David Marsden
Norms of exchange and the socio-economics of labour markets

Book section

Original citation:

© 2016 The Author

This version available at: http://eprints.lse.ac.uk/65808/

Available in LSE Research Online: September 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (http://eprints.lse.ac.uk) of the LSE Research Online website.

This document is the author’s submitted version of the book section. There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.
Norms of exchange and the Socio-Economics of labour markets

David Marsden
d.marsden@lse.ac.uk

17/03/16

1. Introduction ........................................................................................................... 2
2. Enablers, constraints and self-interest ................................................................ 4
   a) Reciprocity ...................................................................................................... 5
   b) Risky transitions and biased risk perceptions .................................................... 6
3. Institutional foundations: employment, sales and hybrid contracts ...................... 7
4. Sustaining reciprocity in the standard employment relationship ............................ 11
   a) The zone of acceptance as a bargain .............................................................. 11
   b) Sustaining reciprocity in the zone of acceptance ............................................. 13
   c) Voice and a common language ....................................................................... 14
   d) Biased risk perceptions, localised self-help, and restrictive rules .................... 16
5. Transient employment relationships supported by social networks and social norms .... 16
   a) High skilled transient employment .................................................................. 18
   b) ‘Precarious’ contracts for low skilled and routine work .................................. 19
6. The production of skills and skill categories ............................................................. 19
   a) Conflicts over cost and quality of training for transferable and specific skills .... 20
   b) Employer collective action to prevent free-riding ............................................. 21
   c) Employment protection and worker investments in skills ............................... 23
   d) Risk sharing among employers to encourage labour flexibility ....................... 24
7. Unions and employers as composite actors ............................................................ 26
8. Conclusion: labour markets as social institutions ................................................... 30
9. References ............................................................................................................ 32
10. Endnotes .............................................................................................................. 39
1. Introduction

Socio-Economics stands at the intersection between Economics, Sociology, and other social sciences. Its primary tenet is that markets are social institutions created by human societies for the purposes of exchange and productive enterprise. This chapter aims to give theoretical and empirical substance to this proposition in the field of employment, and this encompasses both the allocative function of labour markets, and the productive function of employment relationships. There is a perpetual tension between these two as the first tends towards treating labour services as commodities, whereas the team relationships of the latter militate against this. As Alfred Marshall maintained, the demand for labour is a derived demand, derived from its value in its different productive uses. As a result, relations of exchange and production spill over into each other, and this is reflected in the institutions and norms governing workers’ economic activity. Yet the impersonal relations of market exchange are often contrasted with the social relations of team production. Adam Smith’s two great works, his Wealth of Nations (1776), and his Theory of Moral Sentiments (1759), reflect the tension between our thinking on markets and on social ties, hence their treatment in two separate works, and the so-called ‘Adam Smith problem’. Nevertheless, he was acutely aware of the empirical interconnections between them. He was deeply pessimistic about the motivation of salaried managers in charge of joint stock companies, believing their ‘negligence and profusion’ would bring ruin. He was more confident about the future of family firms whose leaders were often motivated beyond their personal gain by providing a legacy for their offspring, that is those with whom the bonds of fellow-feeling, moral sentiments, were strongest.

The Economics Nobel Laureate Robert Solow captured the sentiment of many labour economists in the title of his 1990 lectures which described the labour market as a ‘social institution’. The puzzle that concerned him was why wages do not appear to behave like other prices, and fall in a recession. He was especially interested in why unemployed workers appear to follow a social norm that overrides their individual short-term interest, and so refrain from undercutting the wages of those in employment. In similar vein, Truman Bewley (1999) reported the equally interesting finding that employers do not generally cut wages in a recession because of their concern about employee ‘morale’. From a different perspective, many economic sociologists observe that such norms play a key role in the workings of markets. Arthur Stinchcombe (1986) argues that markets use norms of exchange ‘to see to it that [people] get want they want’ out of transactions. Mark Granovetter (1974, 1985) argues that markets are ‘embedded’ in social relations, such as those supporting job search networks, and Jens Beckert (2009) argues that the social and institutional foundations of markets facilitate the pursuit of economic activity. Behind these two approaches, one may distinguish a strong and a weak sense of ‘institutions’. Each has developed separately, and this chapter will argue for a synthesis.

Douglass North (1990) opens his study of institutions in economic life describing them as ‘the rules of the game in a society, or more formally, they are the humanly devised constraints that shape human interaction’. Searle (2005) offers a similar, if more detailed, definition: that institutions have certain special characteristics that distinguish them from other social phenomena, and set their norms apart from regulatory norms, such as the Highway Code, and from organisational routines and practical ‘rules of thumb’. They are defined by a set of constitutive rules; these rules determine status functions that are
collectively recognised and accepted; these functions are performable only in virtue of this recognition and acceptance; and they should carry recognised and accepted deontic powers, such as rights, obligations and permissions. He argues that statements about institutions in relation to action are of the kind: ‘X counts as Y in context C’. Thus instructing a worker to carry out a task or to apply an organisational routine, counts as exercising managerial authority in the context of an employment relationship. The worker’s obligation to obey that instruction is one of the rules, together with others, which constitutes an employment relationship. Independent contracting, self-employment, is constituted by a different set of rules. In employment law, the presence of an authority relationship is often used as one of the tests applied to determine whether a worker is engaged as an employee or self-employed. This will then be used to determine each party’s rights and obligations.

In contrast, institutions in Solow’s account and those of many labour economists appear to belong to the weak sense, in which norms are self-enforcing. This is illustrated by his analysis of the norm just mentioned that the unemployed should not undercut the wages of employed workers. He argues this is maintained by the willingness of the employed to punish undercutting even at considerable cost to themselves. For simplicity, he assumes workers seek employment at an equilibrium wage, which takes account of factors such as their past investments in skills. However, workers who lose their jobs receive only unemployment benefits, the dole, which determines their reservation wage. They have to choose between two alternatives: either accept unemployment and the dole now, but with the prospect of future employment at the equilibrium wage; or undercut the equilibrium wage in order to gain employment now. However, incumbent workers will retaliate by themselves undercutting, which drives everyone’s pay down to the reservation wage. Thus the two payoff streams facing the unemployed worker are the reservation wage now followed by probable employment at the equilibrium wage, and employment at the reservation wage now and into the future. Provided the first is greater, the norm will hold. Thus, whereas institutions according to the weak sense rely upon the willingness of actors to punish those who transgress the norm, according to the strong sense, there is an additional deontic, obligational, element that commits the actors, such that, once the context is defined, certain behaviours are considered legitimate and others, inappropriate. For both North and Searle, this deontic element enables the actors to dispense with complex calculations as to the other party’s self-interest when choosing their own strategy, and so greatly simplifies economic coordination.

In both cases, rules emerge that constrain short-term individual self-interest, giving rise to stable patterns of behaviour, and these can improve predictability in an uncertain decision environment, which is a considerable benefit in view of the limited cognitive abilities of human beings (Simon, 1976). The weak sense has proved valuable in modelling the emergence of behavioural norms out of evolutionary games (Bowles, 2005), but as Gibbons and Henderson (2012), argue, it depends upon the ability of actors to identify breach when it occurs, and that is often ambiguous. In contrast, the strong sense has more to say about systems of institutional rules, and in particular, addresses the problem of how to determine whether ‘X counts as breach in context C’. This chapter will argue that a socio-economic approach seeks to draw on a combination of the strengths of both senses.

In social science usage, the terms ‘institution’ and ‘organisation’ bear many different meanings, so some clarification is needed. In the present context, it has become common to distinguish between them respectively as ‘games’ and ‘players’ (North, 1990, Scharpf, 1997). Within the game, its rules determine the legitimate roles and moves, and the players apply these with the aim of winning. Organisations are groups of individuals bound together to
achieve a particular purpose, and the institutions set the legitimate parameters for their strategies. A complication is that organisations, as Scharpf (1997) argues, are themselves ‘composite actors’ with their own internal governance rules which assign their leadership specific powers and responsibilities. Following the team metaphor, ‘Manchester United’ is an organisation which has two sets of complimentary objectives: to win football competitions and to make a profit for its shareholders. It assumes its institutional mantle as a football club by signing up to the governance rules of the Football Association (FA), and similarly, for its business activities, it registers in legal form as a Limited Company. The organisation assumes the institutional forms of both a ‘football club’ and a ‘limited company’. In Searle’s terms, we might say: Manchester United (the organisation) counts as a football club (an institutional form) in the context of FA activities. The FA is itself an organisation with the purpose of regulating the sport of football. Historically, it played a key role in establishing a common set of rules of for the sport across England, thus facilitating the growth of football tournaments.

Simply to argue that labour markets are social institutions is not very helpful. Instead, the chapter seeks to show how differently constituted institutions shape the strategies of workers and firms and their collective organisations. In common with the economic approach, it builds up from the micro-level, considering how the two fundamental types of employment relationships, whether workers engage as employees or as self-employed independent contractors, and the growing number of the hybrid forms, respond to the needs of workers and firms, and also how they shape their respective strategies at both the individual and collective levels. It is grounded in the economic theory of competitive markets, as this provides the most developed theory we have of markets as self-regulating entities and with minimal demands on institutional support, and provides a basis for understanding evolutionary norms, but it will be argued that this is not sufficient. It seeks to bring out the changing balance between constraining and enabling rules and thus to shed light on current labour market developments. It seeks an approach that spans both ‘external’ and ‘internal’ labour markets in the sense that employment relationships encompass what North (1990) describes as the ‘transaction’ and the ‘transformation’ components of employment, that is the allocative function of labour markets, the agreement of terms of exchange, and the ensuing productive collaboration, the latter being the purpose of the first two. These range from spot contracts for simple and easily defined services to long-term collaborations with complex governance structures, and as Williamson (1985) observes, the nature of the institutional support varies accordingly.

This approach to institutions needs to avoid two potential dangers: the ‘panglossian’, and the ‘functionalist’. The first treats each institutional intervention as optimal, which makes it difficult to understand change (Streeck, 2009). The second reduces social norms and institutions to their market functions whereas in practice their functions are multiple, and shaped by many competing group interests as Thelen argues (2004).

2. Enablers, constraints and self-interest

In recent years, two very significant advances have been made by means of experimental methods applied to common economic transactions, and both emphasise the enabling as opposed to the constraining role of institutions. They extend and reinforce Etzioni’s (1988) argument for recognition of the ‘moral dimension’ of economic activity, which he based on an older sociological literature. They concern reciprocity and risk. The first gives greater weight to people’s readiness to engage with what the sociologist Alvin Gouldner (1960) described as the ‘norm of reciprocity’, and owes much to the work of Ernst Fehr and
colleagues at the University of Zurich. Following this line of argument, suitable institutions are no less important for encouraging cooperation by fostering the expectation of reciprocity than for restraining opportunism. The second emphasises systematic biases and asymmetries in individual workers’ perceptions of labour market risks and opportunities, pioneered by the work of Kahneman et al. (2000). Building on their work, Schmid (2006) argues that by inducing workers to overestimate the risks, and underestimate the opportunities of employment change, they give rise to individual strategies that can aggravate rather than correct labour market imbalances. In this case, suitable institutions can mitigate these biases by framing risks and opportunities more appropriately, and by providing a safety net of mutual insurance.

