OXFORD

ON SOCIOLOGY

Numbers, Narratives, and the Integration of Research and Theory

JOHN H. GOLDTHORPE

10

Social Class and the Differentiation of Employment Contracts

MY CONCERN here is with the theory of social class: I aim to contribute to the ultimate goal of explaining why social classes exist. However, I start out from a way of conceptualizing class, and in turn of making 'class' operational in empirical research, that I have developed, along with a number of colleagues, over the last twenty years. The fundamental idea that has been pursued is that class positions can be understood—in a way to be explained more fully below—as positions defined by *employment relations*. A 'class schema' has been progressively elaborated that differentiates class categories by reference to such relations, and that can be implemented in research through information on employment status and occupation.

This programme of work was undertaken as the basis for studies of social mobility envisaged within a class structural context (Goldthorpe and Llewellyn 1977; Goldthorpe 1987; Erikson *et al.* 1979; Erikson and Goldthorpe 1992*a*). However, the class schema, in one or other of its several versions, has subsequently become used in many other areas of research.[1] As a result, evidence of its *construct* validity has accumulated: that is, of its capacity to display associations or correlations of a theoretically expected kind with other variables: for example, as regards patterns of class voting (Heath *et al.* 1991; Evans *et al.* 1991, 1996) or class differentials in educational attainment (Jonsson 1993; Müller and Haun 1994; Jonsson *et al.* 1996) or in health (Bartley *et al.* 1996; Kunst 1996). One line of further research that is then prompted—but which is not my present concern—is that aimed at elucidating exactly how class, as conceptualized in terms of employment relations, exerts its influence on different dependent variables. What causal processes, ultimately at the level of social action, are involved? For example, just how does the incumbency of different class positions operate so as to produce the empirical regularities that are apparent in the association between class and party support (cf. Evans 1993) or in that between class and educational choice (cf. Chapters 8 and 9 above)?

This chapter has not been published previously. I am indebted to Sam Bowles, Richard Breen, Jerker Denrell, John Ermisch, Duncan Gallie, Dan Krymkowski and Aage Sørensen for comments on an earlier draft, and to Tony Atkinson, Meg Meyer and Karen O'Reilly for useful information and references.

[1] The schema is known, in consequence of its rather complex genesis, by several different names. In a British context it is usually referred to as the Goldthorpe schema; in an international context, as the EGP (Erikson–Goldthorpe–Portocarero) schema, the Erikson–Goldthorpe schema, or the CASMIN schema, after its use in the Comparative Analysis of Social Mobility in Industrial Societies Project.

At the same time, the wider use of the schema has also encouraged interest in its *criterion* validity: that is, in the extent to which, as operationalized through occupation and employment status, it does in fact capture those differences in employment relations that it is, conceptually, supposed to capture. The findings so far reported on this issue, mainly by Evans (1992, 1996) and Evans and Mills (1998) and based on British data, indicate that the schema does in general perform quite well, and these same authors have more recently extended their research to newly capitalist eastern European nations (Evans and Mills 1999), again with encouraging results.[2] In particular, for individuals of employee status, occupation can, it appears, for the most part serve as an adequate proxy for those features of their employment relations that the schema takes as distinguishing class positions. In this way, then, another line of further investigation is suggested and one that I shall here seek to pursue: namely, that directed to the question of just why it should be that different occupations do tend to be associated with differences in the employment relations of those engaged in them—of the kind that I would in turn see as implying different class positions.

In the conceptualization of class that underlies the schema (see further Erikson and Goldthorpe 1992*a*: 35–47), basic distinctions are made among employers, the self-employed, and employees: that is, among those who buy the labour of others, those do not buy the labour of others but neither sell their own, and those who do sell their labour to an employer or employing organization. Why these three categories should exist is not itself especially problematic, or at least not in the context of any form of society that sustains the institutions of private property and a labour market. However, in such societies in the modern world the third category, that of employees, is numerically quite preponderant, usually accounting for some 85–90 per cent of the active population.[3] Thus, what is crucial to the class

[2] So far as the British case is concerned, much further evidence of the criterion validity of the schema is in effect provided by a series of studies carried out as part of a review of official 'social classifications' by a committee appointed by the Economic and Social Research Council, on behalf of the Office of National Statistics. The committee proposed, and ONS subsequently accepted, a new 'Socio-Economic Classification' that takes over the conceptual basis of the class schema, although with a revised algorithm for its implementation with British occupational and employment status data. This classification was then itself subjected to a series of tests of its criterion (and also construct) validity, which gave generally positive results (see Rose and O'Reilly 1997, 1998).

[3] Why this very skewed distribution of individuals among the three categories should be found *does* of course constitute a very significant problem: as Simon (1991) has put it, that of why there are employing organizations and employees at all, rather than simply independent contractors or, alternatively, of why there is not just a single, encompassing organization. This problem is in fact centrally addressed by the organizational and transaction-cost economics on which I draw later and is one of obvious relevance to understanding how class structures of different 'shape' have evolved. However, it is not that with which I am here concerned.

schema is the further level of distinction that is introduced, applying specifically to the employment relations *of employees*. This focuses on the form of regulation of their employment or, one might alternatively say, on the nature of their employment contracts, explicit and implicit.

In this regard, the main contrast that is set up is that between, on the one hand, the 'labour contract', supposed typically to operate in the case of manual and lower-grade nonmanual workers, and, on the other hand, the 'service relationship', as expressed in the kind of contract taken as typical for the professional, administrative, and managerial staffs of organizational bureaucracies, public and private.

> Employment relationships regulated by a labour contract entail a relatively short-term and specific exchange of money for effort. Employees supply more-or-less discrete amounts of labour, under the supervision of the employer or of the employer's agents, in return for wages which are calculated on a 'piece' or 'time' basis. In contrast, employment relationships within a bureaucratic context involve a longer-term and generally more diffuse exchange. Employees render service to their employing organization in return for 'compensation', which takes the form not only of reward for work done, through a salary and various perquisites, but also comprises important prospective elements—for example, salary increments on an established scale, assurances of security both in employment and, through pensions rights, after retirement, and, above all, well-defined career opportunities. (Erikson and Goldthorpe 1992*a*: 41)

It is recognized that these two basic forms of the regulation of employment may exist with degrees of modification and, further, that 'mixed' forms also occur—typically associated with positions intermediate between bureaucratic structures and rank-and-file workforces: for example, those of clerical or sales personnel or of lower-grade technicians and first-line supervisors. Table 10.1 summarizes the argument in relation to the categories of the class schema.

