Organizing Flexibility: The Flexible Firm in a New Century

Arne L. Kalleberg

Abstract

Research on organizational flexibility should examine the linkages between numerical and functional flexibility. Unfortunately, studies of each type of flexibility generally neglect the other. Moreover, the most popular conception of the interplay between these two forms of flexibility — the core–periphery model — is incomplete in important ways. I discuss evidence and limitations of the core–periphery model of the flexible firm, and outline some promising attempts to conceptualize how organizations may combine functional and numerical flexibility. I focus mainly on the USA and the UK, though I also review evidence and issues involved in cross-national differences in organizational flexibility.

1. Introduction

Social and economic changes in the past quarter-century have underscored the need for organizations to have greater flexibility in their employment systems. Observers in all industrial countries regularly emphasize the importance of human resource management practices that enable organizations to adapt quickly to rapid developments in technology, greater diversity in labour markets, growing international and price competition in product markets, and corporate financial restructuring in capital markets. A popular expression of this concern has been the idea of the ‘flexible firm’, which denotes the kinds of organizational forms that enable employers to obtain the flexibility they need.

Researchers have emphasized two distinct strategies of flexible labour utilization: enhancing employees’ ability to perform a variety of jobs and participate in decision-making, and reducing costs by limiting workers’ involvement in the organization. These two strategies have been referred to...
variously as: functional v. numerical flexibility (Atkinson 1984; Smith 1997; Hunter et al. 1993), internal v. external flexibility (Cappelli and Neumark 2001), clan v. market (Ouchi 1980), dynamic v. static flexibility (Colelough and Tolbert 1992; Deyo 1997), and organization-focused v. job-focused employment relations (Tsui et al. 1995).

Recent research on organizational flexibility has proceeded along two relatively distinct streams, each focusing on one of these two flexible labour utilization strategies. The majority of studies emphasize the correlates of ‘high performance work systems’ that are believed to enhance functional or internal flexibility; the other group examines processes of externalization designed to reduce costs and provide organizations with numerical flexibility. Unfortunately, relatively few studies have examined the interplay between functional and numerical flexibility. Consequently, we know comparatively little about the relative costs and benefits associated with organizations pursuing one as opposed to the other form of flexibility or a combination of the two. In particular, insufficient attention has been paid to whether the benefits to employers of high-performance work systems are shared with their employees in a ‘win-win’ or ‘mutual gains’ enterprise, or whether there are disadvantages to one or more parties to attaining functional flexibility.

Those studies that have examined the interrelations between functional and numerical flexibility in organizations have tended to conceptualize this relationship primarily in terms of a relatively simple ‘core–periphery’ model (e.g. Atkinson 1984; Mangum et al. 1985: 599–601; Osterman 1988: 85–89; Olmsted and Smith 1989; Handy 1990: 87–115). The debate about this core-periphery model has focused mainly on whether it is an accurate representation of employers’ labour utilization strategies. This relatively narrow concern has hindered the development of more refined conceptualizations of the varieties of ways in which the two flexible forms of work organization may be related and of the patterns of conflict and accommodation created by employers’ attempts to balance the advantages and disadvantages associated with them.

In this paper I argue that advancing our knowledge of organizational flexibility depends on our ability to develop models of the relationship between functional and numerical flexibility, and to specify the conditions under which an organization will establish various combinations of standard (i.e. regular, full-time) and non-standard (e.g. part-time, temporary and contract work) employment relationships. Accomplishing these goals will facilitate the examination of trade-offs associated with the various labour utilization strategies and their correlates and consequences for organizational, worker and industrial relations outcomes.

I first critically evaluate research on each of the two forms of flexibility in terms of their failure to consider the other. I then review briefly some of the evidence on the core–periphery model of the flexible firm, and outline some promising attempts to conceptualize employment systems that overcome some of its simplistic assumptions about how organizations may combine
features of both functional and numerical flexibility. I conclude with a discussion of some key elements of a research agenda for the new century that is needed to understand better employers’ labour utilization strategies. My focus is mainly on organizational flexibility in the relatively unregulated economies of the United States and United Kingdom, though I also review evidence from other countries and discuss explicitly some issues involved in cross-national differences in organizational flexibility.

2. Studying functional and numerical flexibility: unidimensional approaches

I first provide a brief overview of the two main research traditions on organizational flexibility, as a prelude to discussing the interrelations between them.

Functional Flexibility and High-Performance Work Organizations

A great deal of attention has been paid in recent years to forms of work organization and human resource management practices that are designed to provide employees with skills, incentives, information and decision-making responsibility that improve business performance and facilitate innovation. The pervasiveness of these studies has led some to suggest that a focus on such new work and human resource management practices constitutes a ‘new paradigm’ that is replacing unions and collective bargaining as the core innovative force in industrial relations research (Godard and Delaney 2000). These new practices have been called, variously, high-performance work organizations (Appelbaum et al. 2000; Osterman 2000), transformed work organizations (Osterman 1994), flexible or alternative workplace practices (Gittleman et al. 1998), employee involvement systems (Cotton 1993), flexible production systems (MacDuffie 1995), progressive human resource management practices (Delaney and Huselid 1996), high-commitment systems (Walton 1985) and high-involvement management (Lawler 1988; Wood 1999).

Common to these labels is the notion that the hierarchical system of control characteristic of Taylorist or Fordist forms of production has been replaced by a type of work organization that empowers workers to participate in decision-making, enables them to work in teams and enhances their commitment to the organization by, among other things, linking their compensation to the organization’s performance. These mechanisms enhance an organization’s functional flexibility by providing employees with multiple skills so that they can be redeployed relatively quickly from one task to another (see the reviews by Appelbaum and Batt 1994; OECD 1999; Wood 1999). Moreover, studies of high-performance work organizations generally focus on work done by regular, full-time workers such as those in
the organization’s ‘core’ occupation (e.g. Osterman 1994, 2000). For these workers, high-performance work organizations are generally assumed to result in mutual gains for both the firms and their workers (especially those in the ‘core’ of the organization) (Kochan and Osterman 1994; Appelbaum et al. 2000).