\textbf{a) Reciprocity}

The work on positive and negative reciprocity demonstrates that people vary considerably according to their expectations that others will reciprocate beneficial actions in a series of transactions. They differ also in their willingness to sacrifice their immediate self-interest and punish a lack of reciprocity in order to maintain that norm (Fehr and Gintis, 2007, Fehr and Fischbacher, 2003). To laypeople this may seem obvious. On the other hand, attempts to prove that reciprocity, as in ‘tit for tat’ cooperation, can build on purely self-interested behaviour depend on very strong assumptions which make the emergence of cooperative behaviour seem unrealistic: they assume the long-term interest of both players, and that the game continues indefinitely (Boyer and Orléan, 1995). Yet Axelrod’s (1984) most striking empirical demonstration of the power of ‘tit-for-tat’ was the emergence of cooperative behaviour between soldiers in the opposing trenches of the First World War. Known as ‘live and let live’, the soldiers deliberately fired to avoid harming their opposite numbers, playing tit-for-tat under very adverse conditions in which overt cooperation was punishable as treachery, and in which the generals actively sought to disrupt such behaviour (Ashworth, 1980). In an industrial setting, one such form of reciprocity, between line managers and their workgroups against senior management, gave rise to Gouldner’s (1954; ch2) account of the ‘indulgency pattern’ in his famous study of workplace relations in a gypsum plant: local management tolerated certain work practices that violated its rules in return for greater cooperation. We witness cooperation all around us despite agents having to manage conflicting interests. The application of experimental methods to simple economic transactions enables us to distinguish more precisely acts that correspond to individual self-interest from those which are both other-directed and incur a significant individual cost. This makes it possible to explore how different frameworks for transactions may encourage either type of behaviour (Bowles, 1998).

The readiness of many agents both to expect, and to engage in positive reciprocity, is well-documented (Fehr and Gächter, 2000, Fehr and Gintis, 2007), and is well-illustrated for the present context by the experiments in the tradition of the trust and gift exchange games. In a series of experiments focused on task assignments in employee and self-employed relationships, Bartling et al. (2013) showed that a high proportion of subjects would anticipate positive reciprocity and agree to task flexibility, even in one-off transactions, and an even higher proportion did so for transactions repeated with the same partner. Nor was this a naive response. Agents’ behaviour adapted as the experiments progressed. In the one-off transactions, cooperation declined with the accumulation of negative experiences, and in the repeated ones, positive reciprocity proved to be mutually reinforcing. Subjects could also switch between contract types, for example, if the ‘employer’ had abused its control over task assignments in the previous round.
It is widely acknowledged by both economists and sociologists that the effectiveness of social norms depends also on the willingness of agents to punish those who transgress them, a form of negative reciprocity (see Gibbons and Henderson, 2012). Such negative reciprocity has already been encountered in Solow’s example of willingness to punish undercutting. Fehr et al. argue that many social norms are like public goods in that the many benefit from their continued effectiveness, whereas the injured party often has to bear a disproportionate share of cost of punishing the transgressor. For example, a worker who feels her employer has imposed excessive job demands without her agreement may punish it by quitting. Such actions reinforce the norm that job demands should be increased only by mutual agreement, but the individual worker has to bear the cost of potential unemployment. Some subjects in Fehr’s experiments act as ‘strong reciprocators’ and are willing to bear such costs in the interest of sustaining the norm, whereas others are ‘self-interested’ and calculate that the costs of quitting outweigh the benefits for themselves. However, Fehr and Fischbacher (2000) show that the willingness of strong reciprocators to act depends on their confidence in its efficacy, and that without this the process can unravel.

Being willing to punish defection at one’s own expense is not sufficient to demonstrate the influence of a social norm. Two additional results are needed, both related to the often ambiguous nature of breach highlighted by Gibbons and Henderson (2012). The cooperative strategy is easily undermined: the longer breach remains unpunished, the more profitable it becomes; and mistakenly treating cooperative action as breach can also cause cooperation to unravel. Thus, first, breach needs to be clearly identifiable if the ‘punishment’ strategy is to work promptly, and not be misinterpreted as opportunistism. This implies reference to criteria that both parties can use. Secondly, such criteria should be understood and thought legitimate by third parties. Their importance is highlighted by the ‘public good’ experiments in which ‘third parties’ are willing to punish unfair actions in transactions engaging others even when they are not directly affected and at cost to themselves. This theme will return in the discussion of voice channels, but for the present, both arguments signal the importance of criteria for breach that extend beyond the two transacting parties and accessible to third parties.

In conclusion, it is not necessary to conceive of cooperation based exclusively on the immediate self-interest of the contracting parties because many human beings even in the economic sphere are willing play the cooperative strategy. The success of this strategy is supported by criteria, or norms, that extend beyond the actors involved in the current series of transactions, which reduce ambiguity, and can be understood by third parties. Field research, outside the behavioural lab, testifies to the prevalence of norms relating to fair behaviour in wage and employment transactions and which are likely to serve as signals as to how cooperative the other party may be (Behrend 1964, Kahneman et al. 1986, Bewley 1999). In a more modern form, these are consistent with older accounts by Keynes (1936) and Hicks (1955, 1974) that attributed nominal wage rigidity during recessions to workers’ fear of loss of relative position in a hierarchy of wage relativities between different groups which had acquired the sanction of custom.

b) Risky transitions and biased risk perceptions

In contrast to the optimistic view on reciprocity and cooperation, biased views of risk appear to point in a more pessimistic direction – the expectation of increased rigidity and
resistance to change, combined with perverse risk-taking, which could impede the adaptive functions of labour markets. In recent decades this has become more pressing owing to the increased diversity of labour market transitions of individual workers, each of which has its associated risks. With the decline of large firm internal labour markets, growing diversity of production and service delivery methods, the spread of project-based work, and consequent changes in contractual forms, there is a need to consider a broader array of transitions than the traditional ones between education, employment, unemployment and inactivity (Schmid, 2006).

Drawing on the prospect theory of Kahneman and Tversky (2000), Schmid (2006) argues that workers do not evaluate their current and prospective well-being in absolute terms, but in relation to certain reference points. Often they focus on the status quo or the recent past, which makes them particularly sensitive to potential losses: making them ‘loss averse’. At the same time, they tend to underestimate the value of future prospects. In combination, these make them more resistant to change and less willing to learn new skills than if they were to focus on the likely benefits of adaptation. When considering potential opportunities they tend to be loss averse, giving disproportionate weight to what they might lose, with a ratio of about 2:1. Equally, when faced with losses, they can be tempted to ‘bet their shirt’, and risk excessive losses in the hope of turning things around. This can lead them to over-invest for example in seeking entry into prestigious occupations, and becoming trapped in low-status entry positions, hoping for a ‘break’, long after their experience tells them that success is unlikely (Marsden, 2011). Schmid also draws out the implications of how people ‘frame’ labour market choices can affect their actions, for example, whether choices are framed in terms of potential losses or gains and how these interact with biases in their perceptions of risk.

The cumulative effect of these individual-level biases undermines the allocative function of labour markets, and can also undermine the patterns of reciprocity discussed earlier if, for example, they cause workers to distrust beneficial actions that call for positive reciprocity. Bowles (1998) for example argues that market exchange emphasizes self-interested actions, whereas production requires a high degree of cooperation. Schmid argues that institutions that help to ‘de-commodify’ labour can ease this tension by sharing and mitigating risks, and by the manner in which choices are framed. Collective action may help correct these. There is plenty of evidence of workers’ self-help to develop coping strategies to mitigate the risks inherent in such transitions. For example, one of the earliest goals of trade unions was to set up hiring halls and establish a form of unemployment insurance (eg. Webb and Webb, 1920).

3. Institutional foundations: employment, sales and hybrid contracts

In free labour markets, the supply of labour services is based on an initial agreement or contract. Two broad types of contract predominate: employees and self-employed independent contractors. In the OECD economies, together with trainee positions, which are normally geared towards work with one or other of these two contract types, they account for practically all kinds of paid employment activity. Although there has been a growth of so-called ‘non-standard’ forms of employment in recent years, as will be seen, these are best viewed as hybrids, in between employment and self-employment, and are born of changing economic circumstances.

It has been common to contrast these two contract types as ‘complete’ in the case of self-employment – meaning that the services provided are fully delineated in advance, ex ante,
and encapsulated in a ‘sales’ contract – and ‘incomplete’ in the case of employees – meaning that the detailed task content is to be defined subsequently, *ex post* (eg. Malcomson, 1997). However, incompleteness is only half of the story. Beyond a certain point, it is simply uneconomic to delineate the attributes of services in ever greater detail, particularly as their complexity increases (Barzel, 1997), and specifying every contingency in advance in a contract will make it unworkably complex (Williamson, 1985). Therefore, we need to think more of degrees of incompleteness, rather than of polar types, and to consider the factors that influence the cost of delineation, of which more shortly. The other part of the story lies in the governance mechanisms associated with different types of contract, and which enable the parties to adjust to changing circumstances. Unlike delineation which admits varying degrees of incompleteness, governance mechanisms, because they are logical constructs, tend to fall into discrete types. The employment contract is incomplete by design, as the British employment law scholar, Hugh Collins (2010), observed. The sales contract is a commercial contract, designed so that its obligations are tradable and can be sold on to third parties for whom only the words in the contract will count (Devlin, 1965: 44). The parties need to know specifically which set of rules will govern their collaboration: in this particular case, should adjustments to the changing mix of work tasks required be dealt with by managerial authority within an existing contract, or by renegotiation of a sales contract?

Figure 1 illustrates these considerations, drawing on Mintzberg’s (1979, 2009) theory of organisations as coordinating bodies. Organisations can coordinate by delineating tasks in detail, what he calls ‘standardisation’, or by delineating only broad elements, and relying upon a flexible process of ‘mutual adjustment’. The former is associated with bureaucratic work organisation, and the latter with what Mintzberg calls ‘adhocracy’. The cells in the figure also provide illustrations of work systems in which both forms can be found under either contractual type.