The main significance, for present purposes, of enquiries into the criterion validity of the class schema is then the following. They reveal that the regulation of employment of different occupational groupings of employees does in fact tend to follow the pattern indicated in Table 10.1. More specifically, when indicators of various relevant features of employment relations are considered for samples of the economically active population—concerning form of payment, perquisites, control of working time, security, opportunities for promotion, and so on—then occupations as differentiated in terms of these indicators are found to map onto the class categories distinguished by the schema in ways broadly consistent with its conceptual basis.[4] The question

[4] The indicators were derived from questions put to respondents concerning the methods by which they were paid (including arrangements regarding overtime and absences), the control of their working time, the period of notice they were required to give on leaving, the opportunities for promotion offered by their jobs, and so on. Various methods of analysis were used, including latent class analysis. Substantially better results than those reported could not, I believe, be expected given that difficulties in eliciting the precise information required from respondents would seem likely to have created a good deal of 'noise' in the data. For example, respondents reporting that they were paid weekly might in fact mean not that they had a fixed weekly wage but rather that, while being paid by the hour, they actually received their earnings weekly; being on an incremental scale might be confused with receiving annual cost-of-living increases; and respondents might report on their own *personal*

TABLE 10.1 *Categories of the class schema and supposed form of regulation of employment*

Class		Form of regulation of employment
I	Professionals, administrators and managers, higher-grade	Service relationship
II	Professionals, administrators and managers, lower-grade, and higher-grade technicians	Service relationship (modified)
III*a*	Routine nonmanual employees, higher grade	Mixed
III*b*	Routine nonmanual employees, lower-grade	Labour contract (modified)
IV*abc*	Small proprietors and employers and self-employed workers	
V	Lower-grade technicians and supervisors of manual workers	Mixed
VI	Skilled manual workers	Labour contract (modified)
VII*a*	Nonskilled manual workers (other than in agriculture)	Labour contract
VII*b*	Agricultural workers	Labour contract

Note: The class descriptions given should be understood as labels only. In any implementation of the schema, the detailed occupational groupings to be allocated to each class are fully specified. For Britain, see Goldthorpe and Heath (1992).

that arises is, 'therefore,' that of how this empirical regularity actually comes about. Why should there be such a tendency for individuals engaged in different kinds of work to have their employment regulated via different contractual arrangements and understandings?

In previous work (Goldthorpe 1982; see also 1995*a*), I have made some attempt at an answer to this question—with particular reference to the idea of the formation of a 'service class' or 'salariat'—but only in a brief and rather *ad hoc* fashion. Here, I aim at a more systematic and theoretically informed treatment. I draw on rational choice, or, as I would prefer to say, rational action theory (RAT), and in particular on such theory as deployed in recent organizational and personnel economics (see e.g. Milgrom and Roberts 1992; Lazear 1995) and in the 'new' institutional, especially transaction cost, economics (see e.g. Williamson 1985, 1996). I would not regard this cross-disciplinary 'borrowing' as in any way implying the abandonment of a sociological perspective (as is suggested, for example, by Pfeffer 1997: ch. 9). What is striking about the economics literature referred to is how much of it can

promotion chances rather than those generally associated with the jobs they held. Better-quality data could almost certainly be obtained directly from employers, provided that sampling problems could be overcome.

in fact be read as a more rigorous development—from the standpoint of a particular theory of action—of observations and insights already to be found in the industrial sociology of the later 1940s and 1950s.[5] Moreover, the version of RAT that is chiefly in use, at least in transaction-cost economics, is not the utility theory of neoclassical orthodoxy but a version in which the idea of objective rationality gives way to that of subjective, or 'bounded', rationality: that is, actors are seen as being, in Simon's words (1946/1961; cf. Williamson 1985: ch. 2), 'intendedly rational but only limitedly so'. As I have argued above (Chapter 6), this modification produces an obvious convergence with approaches to the theory of action that are central to the classic sociological tradition; and all the more so to the extent that rationality is seen as restricted not only by psychological, cognitive constraints on information processing but further by social constraints on the availability of information, or knowledge, itself.

Such efforts as sociologists have previously made to account for variation in the form of employment contracts (see e.g. Edwards 1979; Wright 1985, 1989, 1997: ch. 1) derive largely from Marxist 'political economy' (cf. Marglin 1974; K. Stone 1974; Bowles and Gintis 1976) and are in turn characterized by an almost exclusive emphasis on considerations of power and control. The basic assumption is that, under capitalism, the prime concern of employers will be to maximize the 'exploitation' of their workers: that is, to maximize the extraction of actual labour from employees' working time. Employers will thus in general aim to establish forms of contract that are to their greatest advantage in this regard in that they effectively 'commodify' labour. To the extent that variation in contracts does occur, it is then to be explained in terms of employers seeking to privilege their managerial and supervisory staffs as a means of buying their loyalty in the process of exploitation or to create conflicts of interest among employees as part of a larger strategy of 'divide and rule'. Organizational and transaction-cost economists have, however, been critical of such Marxist interpretations on both theoretical and empirical grounds. The main counterclaim that they have advanced is that, rather than being seen as expressions of power and means of exploitation, most features of employment contracts are better understood in terms of efficiency: that is, as serving not only to ensure the viability of the enterprise within a competitive market context but further to increase the *total value* of the contract, to the benefit of all parties involved (see e.g. Milgrom and Roberts 1992: ch. 10; Williamson 1985: 206–11; 1994).

In what follows I aim to take up an intermediate position that avoids what I would see as ideologically induced weaknesses in the more extreme versions of both the 'exploitation' and 'efficiency' arguments.[6] I treat employment

[5] This statement may appear less surprising if certain common sources are recognized, most notably Barnard (1938) and Simon (1946/1961), and also a similar starting point in the critique of the conceptualization of work, the firm, and employment relations within orthodox neoclassical economics.

[6] In particular, I seek to avoid, on the one hand, the self-indulgence of those Marxists who assume that in some future world the abolition of capitalist institutions will make possible

contracts primarily from the standpoint of employers, with whom the *initiative* in their design and implementation does at all events lie. I first set out what would appear to be certain general problems of the employment contract as such, and I then try to show how the different forms that this contract may take can be understood primarily as employers' responses to the more particular ways in which these problems arise in the case of employees involved in different kinds of work. I make the assumption that in this regard the 'central tendency' is for employers to act as rationally as they are able towards the goal of maintaining the viability and success of their organizations within the context of whatever constraints they may face. This may then lead them, depending on the specific circumstances that obtain, to view their contractual relations with employees in *either* zero-sum *or* positive-sum terms—just as employees may take a similarly varying view of their contractual relations with employers. In other words, I see no reason to treat the interests of employers and employees as being 'fundamentally' either in harmony or in conflict.[7]

General Problems of the Employment Contract

It is has for long been recognized by economists and sociologists alike that the employment contract has distinctive features (see e.g. Commons 1924; Simon 1951; Baldamus 1961). These stem basically from the fact that the labour that is bought by employers on the labour market cannot be physically separated from the individual persons who sell it. What is in effect bought and sold through the employment contract is not a commodity, or at least not in the sense of some definite, objective thing, but rather a social relationship. Employment contracts are contracts through which employees agree, in return for remuneration, to place themselves under the authority of the employer or of the employer's agents.[8] Further, though, employment

the production of 'new' men and women with preferences, orientations to work, and so on, such that present-world problems of efficiency will be entirely transcended; and, on the other hand, the Panglossian tendencies of those economists who assume that 'whatever is, is efficient' and fail to give due recognition to the conflictual or 'contested' aspects of employment relations.

[7] I do not assume, I should stress, that employers' rationality implies that they subject their contractual relations with employees to some continuous process of review and revision; only that they are ready to re-examine and modify these relations from the point of view of organizational effectiveness when prompted to do so by changing circumstances (cf. Osterman 1987; also Chapter 6 above). I also recognize that, although employers have the initiative in the design and implementation of employment contracts, they will typically act under various constraints. These are, of course, likely to include employee responses to their initiatives, whether of an individual kind or as expressed through trade unions or other representative organizations, and also whatever legislative and regulatory framework is imposed on employment relations by the state. The nature of such constraints can, however, be expected to show great variation by time and place. Thus, a focus on the actions of employers in dealing with highly generalized contractual problems would seem appropriate, given that my concern is with explaining broad probabilistic regularities in the association between forms of contract and types of work rather than the deviations from these regularities that will certainly be found.