By emphasizing the work done by regular, full-time workers, studies of high-performance work organizations have tended to neglect flexible staffing strategies involving non-standard forms of employment. As a result, these studies generally do not consider explicitly whether a necessary condition for securing the benefits of stable, functionally flexible employment for a group of highly committed workers is a demand buffer of external flexibility provided by a group of temporary and other non-standard workers, who often have dead-end, insecure and low paid jobs and relatively little commitment to the organization (Atkinson 1984: 31; Goldthorpe 1984: 329–35; Mangum et al. 1985: 609; Osterman 1988: 85; Kyotani 1999: 182–4). Rather than a ‘win–win’ scenario, then, the use of high-performance work practices may result in disadvantages to some, which are rooted in the increased segmentation and polarization between advantaged, regular full-time (‘durables’) and disadvantaged, non-standard labour force members (‘disposables’) (Hyman 1988: 55–7; Vallas 1999: 93–5).

The emphasis of studies of high-performance work organizations on managerial outcomes, a harmony of interests and mutual gains has suppressed attention to conflicting interests and diverted attention from social issues related to workers’ labour market rights and opportunities (Godard and Delaney 2000: 492). Restricting the analysis primarily to regular, full-time committed workers has obscured the fact that achieving functional flexibility may not necessarily require the sole use of such workers, as suggested by Smith’s (2001: 17–20) study of how temporary workers can help the organization obtain functional flexibility and by Drago’s (1996) discussion of ‘management by fear’ in the low-job-security environment characteristic of the ‘disposable workplace’.

**Numerical Flexibility and Externalization**

A second strand of research on organizational flexibility focuses on employers’ attempts to obtain numerical flexibility and otherwise reduce costs by using workers who are not their regular, full-time employees. Forms of externalized labour have been called, variously, flexible staffing arrangements (Christensen 1989; Houseman 2001), market-mediated work arrangements (Abraham 1990; Abraham and Taylor 1996), contingent work (Polivka and Nardone 1989; Blank 1998; Barker and Christensen 1998), and non-standard work arrangements (Kalleberg et al. 1997; Felstead and Jewson 1999; Cousins 1999; Yeandle 1999). A growing literature has sought to document and explain the growth of organizations’ use of flexible staffing arrangements (e.g. Pfeffer and Baron 1988: 284–92; Davis-Blake and Uzzi 1993: 199–206; Houseman 2001; Kalleberg et al. 2001).
An organization’s externalized work-force includes several kinds of non-standard employment relations. Organizations can limit the duration of employment through the use of short-term temporary workers who are hired for finite periods on an as-needed basis. They can also obtain numerical flexibility and often reduce costs by externalizing administrative control through the use of temporary help agency or contract workers. These workers are considered to be employees of the temporary agency or contract company, not the client organization. They include both highly skilled (e.g. consultants and independent professionals) and low skilled (e.g. clericals, food service) workers. Additionally, organizations can also limit the duration of employment by using part-time workers, though part-timers are often relatively permanent members of the organization and so may be considered as part of its core work-force.

Data on trends in organizations’ use of various types of non-standard work arrangements are sparse, making it difficult to adjudicate debates about the extent to which their use reflects a fundamental transformation in employment relations. In the United States, there has been a rapid growth in temporary help agency employment since the early 1970s, though it still constitutes a small proportion of the labour force (e.g. Segal and Sullivan 1997: 118–19). Some writers (e.g. Cappelli 1999) have argued that these non-standard work arrangements represent a ‘new deal’ between employers and employees in which market mechanisms have replaced hierarchical employment systems and wherein notions of loyalty and the social contract, especially for highly skilled workers, are obsolete. Others have maintained that these changes reflect more minor changes in the allocation of risk from employers to employees in light of changing economic conditions (Jacoby 1999). Similarly, British studies have raised questions about whether the use of flexible staffing arrangements is a new phenomenon. Marginson (1991: 35–42) found a great deal of continuity in the use of non-standard work in many large UK companies in the 1980s (see also Casey 1991: 194–8), while Atkinson (1987: 90–1) argued that subcontracting increased in Great Britain in the early 1980s (though Pollert, 1988: 289–91, concluded that there was little increase in subcontracting in the private sector).

Studies of numerically flexible staffing arrangements have generally neglected to examine the impacts of these arrangements on an organization’s regular workers, leaving us with an incomplete view of the disadvantages associated with using non-standard work. Human resource management strategies that try simultaneously to minimize labour costs via contingent workers and to increase commitment and motivation of core workers may not be successful, since management’s recruitment of temporary employees may give rise to conflict and tension between temporary and permanent workers (Geary 1992: 259–61, 267). Nonetheless, using temporaries may enhance the internal job mobility of relatively high skilled, regular workers by reducing the number of such permanent employees who are competing for promotions, as shown by Barnett and Miner (1992: 272–4).
3. Combining functional and numerical flexibility: the core–periphery model

A third group of studies of organizational flexibility have considered the interrelations between functional and numerical flexibility, and have sought to explain how organizations are able to obtain simultaneously these seemingly contradictory forms of flexibility. Three decades ago, Doeringer and Piore (1971: 173) observed that employers in the USA could use subcontracting and temporary work to shift demand into the secondary sector, thereby providing manning flexibility in the primary enterprise. Writers have used various labels to refer to firms’ labour utilization strategies that combine numerical and functional flexibility, including: ‘core–periphery’ model (Atkinson 1984; Osterman 1988; Harrison 1994; Drago 1998), ‘core/ring’ configuration (e.g. Olmsted and Smith 1989), ‘shamrock’ organization (Handy 1990: 87–115), two-tier organization (Christensen in Doeringer et al. 1991), and the ‘attachment–detachment’ model (Mangum et al. 1985).