**Figure 1: Type of contract and degree of incompleteness.**

<table>
<thead>
<tr>
<th>Degree of incompleteness</th>
<th>Type of contract</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardised tasks that admit easy delineation.</td>
<td>Task may be simple, or standardised by occupational design (eg. Craft and professional occupations)</td>
</tr>
<tr>
<td>Mutual adjustment in the face of tasks that are hard to delineate and about which there is uncertainty.</td>
<td>Creative work organised on a project basis as in media and IT</td>
</tr>
<tr>
<td></td>
<td>Machine and professional bureaucracy</td>
</tr>
<tr>
<td></td>
<td>Flexible work patterns with a high degree of employee autonomy: adhocracy</td>
</tr>
</tbody>
</table>

The study of institutions and their constitutive rules comes into its own when it can be show that the ‘games’ they support lead to different outcomes. This can be seen by tracing the different strategies that the self-employment and employee contracts encourage. The rest of this section draws on the experimental evidence discussed earlier relating to reciprocity and to risky transitions in an attempt to highlight the different strategic directions taken by
firms and workers under each type of contract, and according to varying degrees of completeness.\textsuperscript{5}

Figure 2 sketches out a representation of the two contract types, and seeks to combine elements of both reciprocity and ‘risky transitions’. The purpose of juxtaposing the two contract types is to prepare the ground for a broader discussion of how institutional regulation of economic risks contributes to the working of labour markets. It also highlights the value of each contract type respectively to discontinuous and continuous production. It begins by considering the case where delineation is simple, and later considers where greater degrees of mutual adjustment are involved.

Suppose that a firm normally employs two tasks, A and B, but in different mixes according to its sales plans, and that its preferred option is to assign labour to the task that will generate the highest sales revenue in any given period. Suppose this is task A (which Figure 2 shows in upper case). It can hire self-employed workers at the start of each period for that task by negotiating a mutually satisfactory fee. Once the contract is signed, both parties are committed for the whole period. For the firm, the risk is that of committing to A, when B turns out to be the high revenue task. On the plus side, the agreement is for one period only, and can be negotiated from scratch for the next period. Alternatively, if the firm is uncertain about sales and does not wish to commit itself to a particular task, it can hire workers as employees on the understanding that they should be flexible and undertake either task as determined by the employer. Suppose that one task is more onerous than the other, so that its ‘cost of effort’ and the corresponding wage are greater. The parties can agree a wage that reflects the anticipated frequency of both tasks. The advantage for the employer is that as sales progress during the period, it can reassign workers to the task most in demand. It therefore avoids the potential losses associated with the sales contract from having contracted for the less profitable task. However, it comes at a price. For workers to agree to flexible assignment, they must feel confident that the firm will not cheat by assigning them to a third task C, outside their zone of acceptance, and involving a higher cost of effort. The firm might be tempted to do this if C generates higher sales revenues, but becomes the most profitable only if paid at the wage levels agreed for A and B. In other words, workers will agree to an employment contract only if they believe the employer will use task flexibility in good faith. Otherwise, they will insist on the more specific duties of the sales contract. This has implications for pay adjustments between periods. Under the pure sales contract, each period involves a fresh negotiation. Under the employment contract, negotiation at the start of each period is constrained by the need to convince the other party that it will act in good faith – employees will respect the need for flexibility, and the employer will not abuse it. In other words, both have to signal that they will stick to the agreement even if it means some short-term sacrifice of their self-interest. Driving a very hard bargain over pay at the start of each period would signal the opposite: prioritisation of short-term distributive over longer term conjunctive gains (see Section 7 below). This introduces a degree of inertia compared with the self-employed contract because deviations from what was agreed in the previous period need to be justified, and this could be intensified by the presence of loss aversion.
Figure 2: Movement between tasks and contracts for independent contractors and employees

Key: The upper case letter denotes the task with the highest sales, and the lower case, that with lower sales. The arrows denote transitions between periods. In State 1, A commands the highest sales, and in State 2, B does. The solid arrow denotes an expected transition within an employment relationship, and consent for the firm to adapt to the task associated with the highest sales. The finely dotted line to C denotes the possible assignment to task C followed by a quit towards self-employment in the next period. The pecked lines for independent contractors denote the consideration of a new contract together with the outside option.

Both worker households and firms have to be able to function over time as ‘going concerns’, and therefore both need coping strategies to ensure a continuous flow of activity and income. Within certain limits, the employment relationship offers this to both parties. Employees gain a measure of income security, and firms secure their supply of labour. It also provides a framework within which increasing reciprocity can underpin increasing joint investments in skills and knowledge. It therefore favours the ‘inside option’. For independent contractors, their bargaining power depends upon having an attractive ‘outside option’ so that the strongest motivation for workers is to neglect firm specific skills and to cultivate those that open up opportunities in other firms. Workers’ coping strategies to ensure continuity of income have many other manifestations as well. They generate a demand for institutional forms and social supports that encourage closure which conflicts with open, competitive, markets. The Webbs (1920) identified ‘mutual insurance’ as one of the responses of early trade unions to employment instability, and such activities have subsequently exerted a major
influence on social insurance régimes (Esping-Andersen, 1990). Perlman (1928: chs 6 & 7) documented practices of sharing or ‘communism of opportunity’ to stabilise employment among early forms of worker organisation, and also saw such insurance as a means of underpinning workers’ individual bargaining power. In London’s seasonal trades, Webb and Freeman (1912) document established patterns of workers ‘dovetailing’ activities to ensure a continuous flow of income despite having to move between many different jobs during the year, and these were often built upon membership of social and ethnic groups. These can be thought of as methods of self-help to sustain outside options, enabling workers to cope with labour market risks, and insecure livelihoods. A similar line of thinking can be found in Kerr’s (1954) famous account of ‘balkanised’ labour markets. Workers gain access to craft or occupational markets by peer recognition and by passing apprenticeship or professional competence tests which provide both a degree of occupational closure, and restrict access to work to insiders. Di Tomaso (2013) argues that labour market discrimination is a bye-product of self-help networks that are based on a degree of closure within social groups.

The next sections explore the internal dynamics of stable and of transient employment systems, seeking to show how central institutional supports are to their operation.

4. Sustaining reciprocity in the standard employment relationship

In a famous essay, Herbert Simon (1976) argues that in social and economic situations in which uncertainty and lack of relevant information make calculation of costs and returns difficult, it is advantageous to focus on ‘procedural’ rather than ‘substantive’ rationality. In other words, the parties concerned should agree procedures that enable substantive questions to be resolved once better information is available. Managerial authority to direct labour is one such procedure, and is key to his approach to the employment relationship. However, to be adopted, such procedures must be acceptable to both parties, and hence, one has to examine what protections are available to prevent managers from abusing task flexibility allowed by the open-ended content of the contract. There is an extensive literature on the pressures to ‘ratchet up’ work loads after hiring, and on worker resistance to this (eg. Lazear, 1986, Burawoy 1979), yet if the normal state of the employment relationship were to be one of endless haggling, it would surely tip the balance in favour of other contractual models such as independent contracting where mutual obligations are made explicit in advance, as was the case in the nineteenth century before the spread on the employment relationship, when sub-contracting had been the most common arrangement in industry (Mottez, 1966; Jacoby, 1985; Schmiede and Schudlich, 1976).

a) The zone of acceptance as a bargain

How can one combine the exchange process which is based on a market relationship with an open-ended contract which underpins an organisational relationship involving productive collaboration? Simon’s (1951) solution was to develop a simple bargaining model, a precursor to those used later to model union-bargaining over wages and employment (see survey by Oswald, 1993). The key is to identify a set of tasks and their respective shadow prices that are of interest to both parties, and represents a zone of potential agreement. As discussed in the previous section, the employer wants the job to include a variety of tasks that is wide enough to ensure it has both the necessary labour services available when its precise needs become known, and enough flexibility to avoid periods of under-employment. It also knows that on average, some tasks generate more value than others, and that some appeal more to potential employees than others. It therefore has to negotiate a wage level will attract workers for a sufficiently wide range of tasks for its production and cost plans.
This is illustrated in Figure 3. The vertical axis shows increasing wages, and decreasing profits. The horizontal axis shows tasks ordered according to some characteristic valued by both parties, such as their skill requirements or complexity. The firm’s profits increase for a time as these demands increase, but after a point they decrease because workers lack the necessary skills, and become stressed by excessive job demands. The thick inverted-U line $Sf_0$ represents the employer’s break-even point. Above this line, combinations of tasks and wage rates are not profitable, whereas they become more profitable as one moves south-westwards towards $Sf_1$ and below. Similarly, for workers, pairs of wages and tasks below their satisfaction curve $Sw_0$ are not worth taking, and their attractiveness increases as one moves towards curve $Sw_1$ and above. The area of potential agreement is shown by the tasks and prices enclosed by curves $Sf_0$ and $Sw_0$.

The two parties then agree a wage, $W_A$, and the zone of acceptance encompasses the tasks, within the area of potential agreement, between points $A$ and $B$ where it crosses the respective break-even curves of each party. For a slightly higher wage, the employer may gain access to some tasks to the left of $A$ which are still to the right of the worker’s break-even satisfaction curve. These might be less satisfying tasks for which the worker requires a higher wage in compensation. For a lower wage, the worker would be willing to trade tasks close to $A$ in exchange for more frequent assignment to tasks to the right of $B$.

**Figure 3 A negotiated zone of acceptance**

Source: Adapted from Simon (1951), and Marsden (1999: ch 1). For a simple explanation of the shape of the employer’s curve $Sf_0$ see Cartter (1959).
Simon’s model captures some very important aspects of employee’s jobs. First, the great majority of jobs involve multiple rather than single tasks. According to the US O*net database, compiled using field studies for its job placement services, the modal occupation includes about 20-30 different tasks, and many include considerably more. They also comprise a core of high frequency regular tasks, and a periphery of less frequently used tasks. This suggests that, in practice, before hiring, many workers and employers have a common experience of the task diversity of many jobs which may inform their negotiations. As discussed later, this is often related to established patterns of occupations and training. However, there is also an important fringe of tasks that increase flexibility for the employer, but could also prove controversial after hiring. Simon suggests, but does not really develop, the idea that such differences need to be resolved by goodwill with both parties seeking a solution that maximises their joint well-being. This is a necessary condition for efficient coordination by management authority, but it cannot be taken for granted. There are two reasons for this, and both lead to the conclusion that voice channels provide important supports to the operation of the employment relationship.

