Cambridge Journal of Economics 1994, **18**, 181–196

The transaction costs and benefits of the incomplete contract of employment

Kate Cartier*

Transaction costs help explain the incomplete contract of employment, but analysing costs only is static and restrictive. A more complete and objective analysis of the contract requires considering the transaction benefits as well. This is done by applying the theoretical and empirical findings of cognitive psychology to examine how workers learn by experience and use their cognitive abilities. This dynamic analysis of the relation between labour and production at the microeconomic level of the job and the worker modifies theories of labour market behaviour and of economic production and growth.

1. The incomplete contract of employment

The contract of employment as a fundamental and pervasive institution of exchange is a crucial link between the quality of labour and production. However, the contribution of labour to production has been neglected by labour economists preoccupied with the organisation and functioning of internal and external labour markets and wage determination as matters of supply and distribution, while many studies of production assume that labour is a homogeneous quantity, or can be ignored. One condition of equilibrium in neoclassical theory is the assumption that the marginal productivity of labour can be identified and rewarded by wages. In the same tradition is the argument that workers surrender discretion to their employers in exchange for the avoidance of risk and the acceptance of command (Knight, 1926; Alchian and Demsetz, 1972). What determines work effort has yet to be fully explained. The diverse and vast literature in industrial relations, industrial and social psychology, industrial organisation, management, and sociology, and the many attempts to summarise and interpret it (e.g. Whyte, 1955; Argyris, 1964; Vroom and Deci, 1970; Pugh, 1971; Child, 1969; Kanter, 1983; Handy, 1985; Duru, 1986; Tanguy, 1986; Rose, 1988) demonstrate how work effort eludes explanation. Some of these authors agree that the results are disappointing. With few exceptions (e.g. Fox, 1974) the predominant focus of discussion is on extrinsic motivations to work, tacitly assuming that behaviouralist psychology is sufficient. It implies that work effort can and must be controlled directly by monitoring, influenced by the terms of the wage bargain, or subtly conditioned by such things as working conditions, so that managerial intervention is necessary and feasible. Studies of the content of work reveal the lack of an appropriate theory of what is done and how.

Manuscript received 5 March 1990; final version received 26 August 1992.

*New Hall, Cambridge; and Paris. I thank Professor William Brown, Professor Robin Matthews, Dr Jill Rubery, Dr Joan Whitehead and two anonymous referees for their advice and suggestions.

0309–166X/94/020181+16 $08.00/0

This article attempts to construct a conceptual framework to fill this *lacuna*, by examining how people work under the incomplete contract of employment. Economists have recognised the incomplete nature of contract between agent and principal (e.g. Leibenstein, 1966, 1975 and 1982; Mallet, 1980; Hart, 1987; Hart and Moore, 1985 and 1988). Theories of implicit contract have been formulated to try to explain the macroeconomic phenomena of unemployment and sticky wages (e.g. Marsden, 1986), whilst theories of efficiency wages contain assumptions about effort (e.g. Lindbeck and Snower, 1985). This article offers a more thorough examination of both the costs and benefits of the contract.

The contract of employment is part of the exchange of mutual rights and obligations that binds people together in society (Hobbes, 1651; Locke, 1689; Rousseau, 1762). The social contract maintains order by tacit understandings not to harm each other's life and property, which are symmetrical, but negative. They are complemented by implied agreements to work together for the support, development and enjoyment of life, which are asymmetric, given varied endowments of ability and property, and positive. A social contract is a necessary condition of the joint use of labour, capital and other resources, and of the division and specialisation of labour. The contract of employment implies an agreement to contribute to and share output, but does not specify what each will give or receive. What is exchanged is a matter of individual ability, judgment, negotiation and relative power, the subject of persistent conflicts of interest and deep controversy, and a factor to be included in any reckoning of performance.

The contract of employment is inherently incomplete because it is impossible for the bilateral arrangement between employer and employee fully to specify what work is to be done and how. Transaction costs are less for an activity within the firm because it can negotiate one contract with suppliers brought within it rather than a series of outside contracts, and exercise some control over them instead of having to monitor each purchase (Coase, 1937). In the absence of transaction costs neither the firm nor the law would have any purpose (Coase, 1987).