Perhaps the most influential model that combines functional and numerical flexibility is Atkinson’s (1984, 1987; see also Atkinson and Meager 1986: 3–5, 72–4) ‘core–periphery’ or ‘micro dual labour market’ model (Pollert 1988: 283), which has been the subject of lively and often critical debate (e.g. Pollert 1988; Wood 1989: 4–9; Hakim 1990; Penn et al. 1992: 215–16; Hunter et al. 1993; Procter et al. 1994; Cappelli 1995). This model offered managers and government policy-makers a framework for identifying the main practices on which they should focus in order to obtain both functional and numerical flexibility. It suggested that they seek to establish long-term employment relations with part of their work-forces (the ‘core’, regular, permanent workers who are highly trained, skilled and committed to the organization, attributes that are thought to be needed for functional flexibility) at the same time as they externalize other activities and/or persons by means of transactional contracts. Segmenting the organization’s work-force into fixed and variable components is assumed to achieve cost effectiveness, as the numerically flexible, non-standard, ‘peripheral’ workers are used to buffer or protect the regular, ‘core’ labour force from fluctuations in demand, thereby avoiding the morale problems engendered by laying off regular employees and the disequilibria (and illegalities in some countries) associated with treating regular workers differently.

The attractiveness to managers and organizational researchers of the imagery associated with the core–periphery model results largely from its combining various features of management theory and fashions. These include: established management thinking, such as Thompson’s (1967) idea that organizations should buffer or protect the capabilities and resources that are most vital to their competitiveness; strategies such as the need for an organization to define strategically its core competence (see Pollert 1988: 304; Pfeffer 1994); and the popularity of Japanese human resources management (Koike 1978). The model has been thought to be especially applicable to the USA and UK, where labour laws leave employers relatively free to
choose and vary their employment contracts, compared with many other European countries (Hakim 1990: 161).

Research Evidence

Several assumptions of the core–periphery model have been the subject of considerable empirical research, particularly the ideas that organizations utilize both functional and numerical flexibility simultaneously and strategically. There is little hard, direct systematic evidence on these assumptions, though some studies have examined them indirectly. Most of them refute the assumptions, though a few support them.

(a) Do organizations combine functional and numerical flexibility?

Some studies have found a negative relationship — or no relationship — between functional and numerical flexibility within establishments. Osterman (1999: 102–7; 2000: 193–5) found that organizations that made greater use of ‘high-performance’ practices in core occupations were significantly less likely than non-‘high-performance’ organizations to make use of contingent employees and to outsource work (though the former still made use of both practices) in those occupations. He interpreted this as evidence that it is very difficult for organizations to implement high-performance work practices if they employ contingent or temporary workers who will be around for only a short time. Similarly, Davis-Blake and Uzzi (1993: 210–11; 217) found that firms which were more likely to establish long-term relations with employees were less apt to externalize work: their analysis showed that establishment-level predictors of internalization (such as extent of employment bureaucratization, and the degree to which firms face relatively oligopolistic product markets and stable demand) affected externalization in the opposite direction. And Cappelli and Neumark (2001) found that the use of contingent work by US establishments was positively and significantly related to involuntary turnover among the establishment’s permanent work-force, even when they controlled for the standard economic determinants of involuntary turnover and a variety of establishment and worker variables. This finding suggests that contingent work and involuntary turnover are complementary practices, not substitutes, thereby contradicting the core–periphery hypothesis that employers protect their core work-force from involuntary turnover by obtaining numerical flexibility from their peripheral workers.

In Britain, Cully et al.’s (1999: 34–8) analysis of data from the 1998 Workplace Employee Relations Survey found that there was a negative relationship between the various non-standard work arrangements and functional flexibility at the workplace level. In addition, a comparative study of flexibility within establishments in the Scandinavian countries (reported in Nutek 1999) found that indicators of numerical and functional flexibility were only weakly correlated (both positively and negatively) in Norway, though sometimes these correlations were statistically significant.
(see also Olsen 1998). In Finland, functionally flexible workplaces in the private sector with ten or more employees were slightly more likely to use fixed-term contracts at all than traditional (non-functionally flexible) workplaces (46 v. 42 per cent, respectively), but the proportion of temporary employees within the organization was greater in traditional than in functionally flexible workplaces (14 v. 7 per cent, respectively). Moreover, Penn et al.’s (1992) interviews and observations in paper mills found that forms of numerical and functional flexibility were ‘in serious conflict with each other’ (p. 220); they concluded that subcontracting and multi-skilling may be used as alternatives for each other, but not as supplementary practices. Finally, Cappelli’s (1995: 589–91) review concluded that there is little evidence that the increase in numerical flexibility in the United States, Australia, Britain and New Zealand have been driven by managers’ desire to buffer their core employees.

By contrast, other studies have found that patterns of internalization and externalization coexist within the same organization (though these concepts are not always measured in the same way across studies). Japan has long been regarded as the exemplar of the core–periphery model, and Japanese employers have sought to preserve the benefits both of internalized human resources and of cost reduction and deployment flexibility by using externalization and internalization strategies simultaneously. For example, Morishima (1995: 12–13) found that in the Japanese food service sector there was a positive relationship between the internalization of the establishment’s permanent work-force (reward for long service, emphasis on internal training relative to external experience and training, and paying above-market compensation to retain permanent employees) and the proportion of part-time and temporary employees within the establishment.