The first, already mentioned in explaining Figure 3, but also present in reality, is that tasks vary in their value to both parties, and both may seek to ‘cherry pick’ their most valued tasks at the expense of the other. Indeed, if there are significant quit or dismissal costs, then one party could go further and use these to pressurise the other outside the zone of acceptance (Task C in the previous section). The second arises from Simon’s assumption that management authority is exercised simply by selecting the tasks to which employees are assigned, which assumes that management has all the necessary knowledge and information. For simple work environments this may seem uncontroversial. However, where workers apply high levels of skills and professional judgement in their work, and where jobs, in Williamson’s terminology, are idiosyncratic, management frequently lacks this information and is dependent on workers’ cooperation. This may serve as a counter-weight to exploitation by management, but it also introduces the practical possibility that workers can pressurise management to assign them the tasks they prefer within the zone of acceptance, and in a more conflictual environment, both sides can try to extend the zone of acceptance to their own advantage by means of local bargaining pressures. There are several mechanisms by which such pressures can be kept under control with varying degrees of success.

b) **Sustaining reciprocity in the zone of acceptance**

Looking at cooperation within the employment relationship through the lens of evolutionary games (eg. Maynard Smith, 1982, Axelrod, 1984, Kreps, 1990), and following the weak concept of institutions, one may construct a model in which the parties will continue to cooperate provided that defection is always punished, and that the present value of the resulting sanctions outweighs that of any potential gains. Thus, if the employer knows workers will quit if pressurised into less favourable task assignments, possibly even outside their zone of acceptance, then provided the cost of replacing the worker exceeds the employer's gain, it will respect the worker's interests, and vice versa, and cooperation will continue as a self-enforcing behaviour pattern.

Nevertheless, the aggrieved party may not always be in a position to punish breaches, especially if the costs of applying sanctions against the offending party are not symmetrical. For example, in a period of labour shortage, the advantage may swing towards the worker, and in one of high unemployment, it may swing the other way. This could destabilise long-
run cooperation. Likewise, new management may destabilise patterns of cooperation especially if its orientation is more short-term than its predecessor’s. This suggests that while simple local cooperation/punishment mechanisms may contribute to the effective working of management authority under competitive conditions, they are not sufficient and other means of communication are necessary. Two stand out in particular: the need for a common language and for voice channels. Both can strengthen the position of ‘strong reciprocators’ as a common language of say job classifications makes obligations and breach more transparent, and voice provides scope to warn the other party about perceived breach long before sanctions are engaged.

c) Voice and a common language

Considering the need for a common language, Gibbons and Henderson (2012) highlight the problem of clarity and ambiguity of breach in relation to self-enforcing models of cooperation. Given a degree of job idiosyncrasy, and uncertainty about task needs, it is likely that in many instances the occurrence of breach will be ambiguous. As a result, the decision of whether or not to punish may not be clear cut. For reasons already discussed, this may undermine cooperation.

Formalisation may also be counter-productive. Slichter et al (1960) observed that complex rules about job assignments may encourage the haggling they were intended to prevent. Williamson (1975) has argued that complex contingent claims contracts are impractical for the great majority of employment relationships. An alternative solution is for the parties to rely upon commonly understood work rules, and to use these as criteria for boundary tasks to which workers and their line managers can easily refer (Marsden, 1999). This solution meets two essential conditions for boundary criteria: aligning workers' skills with employers' jobs, and being enforceable by ordinary workers and their line-managers. In essence, aligning workers' skills and employers' job needs can proceed either of two directions: from the skills workers bring to the job in which job design is adapted to workers’ skills, or alternatively, from the employers' job demands, which, in the words of one U.S. production engineer, 'moulds the man to the job' (Piore, 1969). Sengenberger (1985) identified the first as the 'training approach', and Lazear (1995: 86) identified the second as the 'production approach'. Work rules identifying boundary tasks for the first could relate to use of certain occupationally specific tools and procedures, and for the second, work associated with specific equipment or administrative processes used by the firm. To be credible, such items must relate to the everyday experience of workers' jobs, and be able to serve as symbols of Searle’s status assignments. The second principle is that such rules should be enforceable, that is breach should be readily identifiable. For this they should relate either to the set of key tasks making up a job, or, in the case of more flexible team working, to the function undertaken. Thus, the first makes its point of reference the tasks making up an individual work post, and the second, to tasks comprising a recognised collective function or team-based activity. Lam and Marsden (2016) show that these four types of rule delineating job boundaries map onto Henry Mintzberg's four patterns of organisation (Figure 4). The reason is quite straightforward. Mintzberg (1979, 2009) argues that the goal of organisations is to coordinate activity, and that this has to obey two principles: whether coordination focuses on inputs, and notably the detail of the tasks employees undertake, or on outputs, that is setting objectives which employees use their expert knowledge to achieve; and whether it proceeds by standardising work roles, and thus making them clear and predictable, or by mutual adjustment between flexible work roles. In summary, these provide a background of guideposts to assist the negotiation process Simon envisages for the employment relationship.
Without these, it is difficult, if not impossible, for the parties to establish boundaries to the zone of acceptance that are meaningful to them, and to identify breach when it occurs. Thus, in the employment relationship, an organisational logic imposes itself on the framework for exchange.

**Figure 4 Four approaches to monitoring work obligations**

<table>
<thead>
<tr>
<th>Enforceability criterion</th>
<th>Efficiency criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training approach</td>
</tr>
<tr>
<td></td>
<td>Task-focused</td>
</tr>
<tr>
<td></td>
<td>Tools of the trade/Professional bureaucracy</td>
</tr>
<tr>
<td></td>
<td>Function-focused</td>
</tr>
<tr>
<td></td>
<td>Qualification Operating adhocracy</td>
</tr>
<tr>
<td></td>
<td>Work post/Machine bureaucracy</td>
</tr>
</tbody>
</table>

Adapted from Lam and Marsden 2016.

Voice channels can play a very important part in sustaining goodwill cooperation within employment relationships, and their operation is facilitated by the development of benchmark rules to which the parties may refer. In recent years, there has been growing interest in various kinds of voice mechanisms: channels between individual employees and their managers (Rousseau et al., 2006), management-led employee voice, such as through quality circles and certain types of goal-setting and appraisal systems (Willman et al 2006, Cawley et al 1998, Levy and Williams, 2004), and collective, representative voice, and how all three interact with each other (Marsden, 2013). Whereas the breach/punishment approach depends upon a formulaic interaction in which these signals are clearly understood by each party, voice enables more nuanced communication about the kinds of task assignments that are felt to be fair, and within the zone of acceptance, and those which push one or other party towards its limits. They also allow the parties to adopt a problem-solving approach to mutual adjustment, and if necessary to consider renegotiating the initial deal if it is felt that conditions have changed significantly since the beginning. Such eventualities are recognized in employment law. Under English law, a legally enforceable contract underpins the employment relationship, but as Collins observes, its purpose is to ‘stabilize expectations’, not to prescribe solutions, which need discussion, and so Collins (2006: 139) argues for participatory channels. In similar vein, French employment law distinguishes between minor changes to the zone of acceptance, which are deemed to be part of the initial deal, and substantial ones that require renegotiation. The line between the two is determined in relation to the initial agreement, and hence to the intentions of the two parties when contracting (Lyon-Caen and Pélissier, 1988, pp. 306ff). In other words, employment law recognises that communication between the parties is an important aspect of the open-ended nature of the employment relationship, and that this depends upon adjustments that remain mutually satisfactory and without the need to treat each change as a potential breach. This still leaves open the question as to whether such communication should be individual or mediated by collective voice channels.
d) Biased risk perceptions, localised self-help, and restrictive rules

As mentioned earlier, Schmid (2006) argues that the biased risk perceptions of individual workers could mean that individual and small-group voice mechanisms are overly restrictive owing to loss aversion and conservative assessments of the potential benefits of change. As a result, they will predominantly choose exchange norms that protect the status quo rather than mutual adjustment. Schmid develops this argument in relation to preparing for labour market transitions, but it applies equally to decisions relating to the zone of acceptance and its flexibility. There has been a large number of studies of workplace and shop-floor organisation the support the potency of his argument. Gouldner (1954) developed his theory of the function of seniority rules over work assignments and progression in relation to workers’ desire to restrict management’s control of uncertainty in the work environment. Greater discretion gives management more power over allocations, and introduces greater uncertainty into workers’ lives, which affects the balance of power between the two groups. Crozier (1965) developed a similar argument: that the bureaucratic work rules of the French public sector were insisted on by workers, not the managers, in order to make their own lives more predictable and to limit management’s local power. Brown (1973) developed a similar analysis of the working of ‘custom and practice’ in the regulation of piecework bargaining in British engineering: workers and their shop stewards took management’s own decisions and omissions as a source of precedent. Thus in all three examples, the decentralised response of the weaker party was to try to bind the actions of the more powerful one by its own rules. Using their own localised self-help strategies, workers could create zones of greater predictability in their work lives which protected them against loss from adverse decisions by their managers, but the resulting rigidities led to greater organisational slack, and in the long-run to lower incomes for both parties.

Biased risk perceptions and general risk aversion can be overcome with a more holistic view of the potential benefits of change. An early champion of this view was Allan Flanders (1964) writing on productivity agreements which were agreed between management and higher level representatives, such as enterprise and in other countries, works councils. Schmid also argues that collective voice channels can overcome some of these biases, and develop more flexible norms of exchange. For example, works councils are in a position to take a longer view of the benefits of cooperation, and in Germany, by virtue of their codetermination rights, can also offer protection against short-term opportunism by management. By virtue of their often long-term relationship with the firm’s management, they can also form a good appreciation of whether or not it is acting in good faith (Jirjahn et al, 2011). As Scharpf argues, recognised institutions can allay concerns about ‘worst case scenarios’ that undermine cooperation. In this case, individual fears that management will exploit every concession would contribute to restrictive worker attitudes to flexibility, whereas higher level institutions such as works councils are able to oversee flexibility and ensure greater protection of worker interests (Kotthof, 2013).

5. Transient employment relationships supported by social networks and social norms

In recent years, the growth of precarious, and ‘non-standard’ employment, has been seen as part of the erosion of the post-war settlement in the advanced industrial countries. According to the ILO (2015: Fig 2.5), across the economically advanced countries, the share of permanent full-time employees has been shrinking while that of temporary workers and the self-employed has been creeping upwards. Moreover, the latter’s associated poverty rates
are two to three times higher, which could be read as suggesting both an element of recommodification of labour services, and that Simon’s model of the employment relationship assumes a more favourable balance of power than applies to many workers.

Closer inspection of ‘transient’ employment reveals a more complex picture. The term is used to designate jobs that have no implied duration beyond the current period, and so embraces both self-employment and temporary employment. A significant proportion of those in transient employment are highly skilled and highly educated, and are engaged in self-employed work that involves a high degree of mutual adjustment. Often their work requires both motivation and creativity. There is also another large group, concentrated in service and sales and elementary occupations, in temporary work whose jobs involve more delineation and control, where the dismissal and non-renewal threats are more common. This is illustrated by Figure 5 below. The vertical axis shows the degree of mutual adjustment during the contract, and the horizontal axis, the implied duration over periods. The hybrid contract forms show the lowest levels of incompleteness and mutual adjustment.

