process we are describing.

A straightforward way of illustrating potential conflicts among goals is to consider flexibility and predictability. A firm which seeks to maximize flexible staffing levels and deployment may reasonably seek to rely on the external market for its labor supply and hence to organize work along craft or secondary system lines. In this way, the firm makes few commitments or investments in particular workers; by the same token, the employees tend not to form firm-based work groups or expectations. When the external market functions well, skilled workers can be obtained without sacrificing flexibility. The external

---

4This may be surprising to those familiar with codetermination in nations such as Germany and Sweden. In fact, the internal labor markets of these nations are remarkably flexible. They have very few job classifications, seniority does not play a large role in allocating jobs, and, although management must consult with unions prior to moving workers to new jobs, the unions or work councils almost always agree to management preferences. The reason this pattern prevails is that job and social security is provided through other mechanisms: active labor market policy and the welfare state in the Swedish example and limitations on layoffs in the German case. For a discussion of these issues and evidence, see Osterman (forthcoming).

5The argument about the value of a deep "stock" of vocational skills is often made in relation to the German educational system, which by most measures would seem to overtrain its youth in vocational skills. The argument concerning the long-term value of a general liberal arts degree can be found in Useem (1986). In the United States, firms with a commitment to "lifetime employment" place a major emphasis on training (Eurich, 1985).
market does not always work well, however, and hence this strategy comes at some sacrifice of predictability. Under many circumstances, skill shortages may develop; so, too, may sharp and unforeseen wage movements. A firm which is committed to maximizing predictability will prefer to develop internal job ladders and on-the-job training systems, i.e., the attributes of the industrial subsystem. In this way, labor supply is under the firm’s control. The cost is limited flexibility.

The potential trade-offs involving cost minimization and the other objectives can be similarly illustrated. Secondary systems tend to minimize cost and to maximize flexibility, but these gains come at the cost of predictability since secondary workers have few attachments and minimal loyalty to the firm or, for that matter, to the labor market. There may also be constraints — especially skill requirements and union opposition — on the ability of firms to implement secondary systems.

Constraints on Choice

No firm is free simply to consider the options outlined above, assign weights to conflicting goals, and design and implement whichever employment system seems most desirable. In most instances, companies have a system already in place. Management is acclimated to working within specific staff patterns, and the labor force is accustomed to a particular set of arrangements. Any serious change usually will be incremental and in response to substantial external stimuli.

The limitations on choice run deeper than short-run inertia. In weighing choices among employment systems, the firm faces constraints imposed by the external and the internal environment. These constraints can be grouped into four categories: (1) the physical technology; (2) the social technology; (3) the nature of the labor force; and (4) the role of government.

*Physical technology.* This category includes the process through which the product gets produced (in the case of manufacturing), the paper processed (in the case of finance or other white-collar industries), or the services rendered (in the case of medicine, education, and fast food). The importance of technology is illustrated by the observation that a production process involving delicate equipment and rapidly changing configurations of that equipment is not compatible with casual secondary employment systems.

While this is true, there is often more than one efficient technological choice available to the firm and each option may have different employment implications. Consider alternative configurations of similar equipment. In automobile manufacturing, for example, pay-for-skill and team work arrange-
ments coexist with the traditional assembly line pattern. In the former case, workers are encouraged to develop broad skills and the link between wages and job classification is weakened. This contrasts with the narrow skills and strict seniority arrangements typical of most automobile production (Katz, 1985). In the white-collar area, the example alluded to earlier of alternative ways of arranging word processing equipment illustrates the same point.

Nevertheless, the most important constraint facing a firm in choosing among employment systems is the nature of the underlying physical technology. There are many conceptual systems for characterizing technology, but they are too aggregative for our purposes. In assessing the characteristics of physical technology which influence the choice of employment systems, useful categories are skill, risk, and configuration.

Some technologies require more skill than do others: Operating a machine tool — either conventional or numerically controlled — calls for more ability than does working on an assembly line. Computer programmers are more skilled than secretaries. As the required skill level rises, the employment system choices available to the firm narrow. In particular, reliance on secondary systems is not likely to be viable in the long run because the workers are so weakly attached to the labor force that long-term investment in skill development is unlikely. In the short run, however, secondary systems may be compatible with high skill. University faculty, for example, teach under various adjunct, part-time, and nontenure track arrangements. The difficulty is that product quality will eventually degrade given the inevitable high turnover and low institutional commitment.