In the United States, Lautsch (1996) found that whether or not an organization had a ‘mutual gains internal labour market’ (which was defined by the presence of practices such as performance pay and profit-sharing, teams, rotation, job security, cross-training and discretion) was positively related to the proportion of direct-hire temporaries and temporary help agency employees within the establishment. Kalleberg’s (1989) survey of high-growth businesses in North Carolina found that companies subcontracting out work were also more likely to promote most of their employees from within the company, an indicator of firm internal labour markets. Similarly, Cully et al. (1999: 34–8) found that functional flexibility was positively related to subcontracting, which is consistent with the argument that employers externalize non-core activities and internalize core functions. A recent study of ‘high performance’ and flexible staffing practices in a representative sample of establishments in the United States (the Second National Organizations Survey — see Kalleberg et al. 1999) provides evidence that in the majority of US establishments numerical and functional flexibility are positively related. Establishments were categorized into four groups based on a cross-classification of whether they used functional and numerical flexibility strategies. Over half of the establishments were
classified as having high numerical flexibility (they used either direct hire temporaries, temporary help agencies or contract companies) and high functional flexibility (they use ‘high-performance’ work practices such as teams and job rotation) (Kalleberg 2001: 6–7). Moreover, Pinfield and Atkinson (1988: 19) found evidence that the Canadian firms in their sample combined numerical and functional flexibility strategies.

A secondary analysis of data from surveys conducted in ten European countries as well as Australia and the United States (OECD 1999: 207–9) provided mixed support for the proposition that functionally flexible practices (such as flattening of management structures, greater involvement of lower-level employees, teams and job rotation) and numerically flexible practices (outsourcing and subcontracting, downsizing, temporary contracts) are positively related. In the European countries, the use of four functionally flexible practices was positively related to the rise (within the last three years) in proportion of temporary contracts and in subcontracting, but using fewer than four practices was not. Workplaces adopting one, two or three (but not four) functionally flexible practices were also more likely to have a rise in the proportion of part-timers within the previous three years. In Australia and the United States, the bulk of the evidence suggested that workplaces with functionally flexible practices had a higher proportion of full-time workers.

An analysis of two Canadian manufacturing plants by Grenier et al. (1997) illustrates the complexity of the relationship between functional and numerical flexibility. They showed that whether or not it is possible to enhance functional and numerical flexibility simultaneously depends on the ways in which these strategies are implemented. They found that varying the volume of labour through internal means (overtime) did not threaten the job security of workers and thus facilitated functional flexibility; on the other hand, the attainment of functional flexibility was hampered in the plant that sought to adjust labour volume by layoffs and external strategies such as outsourcing certain activities.

(b) Employers’ strategic use of the core–periphery model
Another basic assumption of the core–periphery model is that managers’ use of labour utilization systems that combine the two forms of flexibility is deliberate, if not always strategic. The evidence for this assumption is also mixed, though authors often disagree on what constitutes a strategy (e.g. whether it is a pattern or an explicit plan — Procter et al. 1994: 230–8). In Britain, Hakim (1990: 168–73) found that only between 5 and 15 per cent of employers adopted a core–periphery strategy (indeed, only about a third of employers had a strategy for manpower policy at all) (see also Hunter et al. 1993: 396; and Marginson 1991). On the other hand, several US studies support the idea that managers utilize a core–periphery labour use strategy. Mangum et al. (1985: 603) reported that their unstructured interviews with employers and representatives of temporary help firms suggested the existence of a core–periphery relationship within the internal labour market.
Moreover, Harrison concluded (1994: 205), on the basis of a 1988 Conference Board survey of large American firms, that they were ‘deliberately creating both peripheral and internal low-wage labour markets for employees on contingent work schedules’. In addition, over half of the respondents in the recent survey of more than 1000 establishments in the United States mentioned earlier (see Kalleberg 2001: 7) ‘agreed’ or ‘strongly agreed’ with the statement: ‘Your human resource management strategy divides the workforce into permanent and non-permanent employees,’ though of course this begs the questions of whether managers have strategies at all and what they mean by them.

4. Beyond the core–periphery model: emerging conceptions of flexible firms

The core–periphery model is an ideal-type in the Weberian sense (Hakim 1990: 162). This framework oversimplifies the dynamics of the real world, though it provides a useful starting-point for deriving more complex perspectives. Several assumptions of the core–periphery model serve as bases for developing more realistic conceptions of the interrelations between functionally and numerically flexible labour utilization strategies.

First, conceptualizing the varieties of organizations’ labour utilization strategies in terms of a homogeneous core and periphery is too simplistic. Elaborating the elements of the ‘core’ is especially problematic (Pollert 1991: 11–12): Atkinson (1984: 29–31) discussed a variety of work arrangements that comprise the periphery, but tended to treat the core as a fairly homogeneous group. In addition, equating numerical flexibility with the periphery and functional flexibility with the core underplays the ability of organizations to obtain numerical flexibility from core workers (e.g. by asking or requiring them to work overtime) and functional flexibility from subcontractors and other forms of ‘externalized’ labour (see e.g. Ackroyd and Procter 1998: 170–5). Moreover, to view core and peripheral workers as occupying positions in separate parts of the organization is to neglect to consider ways in which these groups of workers may work together within the same departments and may even perform the same jobs within an organization.

Second, the relationship between the core and periphery sector(s) is more complex than is generally assumed by the core–periphery model. It may not always be the case, for example, that workers in the periphery are used to buffer the core, as this model suggests. Rather, the periphery may be related to the core in other ways, such as when an organization uses temporary help agencies to recruit, screen and try out workers for permanent positions.

Third, the core–periphery model is too ‘firm-centred’ and hence pays insufficient attention to dualism between firms. For example, while Atkinson (1987: 90–1) identified subcontracting (‘distancing’) as a component of an organization’s periphery, he viewed distancing as an alternative to flexibility — rather than as another form of flexibility — since he argued that it avoided
the need for a firm to organize its own work-force flexibly. However, as more recent studies have shown (see below), organizations may obtain both functional and numerical flexibility by developing relations with other organizations via networks.