<table>
<thead>
<tr>
<th>Degree of incompleteness</th>
<th>Degree of temporal open-endedness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales contract</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Sales contract with mutual adjustment</td>
<td></td>
</tr>
</tbody>
</table>

In transient employment many of the same kinds of conflicts of interest arise as in permanent employment. These need to be resolved if collaboration is to proceed smoothly, but not necessarily by more detailed contracts. For example, at the high-skill end of project-based work, many such tensions can be seen in Paul and Kleingartner’s (1994) study of Hollywood screen actors in the period after the break-up of the large studio system. These included disputes over intellectual and artistic property rights, how many original ideas to hold back for future projects, ownership of manuscripts, tracking new sources of derived ‘residual’ incomes, actors’ concern about over-exposure and audience fatigue from use of their images, moral rights on the subsequent use of material, favouritism in selection, and so on. In the less glamorous activity of milk delivery, O’Connell Davidson (1994) illustrates the disputes between dairy owners and milk deliverers, both self-employed and employees, over qualitative aspects of performance: how far should deliverers take on additional products on behalf of the dairy or on their own account, how much credit should they extend to late payers, how much time should they devote to signing up new customers, and if they are more geographically dispersed, and so on? Should the dairies undercut door-step prices by those they offer to supermarkets? If the dairies own the delivery vehicles, how much should they
charge for rent and repairs? O’Connell Davidson’s conclusion was that although the precise manifestations of control problems differed between employees and self-employed franchisees, the same underlying problem of engendering cooperation despite divergent interests was present under both contractual forms, but needed to be resolved in different ways. It is to these that we now turn focusing first on high-skilled and then on low-skilled work. The difference compared with the employment relationship is that many of the long-term sanctions and rewards for good performance are absent because of the lack of implied continuity, and the need highlighted earlier to maintain one’s outside options to ensure a supply of work for the next period. One needs therefore to look for their functional equivalents in this different work environment. A number of empirical studies of this kind of employment reveal the frameworks by which the actors seek to ensure cooperative behaviour. We now look at these in the context of high-skilled and low-skilled transient employment.

a) **High skilled transient employment**

The high-skill end of transient employment is often associated with project-based employment. In this, workers combine for a particular project, such as might occur for a new media product, or an IT project, and may disperse when the project is completed. In this environment, projects generally focus on one-off creation, such as a film or a new piece of software, rather than the repeated reproduction of a good or service as in an automobile factory or a government tax service (Baumann, 2003). The roles of project leaders and collaborators are more fluid than those of managers and employees and may change between projects, as Piore and Sabel (1984) observed in analogous small firm collaborations.

The high-skill model requires leaders in a management role to spend considerable time defining and agreeing goals or outputs and agreeing necessary adjustments, and to abstain from managing the details of the work process. Later on, we discuss their relationship with investors. There is a limit to how far the contracts can delineate attributes of the tasks to be performed, which applies more strongly as task complexity and mutual adjustment increase. Thus, just as in the employment relationship, there is an increasing need for each party to be able to count on the other’s good faith. The question is how can this come about in the absence of the repeated exchange in long-term employment relationships. The most important functional equivalent cited in recent studies substitutes membership of an organisation with that of social networks and communities of practice. Even though projects may bring collaborators together for a short period before they disperse again, membership of the community of practice provides a similar continuity of memory to that provided within the firm by the employment relationship. Its members provide the collective memory that supports reputation for past collaboration, and an important source of information for future projects similar to the way social networks are used for finding jobs and for recruitment (Rees, 1960, Granovetter, 1973). They are also an important vehicle for mutual help as fellow members help each other to gain access to key information, but by the same token, may exclude non-members (Di Tomaso, 2013). In line with the previous discussion of weak and strong senses of norms, reputation can function in two ways: it can provide information about an agent’s past performance, and it can also convey information about their commitment to the norms and standards of behaviour of different occupational and social groups. The first might be associated with Williamson’s (1985) approach, whereas the second finds more echo in the work of Saxenian (1996) and others on the role of social networks for sharing ideas, and job information within the occupational communities of Silicon Valley. Likewise, the work of Jones and Walsh (1997), Sydow and Staber (2002), and Baumann (2003) illustrates similar processes at work in the networks of the US, German and British media industries.
Communities of practice also need anchor points, and these are provided, for example, in the Hollywood study by the work of the union in tracking down residual incomes for its members, in Silicon Valley, by links with major research universities, and in the German and British media industries examined by Baumann by links established through former public providers.

b) ‘Precarious’ contracts for low skilled and routine work

The situation is somewhat different for low skilled and routine work where delineation and control are easier to effect. Studies by Muehlberger (2007), and Muehlberger and Bertolini (2008) comparing contract work for tied sales agents in the Austrian and Italian insurance industries highlight the greater extent of dependency and control faced by contract workers whose situation is closer to that of the milk deliverers than the IT workers of Silicon Valley. Nevertheless, like the dairies, the insurance companies experienced the need to motivate the agents to work in the company interest, and to acquire firm-specific knowledge, such as about their products, and this was achieved in part by incentives which improved with longevity of the relationship, and partly by activation of the social networks among the agents. Like the milk deliverers there was also an issue of quality of business brought in, for example, customers who may be bad payers, or to whom products might be miss-sold with possible reputational consequences for the firm. By insisting on tied agents, the companies restricted the outside option associated with self-employment, but it enabled them to develop a form of repeated exchange, but with more tightly prescribed roles than for the highly qualified.

There can also be a relationship between the two types of transient employment. In a study of transient-type occupations in the UK, the author observed extended entry tournaments in which aspirant members of an occupation competed for entry well into their mid-career stage of life (Marsden, 2011). Such tournaments serve the industry well not only by providing incentives for ‘stars’ to flourish, and sanctions should they cease to perform, but also by attracting large numbers of aspirants who take on short-term low-paid work in the hope of a ‘break’ that will enable to cross the threshold into the occupation. That would bring peer recognition, and membership of the high status part of an occupational community (Jones and Walsh, 1997), but failure to do so leaves the worker in a crowded secondary labour market.

6. The production of skills and skill categories

Occupations and job classifications provide transparency and a language for both transactions and transformation, as they guide agents as to both the type investments that can be made, and the performance standards that can be expected from different categories of skills. Indeed, without occupational classifications, it is much harder for organisations to communicate their recruitment needs to job seekers, and to compare performance between employees since one may only compare those in similar jobs. They also provide a framework for investment in skills and training. They are a central attribute of a functioning labour market. Transparency concerning the capabilities attached to a particular occupation supports mobility between firms, as well as investments in skills. It also has profound implications for how we pay for such skills. Lack of transparency and uncertainty about the quality of training can undermine the development of a market and that may inhibit both employers and potential trainees from investing in the requisite skills, a theme discussed at length for other markets, for example, by Akerlof, 1970, and Podolny 2001. This section seeks to explain the different ways in which institutional supports address these problems, and what may happen when they fail. In essence, they provide assurance that the parties will act in good faith, by
protecting against short-term self-interested behaviour. As will be seen, these protections often work less by punishing certain kinds of behaviour, than by tilting rewards to favour cooperation. It begins by looking at some necessary supports for inter-firm markets for transferable skills, and then examines those for firm-specific skills, before looking at job flexibility and risk sharing among employers and employees.

**a) Conflicts over cost and quality of training for transferable and specific skills**

One of Gary Becker’s (1975) many contributions to labour economics has been to identify the fundamental differences in the nature of investments in skill formation depending on whether training leads to *general* skills that are transferable between employers, or whether they are *specific* to an individual organisation. In a nutshell, for training provided in the workplace, firms will not pay for general training because trainees can always quit and take the investment with them. The trainees should pay. The return on their investment is reflected in the higher pay they earn subsequently compared with unskilled workers. Conventionally, they pay for their training by accepting a low trainee pay rate below the value of their productive work during this period. In a competitive market for transferable skills, shortages lead to a higher wage and so induce greater worker investments until the market returns to equilibrium. This mechanism is not needed for firm-specific skills because the skills are not transferable. For these, firms train according to their production plans, but will normally pay their skilled workers enough to induce them to stay, and so will share some of the costs and returns. In addition, workers will be reluctant to invest heavily in firm-specific skills because their employer remains free to fire them.

Becker’s model makes an implicit assumption that training quality for transferable skills is transparent, yet this is often problematic for workplace training because the employer can always skew the mix between learning opportunities and productive work in its own favour. Without some agreed standards, and certification, both the quality of workplace training and trainee productivity are hard to determine, and a mutually satisfactory exchange becomes hard to ascertain. The employer’s method of providing workplace training in exchange for a low trainee wage comes under pressure: skilled workers anxious about trainees being used as cheap substitutes (Ryan, 2004), and aspirant trainees who suspect a ‘market for lemons’ (Akerlof, 1970). As Ryan has shown, if skilled workers fear competition from cheap trainee labour, they may seek to restrict trainee numbers directly or to choke off supply by pressing for a higher trainee wage. Uncertainty about training quality may also discourage trainees from investing in workplace training. Aspirant trainees who are uncertain about the quality of training at an individual firm would require a higher trainee wage, which then discourages the employer from providing high value training. As the firms providing the best training opt out, average quality across all firms declines. In such circumstances, firms may need either to assume a greater share of the cost of training while maintaining its transferability, or to find ways to make their employees’ skills less transparent to outsiders. Both responses put the occupational labour market under pressure. With the first, firms providing high quality training are exposed to free-riding by competitor firms that poach their skilled labour; with the second, the skills lose their transferability and become more like those of firm internal labour markets (Marsden, 1985: ch 8).

Although firm-specific skills are arguably more under the control of individual employers, some analogous problems arise. These hinge very much on the degree of investment required from workers. In Doeringer and Piore’s (1971) influential account, workers acquire such skills as a by-product of working, and learn skills by a process of osmosis. In this case, the
degree of worker engagement needed is fairly limited. These were mostly semi-skills in mass production industries. However, subsequent research on skill formation in Japan by sociologists such as Koike and Inoki (1990) indicates that there are alternative models of on-the-job learning, notably when the process involves more problem-solving than osmosis. This model, which Koike also contrasts with the US model, involves considerable worker investment as the process typically uses rotation among different tasks which have to be learned, enabling problem-solving at a more abstract level, based on analysis of relationships between tasks and how they fit into a broader integrated work system. In the development of what Koike refers to as ‘broad’ or ‘intellectual’ skills, the individual firm remains the central focus, but the ‘cost of effort’ that workers invest is considerably greater than in the older taylorist models. Reflecting back on Mintzberg’s typology, on-the-job training in Doeringer and Piore’s model is very much that of ‘machine bureaucracy’ whereas that described and conceptualised by Koike and Inoki belongs to the world of ‘adhocracy’ where coordination of work roles involves a high degree of mutual adjustment.