Contracts of employment incur relatively high transaction costs: searching for and negotiating with the other partner; doing so for a series of partners, renegotiating with each of them; and the costs of enforcing the contract if one party is in breach, which are the time and expenditure on litigation and the loss of trust. The greater the heterogeneity of labour, the greater the transaction costs to the firm of hiring new workers and the greater the interest in continuity of labour supply. Whilst workers may benefit from changing jobs they, too, incur transaction costs, and so want to maintain continuity of income. In only a minority of jobs (such as sub-contracting or independent craft work) is there a contract *for* employment which attempts to specify all the conditions to be satisfied, such as price, quantity, quality or delivery, under future contingencies. Most jobs are under a contract *of* employment which specifies only some aspects of the employment relationship (Rivero and Savatier, 1981). Those aspects which are specified are usually agreed in collective bargaining between employers and trade unions, not with individual workers. Their terms, which vary, usually concern such matters as the recognition of trade unions; conciliation procedures; standard wage rates and how wages are to be paid; working hours and shift work; vacations; the termination of employment; safety procedures; health care and recreation; and benefits for injury, sickness, retirement and death. According to empirical surveys of employment agreements (e.g. Oswald and Turnbull, 1985) their terms rarely mention the nature or the quality of the work to be done, or how it is to be done. The absence of reference to the content or quality of work

is not an oversight by negligent negotiators. The other conditions can be specified and enforced in law. The content of the work to be done and the quality, intensity and pace of work effort, can usually neither be specified, nor enforced in the same way. Moreover few workers are entirely independent. It has been suggested that ' . . . if it is possible to isolate what each worker produces, workers are paid by the piece and the contract is complete' (Blaug, 1983); but even under systems of piece-rates it is impossible to define or control work (Brown, 1973). The element that remains unspecifiable is the benefit to be realised by entering into the contract. The contract of employment is necessarily and inherently incomplete.

As Marx observed (1865) 'What the worker exchanges is not directly his work, but the ability to work'. What is exchanged is not a specified level of effort nor a predetermined set of results but a worker's ability in return for an agreed initial rate of pay and, sometimes, unspecified rewards in the form of expectations of an *ex post* review of pay or promotion. The contract is incomplete because neither the content of work, nor the future conditions under which it will take place, can be anticipated. The quality of effort is outside the contract (Mallet, 1980). The link between pay and effort cannot be specified: the firm sets pay and the worker chooses effort.

This distinction between the ability to work and the quality of work delivered separates the interests and the obligations of the two parties. Whereas wages are negotiated in money terms, each seeks a different result in real terms: the employer considers the prices of other factors and of the firm's output; the worker judges wages in relation to the prices of those goods and services he or she buys. The parties have no common *numéraire*. The contract also separates the two parts of the exchange in time: the employer begins to pay for the ability to work in advance of receiving the benefits of the contract, or knowing what precisely the benefits will be or what they will be worth in the product market. The employer has to extract a future result that cannot be determined in advance, from each employee's ability to work which is unknown, using methods of management whose effects are uncertain. Consequently, there is no reason to expect a simple relationship between pay and productivity (Brown and Nolan, 1988). The employment relation is inherently controversial.

The contract of employment is established in terms so general that there is uncertainty about what each party will do or can be expected to do. Each party has a general commitment to the other which leaves room for discretion, i.e. the exercise of choice in how to perform. The generality and uncertainty increase with the expected duration of the contract (Coase, 1987).

Because the contract is incomplete there is a mix of prescription and discretion (Simon, 1965). Jaques (1956, 1976) and Fox (1974) define discretion as something allowed by the employer, and describe hierarchies of discretion in the economy. Leibenstein (1975) conceives of the worker interpreting the job and deciding both what to do and how. Empirical evidence about discretion is slight, but some is given by a recent work attitude and work history survey about the incidence and content of discretion in seven levels of jobs (Burchell *et al.*, 1993). Many workers believed they had a significant degree of discretion: *c.* 60% of full-time workers and 51% of part-time women thought they had a great deal of choice, whilst only 15% of the former and 24% of the latter thought they had little or none. All workers agreed that the exercise of discretion was the single most important factor affecting work effort. Managers recognised that all workers had quite a lot of discretion, but rated it much less than did all groups of workers, especially lower skilled operatives and people doing routine clerical

work; and they attached more importance than did workers to responsibility. The workers had a more intimate knowledge of their jobs, which they may have developed by exercising discretion. Only a weak relationship was found between pay and discretion, congruent with the contract being an exchange of pay for the ability to work, and not for subsequent effort.