[8] In what follows, 'employer' should be understood as also covering what is in fact the

contracts are in varying, but often substantial, degree implicit or in fact incomplete, and especially in regard to what employers may demand of employees and what in turn the obligations of employees are. Employers buy the right to tell employees what to do while at work, and minimum requirements may be formally laid down concerning, say, hours of work, working methods and procedures, and so on. But contracts rarely if ever seek to specify just how hard employees should work—what intensity of effort they should make—let alone what degree of responsibility, adaptability, or initiative they should be ready to show in their employer's interest. Such matters would indeed seem largely to defy formulation in explicit contractual terms.

From the employer's point of view, therefore, a major objective must be that not merely of *enforcing* the compliance of employees with the authority that they have in principle accepted but, further, that of *inducing* their maximum effort and cooperation in the performance of the work allocated to them. Another way of putting the matter would be to say that within the employment contract employees will always have some non-negligible amount of discretion; and that it will then be of obvious importance to employers and their agents to ensure that this discretion is as far as possible used in ways that support rather then subvert the purposes of the employing organization.

In the industrial sociology of the immediate post-war years, which had a strongly managerialist orientation, the problems arising in this regard were treated in terms of the degree of congruence prevailing between 'formal' and 'informal' organization—especially, that is, between employers' work rules and work-group values and norms. The greater the congruence that could be achieved, the higher, it was supposed, would be levels of employee 'motivation and morale'; and this goal was to be pursued through 'human-relations' policies, implemented by first-line managers and supervisors at the level of the work group—or, as critics would have it, through social-psychological manipulation.

In the economics literature previously referred to essentially the same issues are addressed, also in fact largely from the employer's standpoint, but in a different idiom. The key issue is taken to be that of how the employment contract may be most efficiently elaborated, not only in its explicit *ex ante* design, to which limitations clearly apply, but, more importantly perhaps, in its *ex post*, and possibly quite implicit, interpretation and actual day-to-day execution; or, that is, in the way in which it serves as the basis for the continuing regulation of employment relations over time.[9] At a minimum, the employer must be given protection against employee shirking or 'opportunism'; but it is a further requirement that employee interests should as far as possible be aligned with those of the employer or, in other words, that appropriate incentive structures should be set in place. And at the same time transaction costs have to be taken into account: that is, the arrangements and procedures involved in actually implementing the contract must be cost-effective as compared with available alternatives.

more likely case of 'employing organization'. It should also be noted that throughout I use 'agent' only in (explicit or implicit) contradistinction to 'principal', and not as a synonym for 'actor'.

[9] The idea of 'implicit contracts' or of implicit provisions in contracts has taken on particular importance in the economics literature. Ehrenberg and Smith (1991: 409) refer in this connection to 'a set of shared, informal understandings about how firms and workers will respond to contingencies' and Gibbons (1997: 11) refers to 'an understanding backed by the parties' reputations instead of law'.

The rational action theory that informs this latter approach does, I believe, endow it with greater intellectual coherence than that achieved by early industrial sociology. However, for my present purposes its most immediate attraction is that it leads naturally to recognition of the fact that employment contracts will need to take on different forms in relation to the different kinds of work task and work role that employees are engaged to perform.

Differentiation of the Employment Contract and Types of Work

What is to be explained here is, to repeat, the association between different occupational groupings of employees and the form of regulation of their employment that can, it appears, be empirically demonstrated on the lines indicated in Table 10.1. To this end, it is necessary to consider types of work analytically, and in a more abstract yet focused way than could be achieved by reliance on occupational designations themselves. The organizational and transaction-cost economics on which I draw would suggest two main dimensions in terms of which potential problems—or sources of 'contractual hazard'—from the employer's point of view can be identified (cf. Weakliem 1989):

(i) the degree of difficulty involved in monitoring the work performed by employees: that is, the degree of difficulty involved both in measuring its quantity and also in observing and controlling its quality; and
(ii) the degree of specificity of the human assets or human capital—skills, expertise, knowledge—used by employees in performing their work: that is, the degree to which productive value would be lost if these assets were to be transferred to some other employment.

In pursuing the explanatory task in hand, I shall therefore find it helpful to refer to the two-dimensional space described in Fig. 10.1. Work that falls into the lower-left quadrant of Fig. 10.1 is that which may be expected to give rise to fewest hazards for employers as regards the employment contract. The absence of serious work monitoring problems means that some kind of 'variable-pay' system can operate or, in other words, employees can be remunerated in direct relation to their productivity. And the absence of serious asset specificity problems means that no understandings need be entered into about the long-term continuation of the contract. No such understandings are required in order to provide employees with an incentive to acquire skills, etc., of specific value in their present employment and then to retain them in this employment. Under these conditions, the employment contract can

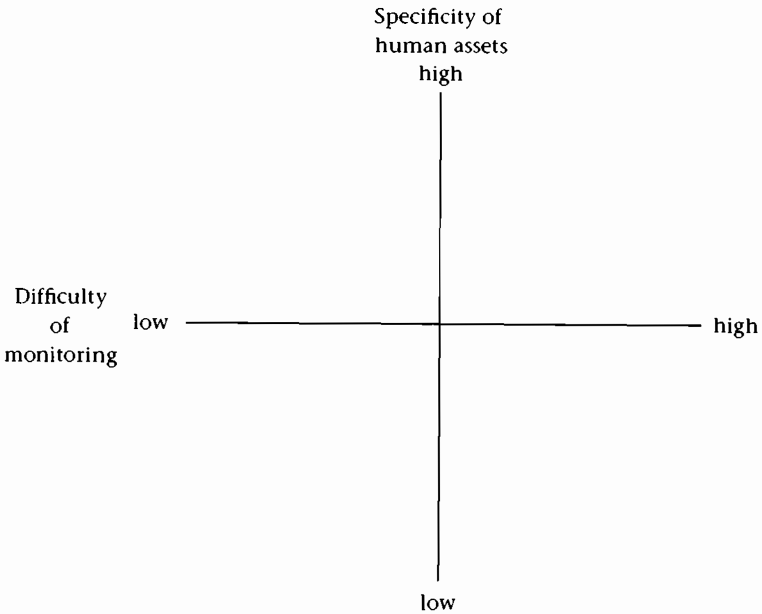

FIG. 10.1. Dimensions of work as sources of contractual hazard

simply provide for discrete, short-term exchanges of money for effort, in the way characteristic of the labour contract as earlier described, and thus come in fact as close as is possible to a simple spot contract—albeit perhaps of a recurrent kind—for the purchase of a quantity of a commodity (cf. Kay 1993: ch. 4).

However, the question does then at once arise of why it is that, despite the evident advantages of the labour contract to employers, the occupational range of its application would appear, as is indicated in Table 10.1, to be rather restricted: that is, in its pure form to nonskilled manual occupations and, somewhat modified, to skilled manual and lower-grade nonmanual occupations. What will here be argued is that this limitation is to be explained in terms of various *concomitant* features of the kind of work in regard to which a labour contract proves viable.