Skill level may be decisive in limiting the choice to industrial/salaried or craft systems, but it does not dictate which of these options is most desirable. Skills can be learned either inside the firm or in external training institutions, and under both systems workers have a long-term commitment (to the firm or to the profession/craft). Thus, it is not surprising that high-skill occupations organized under both arrangements are easily identified. A more important constraint concerns the structure of risk inherent in technology.

There are two significant aspects of risk: its level and its distribution. The more vulnerable is equipment or product to worker error, the greater is the level of risk. It may require relatively little skill to monitor the gauges of a chemical (or nuclear) plant, but the consequences of error are severe. A similar though less apocalyptic point might be made about the operator working with expensive numerically controlled machine tools or a manager designing

---

*Examples of well-known classification systems are Woodward’s (1965) distinction between unit, batch, and mass production; Thompson’s (1967) discussion of long-linked versus intensive technologies; and Piore and Sabel’s (1985) distinction between flexible specialization and mass production technologies.*
a new product line. As the level of risk goes up, the firm will find it more
difficult to cede control to external training and socialization agencies. The
likelihood of industrial or salaried systems therefore increases.

The distribution of risk between the firm and its workers also conditions the choice. In circumstances where the consequences of failure can be shifted
to the employee (e.g., piecework pay systems), the firm will be less concerned
with control and predictability. White-collar equivalents to piecework are
bonus and other incentive schemes. However, in order for these sorts of plans to be viable, a particular category of technology is required, namely
techniques in which individual performance can be isolated and measured.
High turnover employment systems are more likely in these circumstances.
The costs of poor performance are borne by the worker, so the firm will be
less concerned to design internal systems that maintain commitment and
loyalty.

A final characteristic of technology as a constraint includes a miscellaneous
collection of physical attributes which we group under the term “configuration.”
Under some technologies, the flow of work can be more easily interrupted
or the volume adjusted than under others. The more difficult and costly are
such adjustments, the higher will be the premium placed on securing a stable
and predictable workforce. Secondly, while most production processes must
be centralized, there are cases in which work can be distributed or “put out”
(using micro-computers, for example). This form of organization (which is
often tied to wage incentive schemes) permits more flexible labor force systems.
Finally, some technologies — e.g., construction or harvesting — are less
regular than others and this may preclude long-term attachments to firms.

If technology sets the demand side limits within which firms must work in
choosing employment systems, the nature of the available labor force acts
as a constraint on the supply side. Secondary systems do not offer stable
careers, high wages, or advancement. Employers may have difficulty finding
an adequate labor force willing to accept these conditions, and expansion of
these systems may await or depend upon securing an appropriate labor force.
Economic historians have explained the timing of the expansion of clerical
jobs by reference to changing education and family patterns of young middle-
class women (Goldin, 1980), and Piore (1979) interpreted the recent surge
of undocumented workers in terms of secondary employers seeking out alter-
native labor sources. These considerations are also relevant in better jobs.
Craft workers’ commitment to their profession rather than to their firm can
make it more difficult for firms to retain these workers. In order to stabilize
these fields, it may be necessary to draw staff from nontraditional sources
and socialize them using new training schemes.
Most of the economics literature suggests that supply constraints are likely to be very important. Standard labor economics models tend to be supply driven, with worker mobility the main mechanism for equalizing returns and conditions at the margin, and with persistent differentials in wages, working conditions, employment security, and the like ultimately attributable to differences in worker characteristics. As the examples above suggest, supply constraints are important. But, in contrast to the literature, they are more likely to matter in the short than the long run. Firms can be ingenious in finding recruiting patterns, training schemes, and other mechanisms for eventually overcoming whatever limitations the character of the available labor force may place upon work organization.7