The emerging perspectives on organizational flexibility outlined below seek to overcome the inadequacies of one or more of these assumptions by developing a more nuanced view of the mechanisms by which organizations’ labour utilization strategies create and maintain combinations of standard and non-standard work arrangements.

Two general theoretical perspectives are especially promising for describing how organizations simultaneously combine functional and numerical flexibility. One focuses on how organizations may integrate these two forms of flexibility internally, by means of different kinds of ‘human resource portfolios’ or mixes of labour utilization systems. The other emphasizes how organizations can achieve both numerical and functional flexibility via their external relations with other organizations.

Internal Organizational Labour Utilization Systems

Organizations may utilize various combinations of labour utilization systems, or types of human resource ‘portfolios’, that are chosen from a ‘menu’ of possible such arrangements. An informal theory of such portfolios was suggested by Elmer Winter, former president of Manpower, Inc., who noted that: ‘the efficient business firm which desires flexibility and decreased employee costs should utilize a work force mix that is 75 percent full time, 15 percent temporary, and 10 percent permanent part-time’ (cited in Gannon 1974: 44).

This line of reasoning is illustrated by several attempts to develop more systematic classifications of human resource portfolios. First, Way (1992: 333–6) distinguished three ideal-type, non-mutually exclusive staffing strategies to account for why organizations may use temporary employees. He argues that firms are likely to use temporaries to protect and retain regular core employees only under certain conditions: it is a strategy used to control costs of work-force size adjustment mainly by organizations facing potential layoffs arising from product market fluctuations (e.g. seasonality, cyclical, short product cycles, special projects). In order to implement the core–periphery labour utilization strategy, Way argues that it is necessary for this to be supplemented by a philosophy of no layoffs for core employees, high training investments and high technological change. A core–periphery strategy also requires an available supply of peripheral workers, since it can be applied on a widespread basis only if there are a sufficiently large number of labour force participants who are willing to accept peripheral jobs.³

Way identifies two additional models of labour utilization that combine functional and numerical flexibility. Organizations may use ‘optimal-mix’ strategies to obtain sufficient employees (either regular or temporary) to

maximize productivity with minimal costs, but with no objective to protect a core group of employees. This strategy is associated with pressures to minimize costs (e.g. from product markets or politicians) and the motivation is to reduce separation costs rather than to protect core employees. It applies best to jobs with high general skills, so that temporaries are just as productive as regular employees. A third, ‘constrained choice’, strategy uses temporaries to make sure production is not disrupted; this is useful when staffing is needed on a ‘24–7’ (24 hours a day, 7 days a week) basis. The recruitment objective of this strategy is to hire regular workers except when constrained by supply (when it is hard to hire a regular employee at existing compensation levels) and by demand (when it is imperative to hire someone).

Way derives these three scenarios from patterns that he observed in qualitative data he collected in the late 1980s from 50 case studies of organizations in a variety of industries in the Mid-Western region of the United States. As he recognizes, however, quantitative data (e.g. from surveys) are needed to test more systematically his descriptions of these three types of staffing strategies.

Second, Tsui et al. (1995) develop a framework that seeks to explain when organizations are likely to use functional as opposed to numerical flexibility strategies. They distinguish between employment relations characterized by functional flexibility (which they term ‘organization-focused’*) and relationships that provide numerical flexibility (which they call ‘job-focused’). They argue that most organizations are likely to use both job- and organization-focused employment relations simultaneously, though the proportions of each type that are used will vary. They posit three scenarios for combining organizational- and job-focused strategies. ‘Organization-focused’ companies (e.g. the ‘old’ IBM prior to the 1990s, which had a ‘no layoff’ policy) are those for which most employment relations are organization-based, though job-focused relations are used for activities that are peripheral to the core mission of the organization (e.g. security, maintenance). A ‘job-focused’ company uses mainly job-focused relations and has organization-focused relations with only a small core group of employees (e.g. an accounting consulting firm with a large group of junior accountants and a small group of tenured partners). A third model uses both types of employment relations extensively with both core and periphery employees. Job-focused relations are used for jobs that are both central and peripheral to its main activities, e.g. a company may hire an executive on a contingent basis, or contract out its entire training function. However, they suggest that this third scenario, a blend of organization- and job-focused employment relations, is likely to produce conflicting messages for participants, and thus is likely to be particularly unstable.

Tsui et al.’s propositions suggest factors (both external and internal to the organization) that predict whether an organization will use one rather than another of these three scenarios. For example, they hypothesize that organizations are more likely to use organization-focused relations with workers who are in short supply and for more complex jobs. This framework is useful in that it identifies factors that are both external and internal to the organization. But it also needs assessing through empirical research.
Third, Sherer (1996) and Sherer and Lee (1992) extend the core–periphery model by seeking to explain which combinations of human resource portfolios might be used in various circumstances. Sherer (1996) identified a variety of labour utilization strategies that differ in the degree to which workers are internal to the organization: partnership relationships, in which labour is strongly attached to the firm via ownership rights and voice in decision-making; employment relationships, in which labour is inside the firm to some degree; and contracting relationships, in which labour is outside the organization. Each of these types of work arrangements involves distinct forms of control and financial arrangements between employers and employees. Firms may be able to obtain both functional and numerical flexibility by combining these three types of work arrangements in various ways, for example by using employment relationships for some jobs, partnership arrangements for others and contracting-in relations for still others. Sherer’s approach is less restrictive than Way’s in that he considers relationships in addition to those between temporary employees and an undifferentiated group of regular employees.