The occupational restriction of the labour contract

Difficulties in measuring the quantity of work done by employees will be least, and a variable pay system will thus be most easily implemented, where measurement can be based on actual *output*. In this case, a direct link between work and pay can be established through piece rates of some kind. However, work that can be thus measured and remunerated is likely to have various

other characteristics. To begin with, the measurement of work by output implies output that is of a specific, well-defined kind, and efficient payment by output implies a production process that is relatively simple: first, so that output can be clearly attributed to particular individuals or at most to small work groups and, secondly, so that the employer—as well as employees—can have reliable knowledge of how quickly the work can be done and can therefore set an appropriate rate of pay.[10] Moreover, since piece rates give an inducement to workers to concentrate on quantity of work at the expense of quality, it is important, from the employer's point of view, that the quality of the product as well as its quantity is easy to monitor—that is, to observe and assess—and likewise such other aspects of work quality as the use of tools, equipment, and raw materials (cf. Milgrom and Roberts 1992: 394–5). It could then be said that the kind of work with which piece rates are most likely to be associated is work in which workers, acting individually or in small groups, undertake physical (rather than mental) operations that lead in a fairly transparent way to discrete material (rather than symbolic) results. Typical piece-rate workers are in fact fruit and vegetable pickers, various kinds of loaders, fillers, and packers, and machinists in batch-production manufacturing industry.

It may also be possible for work to be more or less adequately measured by *input* in the sense of time spent at work, and thus for a variable pay system to operate through time rates, calculated, say, on an hourly or daily basis. Once more, though, for this to be the case—in effect, for time worked to be informative about output—restrictions would seem to apply to the nature of the work involved. If under piece rates the employer's main monitoring problem is that of quality, under time rates it becomes that of assessing and maintaining the level of worker effort. This problem will be least severe, and time rates thus most attractive to employers, where workers have in fact only limited autonomy in regard to their pace of work: where, for example, this is largely determined by technology, as in assembly-line or continuous-process production, or by the flow of customers or clients, as in the case, say, of checkout operators, ticket-sellers, or counter staff. Otherwise, it will be important that worker effort should be easily observable, and thus open to control through supervision; and this would then again tend to imply work activity

[10] If only workers have such knowledge, employers are exposed to opportunism. Workers may shirk in order to conceal this knowledge from their employer and in this way seek to get a higher rate of pay than they would if the employer were better informed (cf. Gibbons 1987, 1997). One has here a prime illustration of the similarity of concern but divergence of approach of the contemporary economics on which I draw and the industrial sociology of the 1940s and 1950s. In the latter, the restriction of output was recognized as a likely problem of piece-rate payment but was then represented as the outcome of a failure in 'communication' and 'human relations'—specifically, management had failed to build up work-group norms supportive of organizational goals. Interestingly, a yet earlier account of the phenomenon, that of Max Weber (1908), himself trained as an institutional economist, comes much closer in spirit to that of his present-day counterparts in treating the restriction of output under piece rates as in many circumstances an entirely rational strategy for workers to pursue.

with a clear physical component, even if not necessarily of a kind conventionally classified as 'manual' (cf. Fama 1991).

The payment of employees in return for discrete amounts of work done, whether by piece or time, is one defining element of the labour contract. The other is that the exchange is of a short-term nature in the sense that, while it may in fact be many times repeated, there is nothing in the contract itself, explicit or implicit, that is aimed at securing the relationship between employer and employee on a long-term basis. As earlier suggested, an employer is able to operate with such a contract where there is little to be gained in encouraging workers to invest in the acquisition of human assets specific to their present employment, and in turn little to be lost if employees should leave this employment—that is, the costs of labour turnover are slight. Here too, though, the argument may be made that, where such a situation prevails, there are likely to be further implications for the type of work that is involved.

Thus, while in principle a workforce with which no problems of human asset specificity arose could still be a skilled workforce—that is, one reliant simply on general purpose skills—it would seem empirically to be the case that, where general purpose skills are brought to particular employments, it tends to be both possible and advantageous for further, more specific skills, expertise, and knowledge to be developed around them.[11] Consequently, a situation in which employers need take no account of asset specificity in regard to their employees can be reckoned as most probably one in which employers are able to recruit the workers they require from a fairly homogeneous pool of labour, the individual members of which are substitutable for each other without serious loss of productive value on the basis simply of their physical capacities plus, perhaps, minimal literacy and numeracy.

In the light of the foregoing, then, what underlies the restriction of the occupational range of the labour contract should be more apparent. It is work with features that locate it in the lower-left quadrant of Fig. 10.1 that allows employers to resort to this form of regulation of employment. But work that is easily measured and otherwise monitored *and* that in itself offers little potential for the development of specific human capital will have other characteristics too. Its archetype can in fact be regarded as manual work of a non-skilled kind, or what might be thought of as labour in its most basic sense. It is, then, with such work that the labour contract can operate in its purest form or, in other words, that employers can take the 'commodification' of labour to its furthest possible point. Correspondingly, any extension of the labour contract beyond such work is likely to entail some departure from the

[11] Various authors have pointed out that general qualifications are valued by employers not only on account of the skills, expertise, and knowledge to which they directly attest but further, and perhaps primarily, as indicators of the individual's *capacity to learn*. Correspondingly, employees often report that in the course of their work they draw on their prior qualifications only to a surprisingly limited extent (see e.g. Thurow 1972; Wilensky and Lawrence 1980; Cohen and Pfeffer 1986; Bills 1988).

pure form in one direction or another. For example, where the monitoring of work is not entirely straightforward, as regards either quantity or quality, the strict principle of pay in return for discrete amounts of work done will need to be modified in some degree. Thus, a *weekly* wage with, perhaps, provision for overtime pay or time off 'in lieu' for work in excess of a given number of hours is a fairly common arrangement among more skilled manual and lower-grade nonmanual employees. And likewise such workers may be given certain privileges of seniority—such as pay guarantees or a 'first-in-last-out' understanding in the case of redundancies—in circumstances where employers are compelled to recognize some need for the development and retention of human assets of an organization-specific kind (cf. Doeringer and Piore 1971; Weakliem 1989).

The full significance of these latter points can, however, be fully brought out only by changing perspective somewhat. Having begun by asking what accounts for the empirically observed restriction of the labour contract to manual and lower-grade nonmanual occupations, I shall next ask why it should be that, in the case of professional, administrative, and managerial occupations, this form of regulation of employment would appear to be effectively precluded and is typically replaced by what has been called the service relationship. Again, I shall seek to give an answer in terms of employers' responses to the potential contractual hazards that are mapped out in Fig. 10.1 and, more specifically, as these intensify as one moves from the lower-left quadrant towards the upper right.

The rationale of the service relationship

The general problems of the employment contract, as outlined previously, are sometimes represented (e.g. Pratt and Zeckhauser 1984; Eggertsson 1990) as ones of a 'principal-agent' relationship: that is, of a relationship in which a principal (the employer) engages an agent (the employee) to act in the principal's interest in circumstances in which the principal cannot observe the agent's actions, nor share in all of the information guiding those actions. This representation may seem somewhat strained where it is possible for labour to be more or less commodified and some approximation to a spot contract is thus viable. However, it takes on special force where employees act in a professional, administrative, or managerial capacity.[12] Professionals are engaged to exercise specialized knowledge and expertise that they have obtained from a lengthy training; while administrators and managers are engaged to exercise the delegated authority of the employer. In both cases alike, therefore, the nature of the work tasks and work roles that are performed imply some asymmetry of information as between employer and employee

[12] It is in this connection of interest to note that, in the paper that in effect inaugurated transaction-cost economics, Coase (1937) draws attention to the legal distinction made in (then) current British law between an agent and a 'servant', which turns not on the absence or presence of a fixed wage or the payment of commission but rather 'on the freedom with which an agent may carry out his employment'.

and thus, for the latter, an area of autonomy and discretion into which monitoring by the employer cannot feasibly extend. Indeed, effective monitoring would here entail some kind of infinite regress. It would itself require precisely the kind of use of specialized knowledge and expertise and of delegated authority that creates the agency problem in the first place (cf. Simon 1991).