In between Becker’s two ideal types of transferable and specific skills, Franz and Soskice (1995), and Aoki (2010) argue that there exist several hybrid forms applying the two types of skills in combination, such as in large firms in Germany and Japan. In the German case, many technically advanced large firms build firm-specific structures on the foundations provided by the occupational model. General skills provided by apprenticeship training are supplemented by more advanced skills provided by individual employers. These, they argue, shift the balance in Becker’s model as the latter type of skills significantly reduce the transferability of the whole bundle of skills that individual workers apply in their jobs. These compound skills were the foundation of what Sorge and Streeck (1988) described as the German model of flexible specialisation based on ‘diversified quality production’. In this model, workers apply high levels of skill, often with a significant degree of redundancy with respect to their immediate job demands, but it provides them with the ability to work flexibly at a wide range of different tasks, and to develop a more integrated view of their work roles.

**b) Employer collective action to prevent free-riding**

Despite the theoretical strength of Becker’s argument, there is a good deal of evidence that many firms that train for transferable skills incur considerable net costs. Acemoglu and Pischke (1998, 1999) and Dustmann and Schoenberg (2008), show considerable diversity in Germany for apprenticeship training costs, with firms in some sectors, especially in engineering, contributing heavily to the cost of training, whereas in some other sectors, notably some services, they train at a profit. When firms are net contributors how can they avoid the damaging effects of poaching by other firms without undermining the occupational market?

One solution that maintains skill transferability is by employer collective action to organise training and to discourage poaching. Compared with Britain, where a similar apprenticeship system went into long-term decline, undermined by poaching problems, in Germany, the system survived and flourished. There firms and workers have been able to monitor training quality: within the firm by works councils and other workplace representatives; and at the local labour market level, by the activities of local chambers of industry and commerce. According to Drexel, publication of apprentice pass lists enables families and local employers to monitor the apprentice graduation rates of other local firms. Low rates could signal low quality training, or that the firm is free-riding on the training by other local employers.
A more systematic approach was adopted by Johansen (2000) examining the incidence of employer funded transferable training in Norway as a function of the ability of firms in different sectors to act collectively. Transferable skills have many of the attributes of public goods: other employers cannot be prevented from using them, they are non-excludable, and in the long run, a strong demand for such skills may stimulate increased supply, they are non-rival. Johansen then applies Olson’s (1971) theory of how group structure affects the capacity for collective action to mitigate free-rider effects (Figure 7). Accordingly, employers can more easily work together to provide employer-funded transferable training, when they possess an effective coordinating body, such as an employers’ association, and when the group is characterised by a small number of large employers capable of coordinating among themselves. Cross-tabulating these two characteristics, Johansen identified four sectors of the Norwegian economy: two where membership of the sector was dispersed, engineering which was dominated by small firms, and schools, and two where membership was concentrated: private insurance, where a small number of large firms could provide leadership, and county hospitals. Schools and hospitals were members of coordinating bodies, unlike engineering and insurance (Figure 6). As Olson’s theory predicted, engineering employers had the greatest difficulty to coordinate action over training, whereas the other three benefited from either large-firm leadership or well-functioning coordinating bodies. Thus, Johansen was able to show how the nature of the groups of employers in different sectors affected their capacity to resolve the prisoner’s dilemma and coordinate to train for transferable skills.

Figure 6: Group structure and coordination capabilities

<table>
<thead>
<tr>
<th>Presence of a coordinating body</th>
<th>Group membership structure</th>
<th>Concentration capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Dispersed</td>
<td>Engineering: inability to coordinate over skill provision leads to skill shortages</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insurance: large firms provide leadership on skill standards and training provision</td>
</tr>
<tr>
<td>Yes</td>
<td>County schools: coordinating body assists skill provision</td>
<td>Regional hospitals: coordinating body assists skill provision</td>
</tr>
</tbody>
</table>

Based on: Johansen (2000).

In the case of firm-specific skills, the nature of employer collective takes on a different nature. As the worker investments required increase, so does their vulnerability to a hold-up by the employer. For this reason, an established norm of long-term employment has co-evolved with this kind of work system encouraging workers to invest in such skills (Koike, 1997). Appelbaum and Batt (1994) argue that firms with innovative HR policies that require such investments are exposed to similar pressures to those for transferable skills, but they also face a problem of credibility with their own employees, which makes individual employer action more difficult to sustain. They are much easier to uphold when most of the firm’s peers follow the same policies, especially when the resulting obligations are not contractual. This is partly a question of peer norms, but also because promises of employment security are only put to the test during times of crisis when the firm forgoes short-term protection of profits in order to honour its commitments. For an individual firm, such events
occur irregularly. On the other hand, if the norm is widely observed, the number of occasions when firms maintain long-term employment despite financial stress is greatly increased. If one firm were seen to renege on its promise, there is a cost to its peers as it undermines the credibility of their promises to their own employees. Thus, as with transferable skills, there is also a potential collective problem for investment in the more flexible types of firm-specific skills. The capacity of Japanese firms to mobilise informal inter-firm links to coordinate action has been widely observed (see for example Cole, 1989).

c) **Employment protection and worker investments in skills**

Different types of employment protections may favour the development of different types of skills. Estevez-Abe et al., (2001) set out from similar assumptions to those of Becker, but they focus on the security factors affecting workers’ willingness to invest in skills, and notably their income and employment security. They argue that firms need to insure workers against some of these risks if they are to engage in skill formation, and develop a typology shown in Figure 7.\textsuperscript{13} They enrich Becker’s model by considering different types of transferability and specificity, and use of public educational institutions as well as workplace training. Where training quality is certified by educational institutions, workers or the state bear much of the cost, and skills are easily marketable. There is no burden of trust on employers, and so less need to tie their hands over job security and transferability makes for less demand for unemployment protection, as in the US model. In contrast, in Japan, where skills are strongly firm-specific, firms have maintained employment protection to encourage workers to share the cost and engage in on-the-job training, but firms may be less keen on income protection during unemployment as that raises the replacement wage and may favour quits. Where skills are transferable within an industry, this attribute reduces the need for employment protection, because there is a ready supply of alternative jobs, and mobility can be favoured by high income protection during unemployment, as in Denmark. In the hybrid case of Germany, the authors argue that employment protection is needed to encourage worker investment in upper-tier specific skills, and that within industry mobility also generates a demand for income protection during unemployment. In each case, the authors argue that the protective framework chosen stems from the joint concerns of both parties.

**Figure 7. Skill types and employment protection**

<table>
<thead>
<tr>
<th>Employment Protection</th>
<th>Unemployment Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Industry specific skills, eg. Denmark</td>
</tr>
<tr>
<td>High</td>
<td>Industry-specific, firm-specific mix: eg Germany</td>
</tr>
<tr>
<td>Low</td>
<td>General skills eg US</td>
</tr>
<tr>
<td></td>
<td>Firm-specific skills, eg. Japan</td>
</tr>
</tbody>
</table>

Based on: Estevez-Abe, Iversen and Soskice 2001: 154

Apart from the evidence examined by these authors, their findings have additional support from two European studies of the diffusion of modern forms of team and learning work organisation which confirms the link between the more skill-intensive learning model and supporting labour market institutions of the kind just discussed (Lorenz and Valeyre, 2005, and Holme et al, 2010).
d) **Risk sharing among employers to encourage labour flexibility**

Much of the debate about social insurance has been dominated by the idea that it has been imposed on firms, and creating a form of de-commodification of labour by reducing workers’ dependency upon labour market income (Esping-Andersen, 1990). Nevertheless, as Mares (2001) has shown, employers also may derive considerable benefits from the operation of social insurance systems. These are all the more important given the evidence about systematically biased attitudes to risk discussed earlier. By insuring their employees against certain work-related risks, employers find it easier to induce them to undertake certain types of work, to invest in certain types of skills, and to engage more fully with the firm’s objectives, provided benefits are in alignment with firms’ incentive systems. Firms may also benefit from insurance of employee incomes against the costs of industrial restructuring, including active labour market policies to assist retraining, as it can reduce opposition to such measures, and so facilitate economic change. This was indeed the philosophy of the redundancy compensation agreements and laws of the late 1960s and early 1970s (Reynaud 1969).

Mares (2001) proposes a model in which employers invest in social insurance systems according to the degree of redistribution of risks, and the level of control or oversight that they are able to exercise. In Figure 8, the degree of risk distribution of a policy is shown on the horizontal access, and the degree of employer control, on the vertical axis. The inverted U-curve traces out the maximum degree of control available to employers consistent with different levels of risk redistribution. At very low levels of risk redistribution, there is no insurance and hence no employer control, and at very high levels of risk redistribution, for example state universal benefits, there is no employer control. The peak of employer control is consistent with policies that involve a small amount of risk redistribution, for example, within a narrowly defined occupation or for a small group of firms, and in which employers can exercise a high degree of control over costs and the design of benefits. The space beneath the curve represents the viable combinations of employer control and risk redistribution of schemes. Enclosed within the curve, at low levels of employer control, and with increasing levels of risk redistribution, Mares lists a progression from an absence of insurance, through assistance, the contributory Ghent system, to universal benefits. Systems giving higher levels of employer control include contributory insurance which may be targeted to cater for varying degrees of risk redistribution, and firm-based schemes and early retirement which may be subsidised by varying degrees of risk redistribution to other employers or the state.

Mares argues that employers are interested in control for two reasons: cost control, and because they want social policy to be consistent with their own reward systems. Thus, employers of highly skilled labour are likely to want control in order to ensure that benefits and protection are earnings-related, and do not result in a compression of total rewards between skilled and unskilled workers. On the other hand, when their workforce is primarily unskilled, risk redistribution is of more use to them than control. Employers may also be concerned about risk redistribution, because some forms of employment are more exposed to the insured risks than others. In the historical examples Mares provides, sectors with a low risk of industrial accidents were reluctant to share risks with those with a high rate of accidents, and those with stable product markets were reluctant to share unemployment insurance with sectors that served highly cyclical markets. On the other hand, for the employers of highly skilled labour earnings-related insurance provides a double benefit of increasing the incentives to invest in skills, and of easing the pressure on laid-off skilled workers to take the first, unskilled, job available.
Early retirement is a particularly interesting case because of its widespread use in western Europe to enable firms with many older workers to restructure while redistributing the cost more widely to other employers and to the state (Palier and Thelen, 2010). An unintended consequence has been the resulting budgetary squeeze on other forms of public expenditure highlighting, and potentially exhausting, the limits of the capacity of the state and social insurance to underpin concerted action between unions and employer organisations, giving rise to what Streeck has called the long-term ‘self-undermining’ tendency of social institutions as they become over-extended (Streeck 2009).