Discretion makes the incomplete contract dynamic, allowing the worker to develop the job and the employer to realise the gains of productivity in the form of reduced unit costs. The open nature of the contract allows the worker to bargain in terms of the relation between pay and productivity *ex post* and capture gains in the form of higher pay or improved working conditions or both. If an association is observed between pay and productivity it is the result of an *ex post* settling up (Fama, 1980), which obscures the disjunction between pay and effort in the original bargain. The open contract gives more scope for the dynamic development of output and of earnings than would a specific contract, and sets new standards of performance which the employer can use in renegotiating prices and qualities in contracts with subcontractors (Rubery and Wilkinson, 1981).

2. Performance under the incomplete contract

If it is the ability to work, not future work effort, that is exchanged for wages the transaction benefits of the contract are uncertain. What, then, determines the quality of work done?

Management techniques that lean more towards control than to licence rest on the premises that management knows the production function, workers are a source of error in lacking technical knowledge or the expected commitment to work, and workers must and can be controlled. However, experience reveals how elusive is the control of work. *Job descriptions* are primarily to help design pay systems or indicate authority structures, and cannot fully specify *ex ante* a job or how to do it. Some descriptions state objectives only, recognising that the contract is incomplete. Some firms have no written job descriptions. Nor can the division and interdependence of labour be described adequately. Firms devote considerable time, effort and expense to manipulating the composition of their personnel by selecting the workers most likely to perform as they wish, and influencing departures. The use of formal qualifications as a proxy for ability is extensively regarded as a cheap, though unreliable, method of selection. Selection by *ex ante* indicators of performance encounters the invisibility of work and is no guarantee of future performance. With some insight into the functioning of internal labour markets, the Peter Principle postulates that everyone is eventually promoted to the level of his or her own incompetence, and that this is a reason for economic inefficiency (Peter and Hull, 1969). *Direct supervision* is often impossible, even when authority is reinforced by superior technical knowledge. *Controlling time* present is still thought by employers to be necessary and feasible, accepted as fair by most workers, and may influence the sense of commitment to work; but its effect on the quality of work remains unknown. *Job-specific training* in the firm is a powerful, but subtle, instrument of discipline, apparently unnoticed by many workers and their unions. Its effect is to secure conformity to established procedures and norms. The ambiguity of the term 'instruction' conceals an imperative sanctioned by further training or dismissal if the worker fails to learn required practice, by a wage premium if training is completed satisfactorily, by inspection or quality control, and by prospects of promotion to supervisory functions—all of which

perpetuate the system. The function of training as a hidden means of discipline is contained by the existence of discretion.

Uncertainty arises under the labour contract from the outset. Because the production function cannot be specified, the relationships between labour and all other factors, and between time worked and output, are unknown. It is impossible to know to what extent or how the jurisdiction of the employer over the worker can be transformed into a given result, or at what cost.

Uncertain performance arises from *bounded rationality*. Human beings intend to be rational, but encounter limits in formulating and solving complex problems and in processing information (Simon, 1962). They have a finite but unknown capacity for knowledge, understanding and reasoning. The stock of knowledge is continuously increased by research, but society has not learned how to manage the 'information overload'. Searching for and using information has costs, which limit the information that can be obtained if some time is to be reserved for its use. Knowledge can never be certain: it is a principle of scientific method, under the null hypothesis, that a theory can only be refuted or not refuted, never proved or disproved. It is possible to move towards certainty by testing a sequence of hypotheses, but never to attain it. Bounded rationality is the secular equivalent of original sin: finite human beings cannot know what is right. Even if all contingencies could be specified in the contract the worker might not recognise them if they happened, and the employer might not know if any failure by the worker were due to the worker or to an exogenous event (Stiglitz, 1975). Because knowledge has costs and human beings are finite, decisions made in the course of work can only be made in a state of uncertainty; and because some of them have a probability of being wrong, risk exists.

Uncertainty is central to the understanding of organisations (Knight, 1926; Arrow, 1974). The organisation of the firm reflects an implied contract to allocate risk asymmetrically between employer and employee. The allocation of risk explains the existence of authority within the firm, but not its specific structure. The argument that management exists because of the need to supervise workers, and that its form depends on the degree of difficulty of supervision and on the extent to which human assets are specific to the firm (Williamson, 1984), fails to the extent that some aspects of competence are specific to the worker, not to the firm. Uncertainty entails choice, and so requires decision. A worker has to interpret a job by deciding how fast to work, how to allocate time among the component tasks, and how carefully to perform each of them (Leibenstein, 1975) and so makes decisions which incur risk. They entail balancing the potential benefits of a decision against the paid-out costs, the opportunity costs and the possibility of failure. The decision-making process is not the exercise of authority by the entrepreneur or an executive in a process of risk-taking, but of workers jointly and severally helping manage risk.