Sherer and Lee (1992) use survey data on work practices of 853 firms to examine empirically this classification scheme. Their measures included indicators of: partnership arrangements, such as whether compensation was contingent, and practices that provide workers with opportunities to participate in decision-making, such as self-directed teams; modified employment relationships, such as part-time work, flexible hours, home-work and work done at a satellite facility; and the use of independent contractors. They utilize clustering techniques to show that organizations had different mixes of the three types of work arrangements and then tested several hypotheses regarding the correlates of the various combinations. In particular, they found that larger firms were more likely than smaller firms to utilize a mix of the three work arrangements. They also found evidence that ‘periphery’ work arrangements (such as contracting relations) are not always needed to buffer ‘core’ arrangements (such as partnership relationships and modified employment relations).

Sherer and Lee’s cluster analysis identifies the degree to which work arrangements are ‘core’ or ‘inside’ the organization and indicates the range of mixes of labour utilization practices that may be found in empirical situations. However, their analysis did not test systematically hypotheses regarding the conditions under which organizations are likely to utilize one or the other of these combinations of labour utilization strategies; nor did they attempt to assess the consequences of organizations’ use of the various combinations.

Finally, Smith’s (1994) case study of a service company provides an empirical illustration of how organizations are able to integrate regular and non-standard workers. Her study shows that functional and numerical flexibility can exist in the same organization along a continuum, in which permanent workers in functionally flexible jobs often work side by side with temporary workers, rather than being isolated from each other in separate,
dual sectors, as assumed by the core–periphery model. Moreover, she demonstrates that contingent, numerically flexible workers are heterogeneous, with some temporary workers lacking any job security at all and others lacking security at a particular workplace of employment but having relatively secure employment with the firm. She also shows that a company may be able to avoid the conflicts that arise when using temporaries by permitting permanent employees a relatively high degree of autonomy.

Organizational Networks

An alternative, but not mutually exclusive, elaboration of the core–periphery model sees organizations as combining functional and numerical flexibility by forming external relations with other organizations rather than by adopting internally different types of labour utilization strategies. The organizational forms created by these interorganizational relations have been referred to alternatively as quasi-firms (see Eccles 1981: 340 on construction), hybrid organizations (Williamson 1991: 280–1), flexible specialization production systems (see Piore and Sabel 1984: 213–16 on textiles), or network organizations (Powell 1990). Organizational networks represent an alternative to bureaucracy as a way of organizing work. Network organizations are forms of co-operative endeavour characterized by ‘reciprocal patterns of communication and exchange’ (Powell 1990: 295) and are governed by trust relations rather than the logics of price or hierarchical organization. Network organizations are illustrated by strategic alliances, joint ventures and partnerships, and subcontracting. Like the use of internal human resource portfolios, the creation of networks among organizations is in part an attempt by employers to reduce their resource dependencies and uncertainty (associated with skill shortages, for example) and to respond to changes in competition, technology and turbulent environmental conditions (Pfeffer and Baron 1988: 277).

Through networks, organizations can achieve functional flexibility by collaborative relations among specialized suppliers and producers, as illustrated in the Emilia–Romagna industrial district of Italy (Piore and Sabel 1984: 226–9, 266), and by the network of computer-oriented firms centred around Route 28 in Massachusetts and Silicon Valley in California. The close, trust-based relations among organizations in these networks enable producers to reconfigure their relationships with suppliers and to develop interorganizational competencies, thus achieving innovation and multi-use of equipment (and further enhancing trust relations) by being able to redeploy their resources as needed so as to adapt to changes in the environment (Harrison and Kelley 1993: 228–9). Moreover, organizations may obtain numerical flexibility by outsourcing functions such as production, maintenance, repair, clerical and other ‘non-core’ activities, and by hiring workers from temporary help agencies.

Analogous to the core–periphery model of internal organization structure, external networks among organizations may foster a duality between...
primary and secondary firms that differ in job quality (Hunter et al. 1993). Indeed, these differences between firms may be more important than high-wage/low-wage divisions within a firm. Harrison (1994: 196) argues that the flexibility of large firms depends fundamentally on the perpetuation of contingent work (part-time, part-year, temporary and contract work):

Today, business is operating in a world in which best practice increasingly entails vertical disintegration, downsizing, outsourcing, and the formation of networks of companies in order to operate across national borders and sectoral boundaries . . . when we characterize the prototypical business organization of the new era as a lean and mean flexible firm, embedded within networks made up of partners and dependent suppliers and subcontractors, we are implicitly recognizing that the workplace in the new economy is, arguably by managerial intent, being systematically divided into insiders and outsiders.

The insiders are full-time, relatively secure ‘core’ workers, with fringe benefits, training and promotional opportunities. The outsiders are often contingent workers. Harrison also argues that dual labour markets can exist within single firms, especially when the firm is a multinational corporation that is dispersed geographically, such as with Nike. Employees who occupy an organization’s periphery may be employees of other organizations that are connected to the focal organization by means of networks. The inequalities that are generated between insiders and outsiders constitute a ‘dark side’ of flexibility.

The dualism between workers (which may exist between as well as within firms) created by organizing work via organizational networks is deepening as managers often pursue a strategy, if not always intentionally, of identifying some workers they want to keep (‘first class’ or ‘core’) and others (‘second class’ or ‘periphery’) who are expendable (Heckscher 2000: 280–1). Core workers have more alternatives to be employed in other organizations, in part because their skills are portable, and so they may not choose to remain permanently with their employer. Workers in the periphery, on the other hand, have fewer alternatives to obtain comparable jobs in other organizations than core workers, and so these peripheral workers are often relegated to the status of second-class citizens with their employer. In order to improve mobility prospects of peripheral and core workers, it is necessary to establish societal institutions (e.g. education and labour market intermediaries that span firms) that are able to provide the needed inter-firm training and mobility. These institutions are relatively undeveloped in the United States compared with European countries such as Germany (see also Osterman 1999).