<table>
<thead>
<tr>
<th>Control</th>
<th>Risk redistribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private, firm level policies</td>
<td>None</td>
</tr>
<tr>
<td>Early retirement</td>
<td>Assistance Ghent system Universal</td>
</tr>
<tr>
<td>Contributory</td>
<td></td>
</tr>
</tbody>
</table>


To conclude this section, the provision of skills in an economy inevitably involves risky investments by firms and workers, and given the evidence on systematic biases in the appreciation of risks, confronting these on an individual level is likely to lead to an undersupply especially of sophisticated transferable and firm-specific skills. Each of the selected areas in this section, training quality, the distribution of its costs, the need for inter-employer coordination, employment protection and social insurance provision, has sought to highlight the role of collective action, that is coordination among employers and workers to avoid potential market failures. The section has also shown that collective solutions are not always available, for example, when group structures, as in the Norwegian engineering sector, do not favour coordination. On numerous occasions also, Thelen (2004, 2014) has highlighted the historically specific nature of the various national institutions, and Palier and Thelen (2010) and Streeck (2009) have warned of the dangers ‘over-loading’ of these institutions as they also create potential free-rider problems, for example, over-use of early retirement payments to facilitate economic restructuring. In the next section, these institutions are examined as ‘composite actors’, capable of assuming different functions according to the interests and needs of their members, but also often needing to find compromise between the competing demands of different groups among their membership.
7. **Unions and employers as composite actors**

Like Manchester United mentioned earlier, individual unions and employers are organisations pursuing their objectives, but to carry out their functions in the arenas in which they are active, they need to take on an institutional mantle that ensures the capacity to undertake certain kinds of actions on behalf of their constituents, and to enter into certain types of agreements. To be allowed to engage in collective bargaining, for example, in most countries, employment law requires that unions be independent of the employer, and that the employer should refrain from actions intended to undermine this. In most continental European countries, there is a functional division between works councils elected by a firm’s employees, and unions that depend upon recruiting members from many workplaces. Generally, the former engage in conjunctive bargaining over productivity related issues, whereas distributive bargaining over pay is assigned to the unions. The former brings a peace (no-strike) obligation, whereas for the latter both unions and employers are at liberty to engage in strikes and lock-outs. John Dunlop’s (1944) path-breaking work on the economic analysis of collective bargaining set the mould for much subsequent theorising built around three key actors, unions, individual employers and the public authorities, and their respective objectives, but it was heavily influenced by the New Deal framework of the US. Company-level bargaining favoured the transposition of monopoly bargaining theory to industrial relations. It was a model that never fitted the division of labour among European representative institutions where collective bargaining has been organised at industry level between industry-level employers’ associations and unions seeking agreements to cover many different firms. In much of continental Europe this meant that industry bargaining acted, according to Streeck’s observation, to take certain issues ‘out of competition’ between member firms. Thus, competing on low wage costs and minimal training budgets may be an option for individual employers under the US model, but it is largely foreclosed under strong forms of industry bargaining. This is not to ignore current debates among observers of German industrial relations concerning the erosion of industry bargaining in Germany, and changing balance between unions and works councils (Hassell, 2014).

When Scharpf (1997) argues that unions and employers are composite actors, he was challenging the view that the union and the firm might be considered as unitary actors as in Dunlop’s bargaining model, and as used subsequently by many labour economists. He was stressing the manner in which institutional arrangements structure these collective actors and their capabilities. He was arguing also that their actions were constrained by the ‘rules of the game’, and that these were known to both parties. One could therefore legitimately analyse their choices using elements of game theory, but informed by the institutional constraints and capabilities affecting them. The background to his work is provided by neo-corporatist analysis of peak-level collective bargaining and coordination, and its significance to the present context can be best understood by starting from that stream of research. It also provided the intellectual background for the subsequent work on ‘varieties of capitalism’.

The early work on corporatism and wage bargaining is best represented by Calmfors and Driffill’s (1988) model of the relationship between bargaining centralisation and coordination on the one hand, and the trade-off between inflation and unemployment on the other. These authors envisaged an inverted U-shaped relationship between centralisation and the ‘Okun index’ of economic misery (the sum of a country’s inflation and unemployment rates: Okun, 1981: Ch 8). In brief, the underlying theory was that decentralised bargaining at company level placed maximum responsibility for any adverse cost and employment consequences on local managers and union representatives. Once unions and employers begin to group
together in their respective organisations, a collective action problem emerges. Higher levels of organisation bring greater bargaining power and drive a wedge between excessive claims or concessions and their negative impact on employment and sales. It is only when bargaining is centralised that both sets of bargainers once again face the full economic consequences of their actions, and have the capacity to coordinate to avoid them. Hence, Calmfors and Driffil argue, intermediate levels of organisation experience the worst outcomes of inflation and unemployment. Centralised bargaining also increases the scope for coordination with government economic policy, and this occasioned the analyses of scholars such as Scharpf (1997) applying game theory by analysing the pay-offs to each party, but also their respective constraints, in a three-way interaction. If their electors and central banks permit, governments can choose between redistributing the costs of excessive pay bargains by means of permissive macro-economic policy which may increase the Okun index, and tight monetary policy which would translate potentially inflationary pay increases into higher unemployment thus bringing adverse consequences home to those bargaining over pay (see: Iversen, 1999).

With the decline of inflation, interest in macro-economic coordination has gradually shifted away from wage bargaining towards collective frameworks for business performance, corporate governance, skills, and job flexibility. These developments fall under the banner of the socio-economic foundations of different ‘varieties of capitalism’ (Storz et al, 2013). The authors who contributed to the Hall and Soskice (2001) volume identified two contrasted types: liberal market economies, typified by the US and the UK, and coordinated market economies typified by Germany and Japan. Although there has been much debate about the number of different ‘varieties’ (see Hancké et al., 2007), the Hall and Soskice version has the merit of parsimony. In essence, liberal market economies rely more strongly on the use of competitive market forces to coordinate economic decisions, and a set of complementary institutions governing corporate governance, skill formation, industrial relations, and inter-firm networks. In contrast, coordinated market economies rely more on collective action through organised representatives of business, workers, governments and other stakeholders. Neither model has an overall competitive advantage over the other. Instead, each has a comparative advantage in different types of economic activities, so the outcome is specialisation and trade rather than head-to-head confrontation. For Hall and Soskice, the more loosely structured liberal market model is better adapted to activities benefiting from disruptive innovation, whereas the coordinated model has a comparative advantage in those requiring incremental innovation. The authors document this with a study of specialisation of patenting activity in different sectors in the US and Germany: in the US in IT and in Germany, in engineering.

At the heart of the concept of ‘varieties of capitalism’ is an idea that institutions can help to align the choices of the actors by making certain compatible strategies more profitable for them, thus giving rise to mutually reinforcing choices. The consistency of these strategies is critical to the genesis of system-wide comparative advantage (see Höpner, 2005, and Crouch et al., 2005). This mechanism is well illustrated by Amable et al (2005), and it also enables us to return to the earlier discussion of the employment relationship and job flexibility. They show how different institutional arrangements favour different strategies. For example, in wage bargaining decentralisation favours pursuit of distributive goals, whereas coordinated bargaining favours a more problem-solving approach in the workplace, and in corporate governance, exposure to short-term financial market pressures favours pursuit of quick returns, whereas more concentrated ownership enables investors to monitor managers’ performance in greater depth and evaluate harder to measure long-term policies. Figure 9
cross-tabulates these and shows their predicted pay-off matrix for the different combinations of strategies. In this case, the best outcomes for both parties are provided by decentralised bargaining coupled with short-term financial strategies, and cooperative bargaining coupled with long-term financial strategies. The other cells mean that one or other party loses out.

**Figure 9 Complementarity between industrial relations and financial markets**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low financial market pressure: long-term strategies</td>
<td>Capital loses out</td>
<td>Both may prosper &amp; benefit from ‘problem-solving’ negotiation</td>
</tr>
</tbody>
</table>

Based on Amable et al. (2005).

By modelling the strategies of firms and organised labour, Amable and colleagues are also able to explore the conditions under which institutional complementarity might break down, leading one or other party to break out of its established relationship. They represent the respective long term bargaining strategies of labour and employers in terms of the current wage (profit) share and expected future returns, the two expressions on the right hand side of the equations below:

\[
W_t = w_t + s_t W_{t+1}/(1+r_w) \\
\Pi_t = (1-w_t) + s_t\Pi_{t+1}/(1+r_f)
\]

\(W_t\) and \(\Pi_t\) represent their respective objectives for wages and profits, and \(w_t\) and \(1-w_t\) the current share of wages and profits. Expected future returns are a function of the discount rate \(r\) (\(r_w\) for the unions and \(r_f\) for the firm), and firm’s survival probability \(s_t\) (which also reflects its long-term profitability). Unions will seek the highest current wage share subject to the present value of future wage incomes (low values of \(r\) imply higher valuation of future returns), and the firm’s survival probability. The latter depends on how long investors will continue to support it. Inconsistent time frames between labour and investors, where \(r_w\) and \(r_f\) diverge, lead to instability as the party seeking short-term gains is more willing to sacrifice long term gains to get its way. If both parties seek short-term gains the relationship will focus on distribution and will be conflictual, and if both seek long-term gains, there can be a stronger focus on cooperation and mutual gains, such as those arising from more flexible work roles and a more adaptable zone of acceptance.

The authors give the argument an additional twist by considering relative bargaining power - whether cooperation is out of weakness – and the effects of shifts in bargaining power, such as might accompany globalisation. They represent this by the distributive shares, \(w_t\) and \(1-w_t\). Thus, if investors switch their objectives towards short-term gains, weak but cooperative unions will seek to break out as they see the benefits of cooperation being eroded. In contrast, if unions are strong and cooperative, under pressure from investors, management will seek to break out. One may also consider the effect of such changes on job-level cooperation, discussed earlier, as high levels of task flexibility require confidence that both parties share the same commitment, manifest in consistent and high values for \(r\) and \(s\), that is, a long-term orientation to the pursuit and sharing of joint benefits. A hard line on
bargaining over shares is likely to be taken as a signal by the other party that short-term goals prevail.

It is tempting to think of the pressures from investors as reflecting general and exogenous pressures eroding the distinctiveness of the institutional configurations of the two varieties of capitalism. On the other hand, a closer reflection on the nature of skills and knowledge and their effect on corporate governance patterns suggests forces making for their stability. Aoki (2010) argues that governance arrangements have to satisfy the needs of three key actors: workers, managers and investors. Under the classical principal-agent model, the principal’s discipline over agents resides in the ability to dismiss them and hire substitutes. Aoki argues that this characterises his archetypal American firm, the ‘A-firm’. For this to work, the principal must be able to hire workers and managers with the requisite skills and knowledge from outside. In this model, expert knowledge resides with managers rather than workers. In contrast, the research of Cole (1989) and Koike (1997) indicates a common pattern of ‘distributed intelligence’ in Japanese firms, Aoki’s archetypical ‘J-firm’. Aoki argues that learning by joint problem-solving involving both workers and managers leads to the development of joint knowledge assets in which the contributions of each are diffused within work groups and are interdependent in their application. In his words, they are ‘fused’. As a result, substitution from outside is very difficult. It should be said that this form of on-the-job skill development is very different from that found in the classic studies of western internal labour markets in which specific skills were tied to individual work posts, and problem-solving was the work of engineers (Doeringer and Piore, 1971, Maurice et al 1982).