The third category of constraints concerns the "social technology" of the firm or the particular cluster of tasks at hand. The literature on the sociology of bureaucracy suggests that a set of tasks may be viewed as central or peripheral to the purpose of the organization (Thompson, 1967). Tasks which are central are less likely to be entrusted to outsiders or to people who have not had extensive socialization in the purposes and procedures of the organization. For this reason, firms are more likely to entrust programming or clerical work to a high turnover labor force, preserving product development or quality control for more permanent employees.8

Gouldner (1954) and Crozier (1964) have noted that unity of purpose within an organization cannot be assumed and that employment relations may reflect the outcome of power struggles among subgroups. Although both authors had concrete issues in mind (control over knowledge of production techniques in Crozier's case and work rules in Gouldner's), the general point can be extended to include, among other issues, status maintenance. This is particularly important with regard to the pattern of female employment. The work rules and employment arrangements which govern many "women's jobs" tend to impose shorter career lines and higher turnover than appears to be true for

---

7Firms can adjust hiring standards and investment levels in internal training to compensate for unavailability of adequate labor. This is illustrated by recent research on the alleged "mismatch" between the skills of inner city residents and the requirements of downtown office jobs. Some companies successfully followed a strategy of hiring "lower quality" labor and compensating by trading off reduced wages for more elaborate internal training. This is another internal labor market strategy to cope with unpredictable labor supply (Osterman, 1983). In addition to this evidence, Doeringer and Piore (1971) list 11 "adjustment mechanisms" which firms employ to respond to labor shortages. See also Lester (1954).

8Evidence of variation in the employment systems under which clerical workers are organized is provided by a survey of 59 California firms conducted for the Department of Labor. The survey examined conditions under which firms organized work into stable (or core) or high-turnover (or peripheral) systems and concluded that "Occupations associated with full-time permanent positions in some industries (certain clerical and sales functions, for example) were heavily found to be involved in the flexible staffing arrangements of other industries. ... in firms whose workforce was largely of a blue-collar nature where these clerical and sales workers enjoyed the twin protections of specialized knowledge and small numbers. ... they rarely figured in flexible staff arrangements" (Sugarman, 1978, pp. 3-5).
comparably skilled male-dominated occupations. The very identification of certain occupations as stereotypically male and others as stereotypically female raises difficult questions. However, the sociological literature, particularly ethnographic studies of particular employment situations, leaves little doubt that status and power issues are relevant in understanding differential structures for male and female jobs and career lines within firms (Kanter, 1978).

We have said little about the role of unions, not because they are unimportant, but rather because we believe that the most fruitful first step is to focus on managerial decision making in order to make the model generalizable throughout the labor market. Unions are clearly important in shaping employment relationships, and they can be considered a constraint upon firms' choices. Under many circumstances, for example, unions will oppose establishment of secondary systems. A good example from our interviews is the reluctance of some firms to establish in-house temporary help pools even though their extensive use of temporary workers would make such an internalization profitable. Instead, fearing organizing drives, these firms choose to employ outside temporary agencies. Unions are also clearly important in maintaining established internal labor market patterns in the face of company efforts to transform them.

The final category of constraints concerns the role of the government in organizing and regulating labor markets. One subset of government activities includes legislation such as minimum wage, labor laws, and unemployment insurance. Unemployment insurance, for example, has been structured in a way which makes part-time work and job sharing difficult. Another subset includes regulation. A powerful example here is affirmative action programs. These have, on at least some occasions, led firms to redesign mobility paths across particular job ladders (Shaeffer and Lynton, 1979; Osterman, 1982). Finally, government can influence the labor supply through public education, continuing education, and job training programs. To the extent that business can substitute public for private training, the calculations involved in choosing among particular employment arrangements can change.