5. Cross-national differences in flexible firms

My discussion of the two forms of organizational flexibility has focused mainly on the liberal market economies of the United States, the United

Kingdom and Canada. This was primarily for reasons of space: it would require another paper of equal length to review literature on these topics in non-Anglo Saxon countries in Europe and the Pacific. Much of the research on organizational flexibility in the USA, the UK and Canada has tended to be firm-centred, emphasizing employers’ choices regarding how they organize work. This reflects the highly deregulated labour markets and employment relations that characterize these countries, in which state regulations and institutions play relatively small roles in employers’ decisions to adopt one or the other form of flexibility. Illustrating this, Smith et al. (1995: 705) found that there was much less uniformity across industries in the use of external (numerical) flexibility in Canada than in Sweden.

Extending the discussion of organizational flexibility to other countries underscores the importance of considering explicitly the role of the state (e.g. laws and regulations governing trade union influence, employment protection and the operation of temporary help agencies) and that of economic, social and political institutions in shaping employers’ labour utilization strategies. Organizations in all industrial countries need to be flexible in order to respond to competition, technological changes and changes in labour force composition, though the types of flexible labour utilization strategies that organizations are likely to adopt depend on their country’s institutional context.

Whether organizations are likely to utilize numerical flexibility strategies depends on their country’s regulatory regime, such as the amount of protection given to regular, permanent workers and the existence of laws that limit the use of temporary help agencies to certain kinds of work. It has been argued that North American firms, in which employment protections and restrictions on temporary help agencies are relatively low, rely more than European and Japanese firms on external flexibility strategies (Clarke 1992: 239–40). Moreover, Smith et al. (1995: 705–7, 712) find that managers in the telecommunications industry were more likely to adopt numerical flexibility strategies as a response to economic pressures in countries (e.g. Canada) where national institutions involve few restrictions on managerial decisions regarding the level of employment. In Spain, the liberalization of regulations regarding fixed-term contracts in the mid-1980s has been cited as a major reason for the rapid expansion in such contracts (Toharia and Malo 2000: 312–13). In Germany the strength of unions, dominance of occupational markets and strong systems of vocational training combined to provide less scope for the development of peripheral labour forces within firms in the 1980s (Lane 1989: 286–8), and so the emphasis was on ‘new production concepts’ that facilitated functional flexibility (e.g. Kern and Schumann 1987: 159–67).

The likelihood that organizations will adopt functionally flexible labour utilization strategies depends on the existence of institutions that help employers spread the risk of long-term training, development and innovation in work design, as well as on a high level of trust between managers and workers. Germany and Japan had high trust relations in the 1980s and
1990s, and so were better able to achieve functional flexibility than Britain
and France, which had relatively poor trust relations in this same period
(Lorenz 1992: 457).

In countries in East and South-east Asia, firms were able to adopt
primarily numerical flexibility strategies (or ‘static’ flexibility strategies) to
reduce short-term costs due to the lack of significant labour opposition
(Deyo 1997: 108–11). These strategies led firms to subcontract to low-cost
suppliers, which in turn led to the growth of casualization and contract
labour in the large export sectors of the Philippines, Thailand, Malaysia,
Hong Kong and Singapore. These static labour utilization strategies (in
addition to political factors) helped employers to further weaken unions and
contributed to the marginalization of organized labour in East and South-
east Asian countries. The adoption of static flexibility strategies was also
facilitated by the absence of governance institutions that encouraged long-
term investments in technology, human resource and organizational reforms
and research and development. These institutions were absent particularly in
Hong Kong, Thailand and Malaysia, and among small firms in most East
and South-east Asian countries.

In Taiwan, family and associational bonds were able to underwrite
co-operation and investments, creating the trust-based, enduring relations
within and among firms needed for functional flexibility (or ‘dynamic’
flexibility strategies). The Taiwanese model provides an alternative to the
Italian and Japanese models of interfirm co-operation for providing collec-
tive goods and long-term investments, as well as for minimizing risk and
facilitating innovations (Deyo 1997: 100). When East and South-east Asian
countries adopted dynamic flexibility strategies (primarily in product
market niches that involve high levels of quality, batch production and
process technologies), these strategies tended to be more autocratic than
those in more industrially advanced countries (Deyo 1997: 111–12).

6. Towards a research agenda on flexible firms

Organizational flexibility remains an important focus for research on work
and organizations in the twenty-first century. Organizations will persist in
using labour utilization strategies that are functionally and/or numerically
flexible in order to respond to competitive, technological and labour supply
pressures. Employers are likely to be motivated to internalize certain func-
tions and to encourage functional flexibility when they believe that they
can do so efficiently and cost-effectively. At the same time, the greater
institutionalization of organizational networks, temporary help agencies
and other labour intermediaries provides employers with more options to
externalize labour and otherwise adopt numerical as well as flexibility
strategies.

Analyses of the interactions between the two forms of flexibility are
important because of their relevance to research issues in fields of industrial

relations, organizational behaviour and social stratification. For example, organizations’ attempts to protect a core of functionally flexible workers at the expense of those whose relations to the firm are tenuous may help to explain the growth in inequality that has been observed in the United States and other industrial countries during the 1990s. Interactions between permanent and temporary workers may also generate conflict between these groups, hampering organizational performance and putting pressure on the ability of workers to organize in unions. A better grasp of the relations between these two types of flexibility may also illuminate the strategies used by managers — such as the differences between commitment-maximizing and cost-reduction responses to product market competition — that may have important implications for organizational performance, labour–management relations and job quality. Moreover, examining the linkages between functional and numerical flexibility may help us to appreciate better the role of the state, unions and other institutions in setting constraints and parameters on work organization, labour markets and systems of industrial relations.