Where such difficulty in monitoring work arises, it then becomes especially important for the employer to gain the commitment of employees, which in turn implies designing and implementing a form of contract that can as far as possible ensure that their interests are, and remain, aligned with the goals of the organization as the employer would define them. In the case of profit-making organizations, one evident recourse is to link employee compensation to the economic success of the enterprise, as, for example, through stock awards or stock options or profit-related bonuses or other profit-sharing schemes. However, while these kinds of remuneration may often play a major part in the compensation of chief executive officers and other very senior personnel, they are difficult to extend at a similar level of importance throughout the staff hierarchy, and they would appear to have no very effective analogues in the case of employees in public-sector or non-profit-making bodies.[13]

Moreover, further circumstances may well obtain in which serious difficulties for any kind of performance-related pay system are created: that is, where employees are required—as professionals, administrators, and managers typically are—to carry out tasks, or roles, of a very diverse character. In such a situation, payment can scarcely be related equally to every aspect of the work that is undertaken. It will, rather, have to be linked to just one, or at most one or two, aspects—those for which performance indicators can most easily be devised being most likely to be chosen. But such arrangements hold dangers for employers. For the incentives offered will in this case serve not simply to induce greater effort on the part of employees but further to influence their *distribution* of effort, and of time and attention, *among* their different responsibilities. That is to say, those aspects of their work to which pay is in fact related will tend to be favoured at the expense of others, and to a degree that need not be optimal from the employer's point of view (cf. Holmström and Milgrom 1991). Employers may 'get what they pay for' in an all too literal sense (cf. Gibbons 1997). Moreover, given that work tasks *are* diverse, monitoring that is then specifically aimed at preventing such unintended consequences, or 'perverse effects', may well not be cost-effective, even if practical at all.

Since, therefore, in the employment of professionals, administrators, and

[13] Milgrom and Roberts (1992: 413) point out that, while in the late 1980s it was known that around 30% of all US firms operated profit-sharing schemes, the proportion of total employee income received through such schemes was very small, perhaps no more than 1%. Where close employee monitoring is not possible, profit-sharing and group-based bonus schemes are, of course, always likely to give rise to free-rider problems.

managers general principal-agent problems may often be compounded by the further ones posed by 'multi-task' agents, forms of employment contract in which either direct work monitoring or specific performance indicators provide the basis for pay would seem unlikely in the main to answer to employers' requirements. The alternative and generally more appropriate strategy will be for employers to seek to gain the commitment of their professional, administrative, and managerial personnel, or, in Simon's apt phrase (1991: 32), to shape their 'decision premises', through a form of contract with a quite different rationale—that is, one that relies on performance appraisal of only a broad and long-term, though perhaps comparative, kind and that then sets up, conditional upon such appraisal, the possibility of a steadily rising level of compensation throughout the course of the employee's working life. In this regard, the contract provides for compensation primarily through an annual salary that may be expected to increase both in accordance with an established scale and further, and more substantially, as the result of the employee's advancement through a career structure. Given the prospect of such an 'upward-sloping experience-earnings profile', as Lazear (1995: 39) has termed it, both effective incentives for employees *and* effective sanctions for the employer are created—and especially so if, as Lazear would argue (see also Lazear 1981), what is typically entailed is paying employees *less* than they are worth, in terms of their productivity, when they are young or at all events in the lower levels of the hierarchy and *more* than they are worth when they are older or in higher-level positions.

On the one hand, the better employees perform (or at least are perceived to perform) in the pursuit of organizational goals, the further, and more rapidly, they will be promoted out of the levels at which they are 'underpaid' and into those at which they are 'overpaid'. On the other hand, because for most employees higher rewards will still lie ahead, and rewards that they have already in part earned through their previous underpayment, 'hasty quits' are discouraged and the threat of dismissal, as, say, in the case of manifest shirking or incompetence or of malfeasance, becomes a more potent one. Furthermore, appropriately constructed pensions schemes can also be seen as an integral part of such 'deferred-payment' contracts, encouraging employees to stay with their organizations up to the peak of the expected present value of benefits but then discouraging them from staying on too long (in the absence of a mandatory retiring age) in the phase when, relative to their productivity, they are being overpaid (Lazear 1995: 42–5).[14]

In other words, the solution to the problem of agency, as it arises with professional, administrative, and managerial employees, is sought essentially in

[14] Empirical support for Lazear's theory can be found in Medoff and Abraham (1981) and Hutchens (1987). As Lazear notes (1995: 71), the upward-sloping experience-earnings profile would appear in many ways similar to an 'efficiency wage' or, as Marxists such as Wright (1997) would have it, a 'loyalty wage'. But, as Lazear also notes, a major difference is that 'deferred-payment' contracts need imply no breach of the principle of payment in accordance with marginal productivity—nor therefore involuntary unemployment; at the start of the earnings profile wages can be sufficiently low to clear markets.

the service relationship as this was described earlier. In place of any attempt at the immediate linking of performance and pay, the employment contract envisages, even if implicitly as much as explicitly, a quite diffuse exchange of service to the organization in return for compensation in which the prospective element is crucial; and, by the same token, the contract is understood as having a long-term rather than a short-term basis. The key connection that the contract aims to establish is that between employees' commitment to and effective pursuit of organizational goals and their career success and lifetime material well-being.[15]

So far in this subsection, I have concentrated on employers' contractual problems associated with work located towards the right of the horizontal dimension of Fig. 10.1: that is, those of monitoring. However, problems associated with work located towards the top of the vertical dimension are also relevant: that is, those of human asset specificity. I earlier suggested that these problems would be least demanding where employers could operate satisfactorily with workers who possessed no more than commonly available physical or cognitive capacities. In so far as a higher quality workforce is called for, the probability increases that there will be advantage to the employer in ensuring that any general purpose skills that employees bring with them are, as Williamson has put it (1985: 242), 'deepened and specialized' in the particular organizational contexts in which they are to be applied. And, where this is so, a form of contract that does nothing to help secure the employment relationship on a long-term basis will obviously be deficient: it will fail to provide incentives for employers to engage in the training, or for employees in the learning, from which they could alike benefit.

Problems of human asset specificity are not confined to professional, administrative, and managerial employees—as was indicated earlier and as will be seen further below. However, such problems may be thought to take on an importance in the case of these employees that is proportional to that of the organizational roles that they perform. That is to say, failure to provide for either the development or the retention of organization-specific skills, expertise, and knowledge on the part of professionals, administrators, and managers is likely to be especially damaging. Here again, then, the rationale of the service relationship, with its implication of continuing employment, is apparent. Through such a relationship, in contrast to one in which renewal

[15] The above differs from my earlier discussion of the service relationship (Goldthorpe 1982) chiefly in that it dispenses with the idea of 'trust' as being central to this relationship. I found my own increasing doubts in this regard expressed in a more cogent way than I was able to formulate myself in Williamson's critique (1996: ch. 10) of the concept of trust as applied in recent sociological work (e.g. Gambetta 1988; Coleman 1990). The gist of Williamson's argument (1996: 256) is that 'it is redundant at best and can be misleading to use the term "trust" to describe commercial exchange for which cost-effective safeguards have been devised in support of more efficient exchange. Calculative trust is a contradiction in terms.'

is entirely contingent, employers can more securely embark on costly training and 'planned-experience' programmes aimed at increasing employees' organization-specific abilities, and employees can in turn more securely devote the time and effort necessary to acquire such abilities. In other words, it becomes possible for both the costs of and the returns on investments of this kind to be shared between employers and employees (Milgrom and Roberts 1992: 363; Lazear 1995: 74). The advantage of the long-term character of the service relationship as the basis of a solution to agency problems—that is, via the rising worklife compensation curve—is therefore reinforced in that at the same time the basis for a solution to asset specificity problems is also provided.[16]

Mixed forms

Finally in this section some explanation should be attempted of the prevalence within certain occupational groupings—as indicated in Table 10.1—of mixed forms of the employment contract: that is, of forms that, in various ways, combine elements of both the labour contract and the service relationship.