Summary

The key assumptions underlying the model sketched here are that firms choose among alternative ways of organizing work and that it is helpful to view these choices in terms of categories or employment subsystems. We identified three central goals that firms consider in making these choices: cost minimization, predictability, and flexibility. Each has different implications for the choice of subsystems and some balancing or weighing of conflicting objectives is necessary. Four factors constrain the choices made: physical
technology, social technology, the characteristics of the labor force, and government policies. The interaction of objectives and constraints leads firms to implement a specific subsystem (or subsystems). These choices constitute the internal labor market and hence the discussion presented here is the beginning of a theory of the determination of internal labor market structure. This theory is tentative and must be subjected to more systematic empirical research. One recent example of such research is Baron et al.'s (1985) study of job ladders within a range of California firms. The authors describe a pattern of internal variation consistent with the model presented here: "... among the 84 establishments... there is tremendous diversity in the amount of opportunity and the way it is structured" (p. 16). Rosenbaum (1984) employs extensive personnel data on a single firm and finds very different career patterns for employees in different locations within the firm's internal market. He concludes that systematic structural differences are important. These examples suggest that variation and change in contemporary internal labor markets offers a largely unexplored arena for empirical study.

A logical strategy for further work would be comparative research, examining, for example, changing employment conditions across a range of job families (occupational clusters) and across different firms. Developing a matrix of job clusters by firm would achieve variation in constraints and weights among the competing goals, and the consequent ability to hold firm and also occupational characteristics constant would permit a determination of the effect of key variables. For example, by holding job cluster constant and looking across firms, it should be possible to understand the impact of labor supply or social technology. By holding firms constant and looking across job clusters, the impact of technology can be explored.

This leaves open how to operationalize measures of alternative employment systems. One fruitful, albeit time-intensive approach is described in Osterman (1984). In that study, we collected data on the length of job ladders; the extent to which they were open or closed to the outside and to other ladders within the firm; turnover rates; the location of training for jobs; the length of time it took to achieve proficiency; and other relevant variables. We were able to differentiate job ladders in ways which are consistent with the theory developed here.

Once validated by more empirical research, the model would provide a more fully developed institutional theory of internal labor markets, and it would also offer insight into the determinants of changing employment conditions in the economy. An example is the changing balance between the industrial and salaried models.
As we have seen, the salaried and industrial models each entail distinctive costs and benefits. The industrial model often implies considerable rigidity in wage structure and the deployment of labor. However, it provides firms with flexibility in adjusting employment levels via layoffs. The attention to employment security in the salaried model restricts the flexibility of the firm concerning employment levels. In return, the firm gains a measure of internal flexibility and commitment which may be lacking in the rival system. In the current period, the valuation which firms place upon these alternatives may be changing.

The industrial model is under extreme pressure from several sources. New technologies are altering the traditional job classification systems and are pressuring the firm toward the more flexible salaried model. These technologies require greater flexibility in internal deployment than the industrial model permits.

Employers face tension from a different source for white-collar work. These jobs are already organized along the lines of the salaried model. But that model, too, is in some jeopardy. Employment reductions resulting from competitive pressures and the nature of new technologies may erode the job security associated with the salaried model. If employment reductions are too severe, then employers may lose the commitment of their white-collar employees and other consequences, such as union organization, may follow.

Employers are therefore in a bind. In the case of blue-collar work, many want to shift to a salaried model. Yet, macro-economic uncertainty and unwillingness to permit labor to become too great a fixed cost generates fears of the high price of employment security. In the case of white-collar employment, firms want to maintain the salaried model but at lower employment levels. These conflicting pressures and firms’ efforts to resolve them help explain much confusion and mixed signals about the current direction of work organization. Some employers are attempting to finesse the problem by imposing flexibility via concession bargaining and employer militancy. Other firms are experimenting with a core-periphery model in which they offer job security (or a commitment to make every effort to avoid layoffs) to a core of more or less permanent employees and surround that core with a “periphery” of temporary, contract, and part-time workers who enjoy less protection. The workers in the “core” will be willing to work under the salaried model and to provide both flexibility and commitment to the firm. The peripheral labor force provides the firm with a buffer against either cyclical downturns or labor force reductions necessitated by technical change.

It is highly uncertain whether either the employer militancy or the core-periphery strategies can provide a stable solution to the pressures facing
employers. To explore this question in detail requires further analysis. Our central concern here has been to provide an improved understanding of the internal labor market alternatives available for organizing work and of the considerations that lie behind the choices made. This improved understanding hopefully contributes both to the theoretical literature on labor market organization and to understanding recent developments in the workplace.

References


------. "Employment Structures Within Firms," _British Journal of Industrial Relations_, XX (November, 1982), 349-361.