The idea of the ‘flexible firm’ as a framework for research on the interplay between these two forms of organizational flexibility is likely to remain important for social scientists studying work, organizations and industrial relations. The theoretical perspectives on mixed labour utilization strategies discussed here offer two alternative perspectives to the core–periphery model as to how organizations may achieve the seemingly contradictory goals of obtaining functional flexibility from a committed group of highly skilled, trained and committed workers, while at the same time having the numerical flexibility to adapt to changes in demand. The internal and external perspectives are not mutually exclusive: for example, organizations seeking to adopt a mix of human resource portfolios by using temporaries or subcontractors will typically obtain these workers from organizations with whom they establish network relationships.

Much remains to be done in order to refine these perspectives on the flexible firm into a systematic theory. We need to specify and test theories of how organizations are able to combine the advantages of functional and numerical flexibility, while minimizing the conflicts and other disadvantages created by the coexistence of these two types of organizational flexibility. Research has made little progress in showing how the use of peripheral workers undermines trade union organization (Atkinson 1987: 98–102), or how employers’ attempts to obtain functional and numerical flexibility from different groups of workers deepen labour market segmentation (Atkinson 1987: 91–2).

Specifying further internal and external theories of mixed labour strategies has been hampered by the lack of requisite data. Within a country, studies of organizations’ labour utilization strategies are generally based either on aggregate analyses of labour market data (such as the proportion of temporary workers in the labour force) or on surveys of employers. Such research designs usually provide fairly general information that is often not
very useful for examining within-organizational structures and may obscure the mechanisms underlying these strategies. At the other extreme, case studies are often rich in detail but limited in generalizability, since cases are often chosen because they were known to have been innovative and to have introduced changes designed to increase flexibility.

The data needed to examine the nature and consequences of labour utilization strategies should go beyond surveys that simply ask senior managers about their strategy or use of various practices, as Procter et al. (1994: 236–8) argue. Information needs to be obtained from multiple informants about how labour utilization strategies are actually implemented within the organization (assuming they have such strategies at all) and how work is organized and staffed. These data should be collected at multiple levels within the organization, so as to permit an assessment of the actual implementation of flexible staffing strategies and non-standard work arrangements and their impacts on employees’ experiences and the conflicts that may arise between regular and non-standard workers.

Studying issues of strategy and employer motives would of course be facilitated by the collection of historical and longitudinal data on the combinations of labour utilization strategies used by employers. Most of the longitudinal studies (e.g. Casey 1991: 194–8) have focused on trends in externalization of employment. Such data are somewhat suggestive of changes in flexible firms, since externalization represents a departure from the standard employment relationship and since most organizations are likely to internalize at least some of their members. Nevertheless, data on trends in externalization do not speak directly to the debate about whether or not there has been an increase in segmentation within firms. Nor do such data shed much light on how network linkages between firms develop and change over time.

Central to many of these research objectives is the identification of the appropriate unit of analysis for examining arguments related to organizational flexibility. A limitation of most previous studies is that they measure these two types of flexibility at the establishment level. The establishment may not be as useful as the firm for studying organizational flexibility, since firms may often obtain flexibility by dividing work differently among various establishments, as is often the case with profit as opposed to cost centres. However, to the extent that organizations obtain functional and numerical flexibility by means of their relations to other organizations in networks — as is the case in subcontracting relations or linkages to temporary help firms — the most appropriate unit of analysis may be neither the firm nor the establishment, but the network defined by the relationships among the organizations and the labour market intermediaries from which they recruit their workers and subcontract some of their functions. Moreover, as the way in which work is organized shifts from bureaucratic structures to networks, people’s careers may be more likely to span individual organizations in the future, making sequences of projects or jobs more relevant units of analysis for studying flexibility.
Further research is needed especially on the characteristics of the networks by which organizations obtain functional and numerical flexibility. There have been relatively few studies of network dimensions such as the durability, intensity, formalization and overlap of the relations between client organizations and the temporary help agencies and subcontractors from which they obtain workers and services. Long-term contracts with temporary help agencies, for example, may enable organizations to obtain skilled workers without making investments in training these employees internally. Moreover, studies of the interrelations among labour market intermediaries (such as temporary help agencies) would illuminate the competitive dynamics of this growing and increasingly important sector of the economy.

Examining the role of supra-organizational institutions in shaping employers’ labour utilization practices is particularly important in countries where labour markets are highly regulated (unlike the United States and United Kingdom), since employers’ choices in these countries are more constrained by economic, social and political structures. Cross-national studies of institutional variations in organizational flexibility offer considerable potential for revealing the range of variation in the organization of work and employers’ labour utilization practices. Comparative analyses of the interplay between functional and numerical flexibility are likely to enhance considerably our understanding of organizational performance, labour–management conflict and workers’ career experiences, attitudes and behaviours.

Final version accepted 20 July 2001.

Acknowledgement

This article was written while I was a visiting scholar at the Russell Sage Foundation in New York City, whose support I gratefully acknowledge. I also thank Stephen Wood for his perceptive and persistent comments on earlier versions.

Notes

1. A third type of flexibility discussed by Atkinson and others is pay or wage flexibility, whereby managers lower workers’ wages as business conditions warrant. This form of flexibility has not been as extensively studied as the others, owing in part to its having proved to be a less useful flexibility strategy within organizations than the other two forms. Wages tend to be ‘sticky’ downward and resistant to managers’ efforts to lower them.
2. Organizations also achieve numerical flexibility from their regular, permanent work-force by requiring them to work overtime (Streeck 1987: 294).
3. If the peripheral labour force is not sufficiently skilled or committed to perform the same tasks as the core group, then the firm may not be able to create buffers in the key positions. Periphery workers are not committed to organizations or socialized in them, and efforts to socialize them or to enhance their commitment may destabilize this system, as the peripheral labour force may seek employment relations comparable to the core group (Osterman 1988: 85).

References