By restricting the dismissal threat, fused knowledge assets require a different form of corporate governance to avoid a potential coalition between workers and managers against investors, and Aoki argues that this is provided by relational contracting and the group structure of large Japanese firms. He argues that large German firms practicing the hybrid occupational and specific skills model of Franz and Soskice also have to balance the need for workers to invest in firm specific knowledge and investor protection, and consequently, the institutions of company-level codetermination which involve worker and investor representatives play a key role in providing sufficient protection for each party to facilitate cooperation, giving rise to his archetypal ‘G-firm’. Following from the previous discussion, one might ask whether similar pressures for increased short-term returns would cause firms to break out from the ‘J-’ and ‘G-firm’ models. Such pressures appear to lie behind a widely observed segmentation within these economies, between sectors for which fused knowledge assets confer a comparative advantage, and those for which it either does not, or has ceased to. Thus the present author finds that the ‘G-firm’ model and its related representative institutions have been more robust in sectors that have corresponded to the German economy’s traditional comparative advantage, but has eroded in other less-skilled and service sectors (Marsden, 2015).

Finally, these models all relate to the classical employment model, but what can be said about interest protection in the project-based model with its transient employment relationships? As many of these are high-skill ventures between collaborators, project leaders and investors, there is a risk that control of skill and knowledge by project leaders and collaborators will generate exploitation of investors, thereby discouraging their participation. Based on work on collaborations in Silicon Valley, Aoki suggests that in the ‘S-firm’ teams of collaborators often come together with a project in mind, and then seek finance from venture capitalists, in a similar way to university colleagues bidding for research funding. Guide posts for the realisation of different stages of the project when additional tranches of finance can be released provide some protection to investors against collusion among the
project leaders and collaborators. An interesting question is whether the more decentralised liberal market economy model provided a more supportive environment for this kind of institutional innovation.

8. Conclusion: labour markets as social institutions

A good deal of recent debate about the nature of economic institutions, for example in the *Socio-Economic Review*, has been distinctly ‘macro’, focused on different ‘varieties of capitalism’, looking especially at the institutional frameworks that underpin them. This chapter has taken a different approach, starting from the ‘micro-foundations’, building on the role of institutions in the two central activities of labour markets: transactions and transformation. This approach opens the way for a more direct engagement with the analytical tools used by economists. At the same time, it makes way for a more ‘actor-centred’ approach which has been gaining ground among economic sociologists. Together, these enable one to understand better how macro processes build up from the choices and decisions of individual actors, but which reflect both their interests, and the norms of exchange, the rules, which they choose or are obliged to follow. Reflecting on these interests makes it possible to build on the advances in our understanding of human behaviour in economic transactions, notably in relation to reciprocity and risk, and to see how the adoption of certain ‘rules of the game’ can foster the first, and mitigate some behavioural consequences of the second. This opens up a richer palette for our understanding of collective action.

Although the ‘weak’ sense of institutions has demonstrated its usefulness by providing parsimonious explanations of how social rules, such as the not-undercutting norm, may emerge out of decentralised interaction among individual agents, aquably, it assumes too much, by relying purely on calculation of the parties’ immediate self-interest. As North observed, the knowledge that one’s partners in transaction or transformation will follow certain rules greatly simplifies the task of cooperation. If agents know that positive reciprocity is a widely respected norm, then they can more readily engage in ‘tit-for-tat’ cooperation than if they have to calculate continuously whether or not it is still in the other party’s interest. Institutional rules rely on more than a balance of self-interest. They involve an element of obligation. Keeping one’s promises, and contracts, is a key norm of exchange. If people kept their promises only when it was in their interest, the institution could not function, yet we observe that it is a very widespread practice.

There is another limitation of the weak sense of institution, namely that reliance on calculation of self-interest leaves little room for reasons of the kind given when an agent invokes ‘context C’. The reason why employees should follow their managers’ instructions is that they have a contract of employment. The context is different for a self-employed worker whose contract does not admit such actions. Giving reasons is also a key component of voice relationships. The boss can explain to the employee that task A has now become the more profitable one and that it is in their joint interest to switch their effort to that one. The employee can explain refusal to undertake Task C pointing out that it is not part of the job, and it is not necessary to issue an immediate threat to quit. Context also plays an important part in Kahneman, Knetsch and Thaler’s study of fairness constraints on maximising behaviour. For example, presenting a cut in the wage offered as relating to incumbent employees as opposed to lesser wage offers to new hires evoked different responses from their sample: the first was predominantly judged unfair, the second, mostly acceptable.
A final limitation of the self-interest calculus arises from the inherent difficulty that Gibbons and Henderson identify with breach in an employment relationship: its flexibility often makes breach very ambiguous. Context also enables one to judge more easily whether actions are in good or in bad faith, and to identify the ‘trigger’ situations for unleashing punishment strategies. Relying solely on individual perceptions and punishment strategies to sustain the balance of the relationship is very risky. Allowing small and persistent breaches to accumulate may benefit the transgressing party, and encourage further breach. It may also influence the norms governing the exchange as what is tolerated can become ‘normal’. Explaining a decision in relation to its context – ‘we had to assign you exceptionally to task C because of a production emergency’ helps the other party to assess whether it is in good faith, before considering whether any action should be taken. Treating norms of exchange as rules that are expected to ‘trump’ immediate self-interest, and considering seriously the role of context in determining their application enriches the palette of tools available to understand the role of institutions in the transactional and transformational processes associated with labour markets. In particular, they open the way for a fuller consideration of labour markets as social institutions.

Karl Polanyi’s (1944) *Great Transformation* has played a defining role in the recent growth of socio-economics, and in particular, his idea of a double movement, of an oscillation between greater freedom for markets and greater protection of the social fabric in which their actors are embedded (Krippner, 2004). Such movements are visible in the ‘commodification’ and ‘de-commodification’ of labour in different labour markets: the growing precarity of employment in some sectors, and the maintenance of long-term employment in others. Polanyi argued that the political and economic chaos of inter-war Europe was the fruit of having allowed the pendulum to swing too far towards self-regulated markets in the preceding decades. Stiglitz, in his preface to Polanyi’s 2001 edition, warns that markets can require greater and faster price and wage adjustments than the social fabric can bear. The social norms that underpin labour markets and employment change, but often gradually, as do the coping strategies that workers develop to deal with uncertainty. Given their importance in simplifying transactions and transformation activities, the risk of imposing over-rapid change is that it leads not to a world of great flexibility, but one that Durkheim described as ‘anomie’: normative breakdown accompanied by a breakdown of the predictability on which much of economic and social life depend.
9. References


10. Endnotes

1 The author would like thank Bruce Kaufman for many helpful suggestions along the way, and also participants at the 2015 SASE conference for their comments and suggestions.

2 See Smith 1786, Book V, ch 1 part 3. I am grateful to Saul Estrin for pointing this out to me.

3 Solow also considers the case where a high equilibrium wage is associated with high unemployment which reduces the probability of finding a new job for the unemployed and reduces their chances of benefitting from the equilibrium wage.

4 The difference between the weak and strong senses can be further illustrated by considering the stable behaviour patterns in the animal world that led Maynard Smith (1982) to develop his analysis of ‘evolutionarily stable strategies’, that is, behavioural strategies that will spread and drive out other possible strategies. In his example, animals develop territorial behaviour for example for food and mating because these strategies are more likely to lead to survival of their species than the alternatives. He considers the ‘hawk’ strategy in which an animal will always escalate a conflict to get its way, a ‘dove’ strategy in which it may tussle with the other, but will never escalate, and the ‘bourgeois’ strategy which is a hybrid of the two: play ‘hawk’ if you are the incumbent against an interloper, otherwise play ‘dove’. Provided that bourgeois players always retaliate against interlopers and the cost of escalated conflict is greater than the gain, their strategy emerges as superior to the other two. Human beings may also play such strategies, but the point about its robustness in the animal world is that it makes no reference to status assignments of the kind emphasised by Searle as an essential ingredient of institutions. From our current knowledge, only human animals have the linguistic sophistication to assign status rules.

5 In an interesting experimental study, Provenzano (2013) shows that employees were more likely to show positive reciprocity if their tasks involved a degree of discretion than if they were tightly defined.

6 In their classic text on industrial democracy, the Webbs (1920) identified two key methods of trade unions: collective bargaining and mutual insurance. The first was taken up by scholars such as John Dunlop (1944), who was among the first to use the economic model of bilateral monopoly to analyse collective bargaining between a union and a firm or an employers’ association. Perhaps because unemployment and sickness insurance were subsumed into the welfare state from the mid-twentieth century, the method of mutual insurance went into eclipse.

7 The O*NET data base established by the US Bureau of Labor Statistics to support its labour market services has built up a description of the detailed work activities associated with each of the nearly 1000 occupations. It does this by means of field surveys of managers and workers who not only list tasks but indicate how important they are. http://www.onetcenter.org/aboutOnet.html The O*NET classification is of particular interest because it is constructed from field surveys of workers and managers about the tasks comprising their jobs and how important they are. It therefore provides a rough image of the zones of acceptance associated with detailed occupations in the US labour market. Some of these are core tasks, and regularly undertaken, but others are related, and while important to successful execution of the former, are harder to specify and to predict, and likely to vary more between workplaces. http://www.onetcenter.org/dl_files/DWA_summary.pdf

8 Simon uses the term zone of ‘acceptance’ implying that it has been agreed, but Barnard (1938) makes the stronger claim that workers are indifferent between the tasks within it, and uses the term ‘zone of indifference’.

9 Deakin and Wilkinson (1998) provide an interesting three country comparison of how small firms can work with varying degree of contract incompleteness, and mutual adjustment, depending on the legal framework.

10 Tolbert (1996: 339) defines a ‘community of practice’ as follows: ‘individuals who actively share a core body of tacit knowledge that is necessary for the execution of concrete, everyday work tasks . . . provide the foundation for occupations’.

11 Acemoglu and Pischke (1998) argue that an externally recognised qualification is an important incentive for workers to engage in workplace general training, as in the German apprenticeship system. See also Höpner, (2005).
The Japanese labour economists, Shirai and Shimada, argue that the consolidation of Japan’s system of ‘life-
time’ employment occurred during the first oil crisis in 1973-74 when their employees saw the lengths to which they went in order to preserve employment at a time when many western firms were cutting jobs in order to protect profits.

‘Because investment in specific skills increases workers’ exposure to, only by insuring against such risks can firms satisfy their need for specific skills’ (p. 181).

Research justifying the corporate governance and financial market side can be found in (Aguilera and Jackson (2003), and Aoki and Jackson (2007).