Referring back once more to Fig. 10.1, a tendency may be supposed, in the light of the discussion so far, for empirical instances to be concentrated on the lower-left to upper-right diagonal: or, in other words, for there to be some correlation between the severity of the monitoring problems and of the human asset specificity problems to which different kinds of work give rise. None the less, instances lying off this diagonal can certainly be envisaged, and the suggestion I would advance is that it is in terms of work thus located

[16] In this connection, it is of some interest to note that Savage and his associates (Savage *et al.* 1992; Savage and Butler 1995) have argued, in critique of the idea of a 'service class', that a line of class demarcation should be seen as falling between professionals, on the one hand, and administrators and managers, on the other. This is on account of the different kinds of asset that they typically possess—in the case of the former, *cultural* assets, and of the latter, *organizational* assets. Professionals are then taken to be advantaged over administrators and managers in that cultural assets are less specific than organizational assets and therefore more 'storable': that is, more readily accumulated in transferable form. This analysis can be related to, and may perhaps enhance, that which I have made in suggesting that agency problems may be more acute than asset specificity problems among some kinds of professional—such as, say, those employed in less senior positions in health, education, and other state welfare services; while the converse situation may apply with some lower-level administrators and managers whose decision-making role is more local than strategic. But what, from the position I adopt, must tell against the claim that different class positions are thus created is the fact that the service relationship would appear an appropriate response to the contractual hazards that arise in *either* eventuality, even if certain features of this relationship are more fully expressed in some instances than in others. Moreover, Savage and his colleagues seem to miss the key insight of transaction-cost economics that asset specificity gives rise to *bilateral* dependency. Thus, in so far as the assets that administrators and managers acquire are of an organizationally specific kind, they may indeed be lost or much diminished if these employees lose their jobs; but, at the same time, there will be costs to employers who fail to retain employees with such assets. I have developed elsewhere (Goldthorpe 1995*a*) a more general critique of the 'assets theory' of the differentiation of the service class (see also Li 1997).

that the occupational distribution of mixed forms of the employment contract is to be explained. Work falling in the lower-right quadrant of Fig. 10.1—that is, work that confronts employers with real difficulties of monitoring but not of asset specificity—could be expected to lead to a form of contract in which features entailing some departure from the exchange of discrete amounts of money and effort characteristic of the labour contract would be more apparent than ones directed towards furthering a long-term relationship. And, conversely, work falling in the upper-left quadrant, where asset specificity problems are serious but not monitoring problems, could be expected to lead to a form of contract in which a fairly specific money-for-effort exchange is preserved but some understanding of the long-term nature of the contract is at least implied.

Some empirical support for this argument can in fact be provided, and indeed for a rather more precise version of it: namely, one that would associate the first of the two situations outlined above primarily with routine nonmanual occupations in administration and commerce that exist, so to speak, on the fringes of bureaucratic structures, and the second situation with manual supervisory and lower-grade technical occupations.[17] Thus, routine nonmanual employees—clerical workers, secretaries, and so on—would appear to enjoy fixed salaries and also relatively relaxed or flexible time-keeping arrangements almost to the same extent as the professionals, administrators, and managers with whom they typically work in ancillary roles, although for the most part deploying only rather standardized skills (cf. Fama 1991). However, they are not to the same extent involved in career structures within their employing organizations that would hold out the prospect of steadily increasing rewards over the entire course of their working lives.[18] Conversely, supervisors of manual workers and lower-grade technicians tend to be paid weekly, with overtime pay and other adjustments being based on the monitoring of hours worked, in much the same way as many of the rank-and-file employees over whom they exercise direction and control. But their distinctive value to their organizations is more often recognized through agreements or understandings on employment and income security and also perhaps through opportunities to progress up 'job ladders' that are based primarily on seniority—that is, in the manner provided for by the institutions

[17] Some findings of the validation exercises referred to in the text above and n. 2 point to such a conclusion, as also do unpublished data on employment relations at the level of basic ONS occupational groups resulting from the work of the ESRC committee described in n. 2, and kindly made available to me by Karen O'Reilly.

[18] From this same point of view (as has been pointed out to me by Jerker Denrell), it is then not surprising that secretarial and similar workers should often be engaged on a temporary basis via employment agencies. A further occupational group that is of interest in this connection is that of sales personnel. With these employees, payment on the basis of a fixed salary *plus* commission is often found. The rationale here would appear to be that while quantity of output, that is, sales made, can rather easily be measured, thus allowing piece-rate payment, problems of monitoring 'quality' do also arise, notably as regards maintaining the firm's reputation with its customers. I would, therefore, hypothesize that it is with the employees of firms having least to lose in this respect—for example, firms selling double-glazing or door and window frames—that the largest commission element will occur.

of the classic 'firm-internal' labour market (Doeringer and Piore 1971; Osterman 1987).[19]

Taking this argument together with those previously advanced in this section of the chapter, it is then possible to give an overall representation of how the analytical dimensions of Fig. 10.1 are seen to relate to the empirical regularities implied in Table 10.1: that is, regularities in the association between different occupational groupings of employees and the form of regulation of their employment. This is done in Fig. 10.2. On the understanding that the regularities in question are to be regarded as only probabilistic, and that exceptions to them will thus certainly be found, this latter figure can stand as a summary of the explanation that is here offered of how they are in fact generated and sustained through employers' responses to the problems of contractual regulation that arise from the engagement of workers to undertake differing kinds of work.

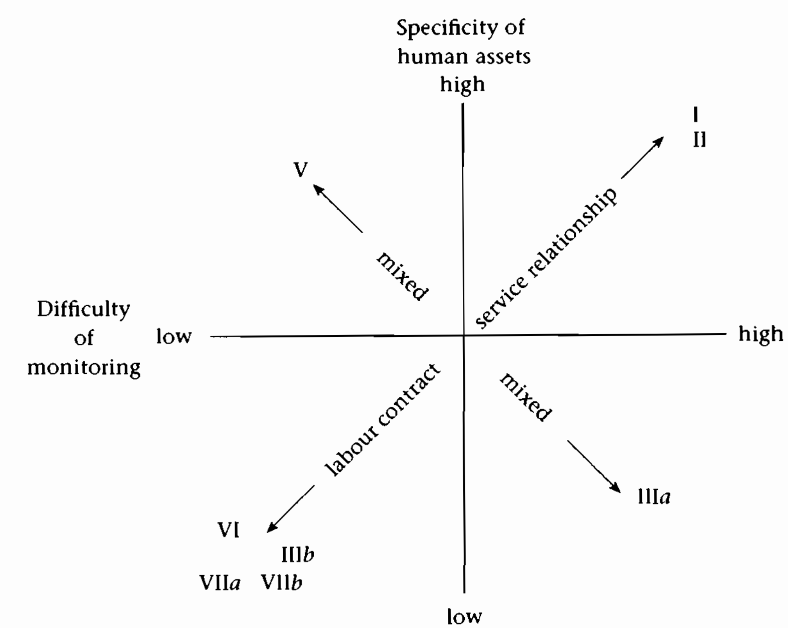

FIG. 10.2. Dimensions of work as sources of contractual hazard, forms of employment contract, and location of employee classes of the schema

[19] It should, however, be noted that the job ladders in question are typically much shorter than those available to professional, administrative, and managerial staff. Cf. the distinction made by Osterman (1987) between the 'industrial' and 'salaried' forms of internal labour markets.