3. The commitment to effort

Because the contract is incomplete, the quality of work is partly a matter of the exercise of individual will (Mallet, 1980). But many employers state explicitly or imply by their behaviour that workers shirk, either by slacking or by working towards their own interests. Shirking is so frequently mentioned in the academic literature that several economists, too, appear to treat it as a general phenomenon. 'The . . . factory system is nothing more than a response to the control problem of shirking in team production'

(North, 1984); ' . . . employees . . . retain an interest in working less hard or slowly enough to ensure they are paid overtime rates' (Malcomson, 1984). The only constraint is the fear of retaliation: ' . . . it always pays for a party not to perform its part of an agreement if that non-performance does not lead the other to do the same or is not penalised'. (Malcomson, 1984). Opportunism has been defined as 'self-interest seeking with guile': ' . . . an effort to realise individual gains through a lack of candour or honesty in transactions' (Williamson *et al.*, 1975).

The notion of shirking is problematic. Two contradictory propositions rest on the casual and partial observation of anecdotal evidence: because workers shirk they must be monitored, and they shirk because monitoring is impossible. Opportunistic behaviour is also found among employers who seek an unfair advantage, or managers who hire to create empire, power and prestige. But the principal difficulty with notions of shirking, cheating or of opportunism is their negative bias. It is evident that laziness, cheating, absenteeism, pilfering and misappropriation, abuse and exploitation exist, and may be induced if the original contract were made under duress; but to postulate a general principle of shirking rests on a view of human behaviour and a lack of objectivity which does not contribute to serious economic analysis.

The assumption of shirking is at variance with theories of efficiency wages. They implicitly assume that the quality of work is a function of wages, that the worker knows more about the job than the employer, that it is costly or impossible to monitor work, and the firm makes wage decisions unilaterally. Therefore, the firm pays a rent in the form of relatively high wages to keep turnover costs low, or induce effort (e.g. Lindbeck and Snower, 1985). There is much evidence of the converse of shirking, responsible behaviour under the contract, and of one or both parties delivering more than could reasonably be required or expected. There is no reason to assume that it is more or less common than shirking. Conscientious behaviour in unsupervised work may reflect respect for oneself, for the contract, for other workers or the employer; or responsibility under the law or for the standards and sanctions of a professional association. It can spring from the sense of identity, craftsmanship, satisfaction and pride which work can confer on workers at all levels.

A more neutral view of effort is that the behaviour of the two parties to an exchange is reconciled when the exchange is frequent. Adam Smith reasoned that 'Whenever commerce is introduced into any country probity and punctuality always accompany it . . . because the trader deals so often that he finds honesty is the best policy' (Smith, 1763). Marshall, too, held that modern methods of trade implied habits of trustfulness and the power to resist temptation (Marshall, 1920). It has been suggested that the behaviour of the two parties to an employment contract is reconciled by an 'invisible handshake' of trust and loyalty (Blaug, 1983; Garnier, 1984) which replaces the 'invisible hand' of the market (Okun, 1981). Reciprocity is the key to a balanced view of pragmatic behaviour: ' . . . it saves a good deal of trouble to have a fair degree of reliance on other people's word' (Arrow, 1974). Yet trust is elusive. A 'first-mover' who breaks trust is likely to find the other party retaliates; and trust is hard to create when it is absent.

It is because the production function is neither specified or known, not because of laziness or opportunism, that 'X-inefficiency' exists (Leibenstein, 1966). Each worker makes 'an internal moral compromise' among personal standards and desires and pressures from the firm and from peers; and chooses an 'effort position' which reconciles it with the choice among alternative activities, the pace and quality of work, and the time devoted to each task. The effort position can be improved, but is unlikely to reach an

optimum because of costs of information, disutility for the individual worker, traditional practices that are hard to change, and vertical constraints from supervisors and horizontal ones from peers. So, a worker tends towards an inert area, and the distribution of effort among all workers is likely to be far from a Pareto optimum. Once a manager establishes a stable, cooperative group he or she is likely to be satisfied.