New Employer Strategies and the Future of the Service Relationship

One objection that might be raised against the foregoing analysis—and to which a response should at least be outlined—is that it seeks to explain in very general theoretical terms a pattern of differentiation in employment contracts that could well prove specific to a particular historical era. Several authors have indeed already broached the question of the continuing viability of the service relationship under conditions of rapid change in technological and market conditions and of intensifying global competition that impose ever greater requirements of organizational 'flexibility' (e.g. Halford and Savage 1995); and others have gone further in maintaining that under these conditions the service relationship is actually being eroded, and in the public as well as the private sector as competitive tendering and other forms of 'market discipline' are increasingly imposed (e.g. Brown 1995). There is constant pressure, it is argued, to 'downsize' or 'de-layer' management structures, to buy in professional services rather than to provide them in-house, to engage staff on fixed-term contracts, and to introduce performance-related pay systems at all levels of employment. From this point of view, then, the service relationship appears not as a form of contract with a rather sophisticated underlying rationale but as an expression merely of a conventional status distinction that could be sustained as an aspect of 'organizational slack' during the long boom of the post-war years but that is now being swept aside in a far more demanding economic environment.[20]

Direct evidence of the undermining of the service relationship is in fact still hard to come by; and what is sometimes taken to be indirect evidence—for example, that of more frequent job losses among professional, administrative, and managerial employees or of their greater mobility between employing organizations—would appear open to more than one interpretation (see Gallie *et al.* 1998). None the less, in so far as attempts *are* under way radically to reshape the form of regulation of employment of these personnel, a valuable opportunity is afforded for the empirical testing of the account I have offered of the logic of the differentiation of employment contracts; and further clarification of this account may then be achieved by considering what would, from the point of view of empirical evaluation, have to be regarded as centrally at issue.

To begin with, it is entirely consistent with the general position that I have taken that employers *should try* to exploit any changes in labour market or

[20] A somewhat related argument is that advanced by Sørensen (1999; cf. also 1994, 1996) to the effect that, under competitive pressures, employers will seek to reduce problems of human asset specificity by the redesigning of jobs to allow them as far as possible to be carried out on the basis of general purpose skills. Sørensen sees the 'composite rents' to which asset specificity gives rise as a particular problem for employers, since the division of these rents can be settled, in Alfred Marshall's words, only by 'higgling and bargaining'—through which employers have no guarantee that they will get what they would regard as their fair share.

other economic conditions that might enable them to modify contracts of employment, explicitly or implicitly, in ways that would be to their advantage or, more specifically, that would reduce their contractual hazard. And it is further consistent that such modifications should then, all else being equal, be ones leading away from the service relationship and towards the discrete and short-term exchange of the labour contract; or, in other words, away from forms of the regulation of employment that presuppose a diffuse and continuing exchange and towards ones in which labour is to a greater degree commodified. As Breen (1997*a*) has observed, one can in this respect think of employers as seeking to transfer risk from themselves to their employees: that is, to free themselves from the inflexibility entailed by the 'quasi-generalized reciprocity' of the service relationship and to secure instead an 'asymmetric commitment' or in effect an *option* on the supply of labour, which they can then decline, if necessary, in order to avoid 'downside' risk while preserving the possibility of profiting from 'upside' risk. Evidence simply of employers being alert to the possibility of revising forms of contract in ways they would see as being in their interests is not, therefore, in itself, of any great consequence. What matters is how far employers are thus led actually to abandon the service relationship in cases where it had previously applied.

In this connection, it has also to be recognized that some of the strategies that employers may pursue in search of greater flexibility need have little or no impact on the service relationship *per se*, and indeed may even help to make this relationship *more* viable. Thus, by downsizing and de-layering management structures and also by buying in professional services, employers can hope to reduce the proportion of their total workforce to whom the service relationship is extended, and in turn the degree to which they are involved in 'quasi-generalized reciprocity'. Likewise, by creating greater flexibility in the employment of *other* grades of worker—as, say, by modifying features of internal labour markets for skilled manual workers (cf. Capelli 1995)—they may be better able to sustain the service relationship in the case of those employees for whom they would see it as specifically appropriate. In some influential models of the 'flexible firm' (e.g. J. Atkinson 1985) the emphasis is in fact on the *divergence* between the employment relations that apply with the 'core' and with the 'peripheral' workforce.[21]

What, therefore, emerges as the key question is that of how far in prevailing economic circumstances, and on what basis, employers do come to regard the service relationship itself as expressing a form of contract that they should aim in general to terminate rather than to preserve. And in so far as this relationship is in fact undermined in the case of professional, administrative, and managerial staff, as, say, through the introduction of short-term contracts or

[21] De-layering could perhaps make the service relationship more difficult to implement in that promotion opportunities would be reduced; but it would still be possible to maintain the upward-sloping earnings-experience profile by allowing incremental salary scales or appraisal-related pay increases to operate *within* given hierarchical levels (cf. Lazear 1995: 79–80).

performance-related pay, it will in turn be important to know just how employers then seek to handle those problems of work monitoring and of human asset specificity out of which, I have argued, the rationale of the service relationship and of its occupational range initially arise.

Breen, for example, has pointed out (1997*a*) that, despite employers' concern to offload risk within the employment contract onto their employees, there are still good grounds for supposing that the service relationship will prove durable since there is no other obvious solution to the agency problem deriving primarily from asymmetric information. Professionals, administrators, and managers are employees in regard to whom it is generally less important that the employment contract should provide for flexibility (from the employer's point of view) than that it should ensure that the employee has strong incentives to show commitment to organizational goals (see also Gallie *et al.* 1998: 312–13). And to this I would add that, even if improved techniques of monitoring may in some instances be capable of reducing agency problems (Halford and Savage 1995: 129), attempts at basing pay on performance will still threaten to give rise to 'perverse effects' in the way earlier noted where employees are engaged in work tasks of a multi-faceted kind.[22] In this respect, the requirements of contractual flexibility would seem to come into direct conflict with those of 'functional' flexibility (J. Atkinson 1985), which in fact lead to demands on employees to be ready to take on an ever-wider range of tasks and responsibilities.

What is perhaps more plausible than the idea of the general abandonment of the service relationship is the suggestion that, at least with some employee groupings, the 'deal' that it comprises may be reformulated: in particular, so that what the employee is offered, in return for commitment and a readiness to develop organization-specific assets, is an understanding not on continuity of employment but rather on continuity of *employability*. In this case, it falls to the employer to provide employees with training and experience that, as well as enhancing their organization-specific skills, expertise, and knowledge, will also equip them for future career progression in the external labour market. However, it is still far from clear that even this modification of the service relationship would necessarily be to the employer's advantage. The evident risk that the employer incurs is that, if investment is made in the development of human assets that are *not* organization specific and if the employees who benefit from this investment are not then retained, the returns on the investment will largely be lost: they will be divided between the employees and their *subsequent* employers. Again, then, the force of the original rationale of the service relationship is brought out.

In sum, for a compelling argument to be advanced that the service relationship is in general decline, two things would seem to be required: first, direct evidence—which, to repeat, is so far lacking—that employment contracts expressing this relationship are indeed being discontinued across the range of employee groupings for which they were previously typical; and, secondly, evidence—the need for which seems not even to have been appreciated—that such a change can be regarded as permanent, rather than being, say, a response merely to short-term economic exigencies or to the current vogue among management consultants, *because* the logic that previously underlay the service relationship has now ceased to apply or has in some way been transcended.[23] The analysis that I have earlier presented obviously leads me to the view that evidence of the latter kind at least will not be readily forthcoming.

[22] As Holmström and Milgrom (1991) stress, the ultimate source of difficulty here lies in the diversity of tasks performed and in finding a way of appropriately relating pay to all of them, rather than in work monitoring and measurement *per se*.

In this chapter I have started from an empirical regularity that has emerged from attempts to assess the criterion validity of the class schema that colleagues and I have developed as a research instrument. At least for the British case, a pattern of association has been established, among employees, between their broad occupational grouping and the form of regulation of their employment. I have then suggested, drawing chiefly on theoretical ideas developed in both more and less orthodox branches of modern economics, how this regularity might be explained at the level of social action: that is, in terms of employers' rationally intelligible responses to the different problems that they face in devising and implementing employment contracts for workers engaged in different types of work—in particular, problems of work monitoring and of human asset specificity. I have also indicated how I believe that attempts at testing the explanation offered, within the contemporary economic context, might best be focused. To end with, I make two observations concerning what is and what is not implied by the central argument of the chapter for the more general understanding of the stratification of modern societies—on the assumption, of course, that the argument is basically sound. Since this assumption is, for the present at least, obviously open to challenge, the remarks will be brief, although they relate to large questions.

First, if the analysis I have presented is valid, it must follow that modern societies, at least in so far as they retain capitalist market economies of some kind, will have a relatively complex class structure as one of their concomitant and abiding features. In addition to the differentiation of employers, the self-employed and employees, the latter will themselves be differentiated in terms of the employment relations in which they are involved as a result of a highly generalized 'situational logic' that applies, and that will have its effects, across a wide range of institutional and cultural contexts. The further implication then is that, although such societies may well show variation in the historical evolution of their class structures—that is, in the rates of growth or decline of different classes and thus in their proportionate sizes—they will at the same time be characterized by a similar pattern of class-related social

[23] It might be added here that satisfactory evidence of the kind in question could not be derived merely from *ad hoc* case studies. Research based on representative samples both of employing organizations and of their employees would seem essential (see further Kalleberg 1990).

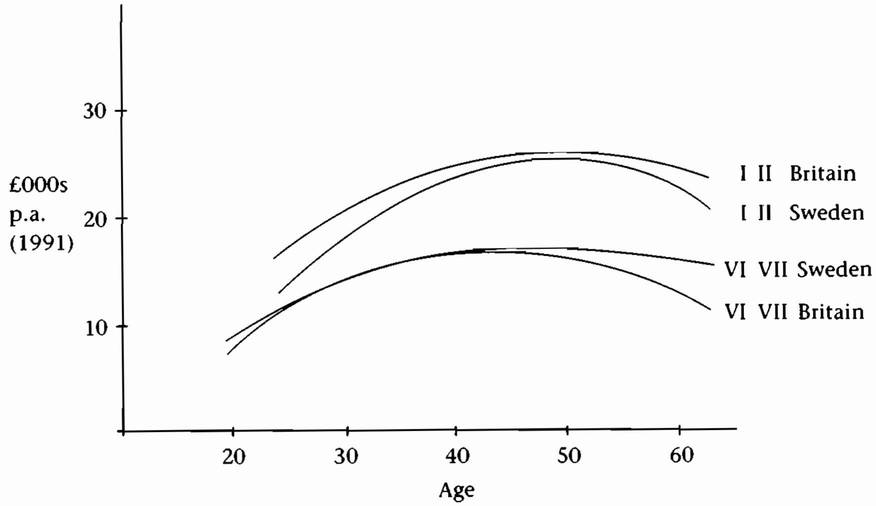

FIG. 10.3. Age-specific earnings curves (smoothed) for Classes I and II and VI and VII, Britain and Sweden, men in full-time employment
Sources: Britain: *New Earnings Survey* (1992); Sweden: Level of Living Survey, 1990 (unpublished data).

inequalities. The form of these inequalities is likely to be no less complex than the structure of class relations from which it derives, and therefore not capable of adequate representation in any one-dimensional fashion. None the less, it may still be said that, in so far as employees are involved in a service relationship rather than a labour contract, they will be advantaged not only (in most cases) in their general level of income from employment but further in that their incomes will be less subject to interruption or short-term fluctuation and will tend to follow a rising curve over the larger part of their working lives.[24]

Secondly, though, it has to be recognized that the analysis advanced carries no particular implications for the actual *degree* of inequality in rewards from

[24] It is in the case of intermediate classes that the problems of any one-dimensional ordering become most apparent. For example, while employees in manual supervisory and technical positions may have higher average earnings than employees in routine nonmanual work, their earnings are likely to show more short-term variation; small employers and self-employed workers are generally exposed to more insecurity than employees but have better chances of accumulating capital, and so on. It has sometimes been seen as a disadvantage of the class schema that my colleagues and I have developed that its categories cannot be fully ordered in any unambiguous way. However, studies of its construct validity have shown its capacity to reveal *both* marked class differences in regard to political partisanship, educational choice, health, etc., as between the service class or salariat (Classes I and II) and the working class (Classes VI and VII) *and* at the same time more subtle but still intelligible differences in these respects involving the intermediate classes (Classes III, IV, and V) of a kind that would often have been obscured by one-dimensional measures of inequality such as synthetic 'socio-economic status' scales.

work (let alone in living standards more generally) that will be found among the members of different classes. There is indeed evidence to suggest that in this respect some significant cross-national variation can coexist with essentially the same pattern of class differentiation. An illustrative example is given in Fig. 10.3. From this it can be seen that in both Britain and Sweden age-specific earnings curves for (full-time male) employees in service-class and working-class positions show the same characteristic shapes: for the former, rising strongly up to around age 50—that is, revealing the expected 'upwards sloping profile'; for the latter, flattening out after age 30. However, at all ages, the gap between the two curves is narrower in the Swedish case than in the British.

An adequate explanation of this latter kind of finding would then need to be pursued on quite different lines from those that I have followed in the analyses above. Rather than being of a highly generalized character, such an explanation would have to give serious consideration to the effects of nationally specific features of, for example, industrial structure, trade union organization, collective bargaining institutions, and public policy, and might thus in the end take on far more of a historical than a sociological style (cf. Chapters 2 and 3). As will be seen in the chapter that follows, such a distinction between kinds of *explananda* and appropriate styles of explanation is of direct relevance when the attempt is made to account for further notable regularities—and deviations—that are by now fairly well established in the degree to which the class structures of modern societies display openness or fluidity: or, that is, in the degree to which individuals of differing class origins have equal chances of experiencing social mobility or immobility within these structures.