Because workers acquire different abilities and degrees of expertise, collective action can extend the boundaries of individual rationality (Arrow, 1974). Its effectiveness depends on the extent to which teams have a common purpose which may not be explicit. At best, personal relationships are the foundation of harmonious effort which can develop into synergy. At worst, they can be dysfunctional to the point of taking precedence over the economic purposes of the firm. Between these extremes, workers may develop a team effort if their various efforts are mutually consistent. Workers who exceed the standards of the group as well as those who do less often earn the disapproval of the group (Leibenstein, 1975). Efforts are dispersed about an accepted norm which the employer recognises in adjusting pay.

There is in every firm a mix of cooperation and conflict, influenced by incentives of pay, promotion, status, honours, assessment, excessive or insufficient supervision. Conflict is inevitable between the interests of the team and those of the individual workers, each of whom belongs to an interest group outside the firm, and has different values and motives for working. Constructive discord can arise when specialists in a firm have different professional knowledge and principles or values which conflict. Disagreements between the chief executive who has to make decisions in a state of uncertainty and accept the risk of being wrong and a firm's advisers on science, law and accounting who reason from different premises, are well known. Perfect collective rationality is impossible. Performance is inevitably sub-optimal.

We may conclude that, whilst some transaction costs arise if workers shirk, others cannot simply be ascribed to enforcing either individual or collective action. If workers exercise discretion in the interest of the firm, effort becomes a benefit of *implementing* the contract. An objective view of the incomplete contract and of the dispersion of human effort recognises that the contract carries both transaction costs and benefits. The boundary between the firm and the market is determined by the extent to which the firm can realise the benefits of bringing transactions within the firm, and secure the gains which can be realised by individual and collective action.

4. Learning from the experience of work

The transactions benefits of the incomplete contract are governed by the ability of the worker to exercise discretion in interpreting the job. How, then, is that ability acquired? Economists recognise the existence of learning-by-doing, but have neglected the nature of the process and the cognitive abilities thus acquired. Theories of learning and knowledge have been discussed elsewhere (Cartier, 1993), and only their main implications for work are summarised here.

Modern cognitive psychology offers an explanation of the relationship between the human mind and experience that is not necessarily conscious or explicit at all times. Its central proposition is that new, initially unstructured observations and ideas from the environment are perceived, organised and interpreted by the mind in relation to a context of prior understanding. It derives from Gestalt theory, developed since the 1950s, which posits that all information coming into foreground of the mind is shaped and fitted into

a background, i.e. placed in a context. The understanding acquired at any time constitutes a whole that has properties, in varying degree, of regularity, symmetry and coherence, but which can never be complete. To describe this coherence of acquired understanding, different disciplines use a variety of terms such as representation, schema, vision, construct, model, theory or paradigm, all of which express the same concept: a modifiable frame of reference which organises knowledge into forms available for use. This schematic understanding differs from scientific theory to the extent that it can lack rigour.

Behaviouralist and cognitive theories agree that experience is a pillar of understanding, but differ in the ways that it is interpreted in the mind for behaviour. It is a semantic misunderstanding of no little consequence to assume that because economics is a behavioural science, it is behavioural*ist* psychology which is relevant to economic behaviour. It is, however, more relevant to extrinsic conditioning of the mind than to intrinsic discriminating behaviour. The process of acquiring cognitive ability is quite different from learning as automatic responses to extrinsic stimuli.

People continuously reformulate their understanding because curiosity and the will to learn are intrinsic (Bruner, 1966). If information coming from experience is congruent with an existing schema it is assimilated. If it is incongruent, either it is rejected or the schema is altered to accommodate the incoming information (Piaget, 1969; Whitehead, 1976). Learning proceeds by the continuous marginal reformulation of understanding or by the conception of a new paradigm (Kuhn, 1968; Lakatos and Musgrave, 1970), and not by simple accumulation, as some economists appear to imply in discussing learning-by-doing (e.g. Arrow, 1962; Rosen, 1972).

Cognitive abilities have three key characteristics. They may be subconscious: everyone has knowledge of which he or she is not aware, and can think and act without being able to explain precisely how (Polanyi, 1967). People can discriminate by deciding whether, when or how to use their knowledge. Cognitive abilities are idiosyncratic: differences in individual understanding are inevitable because everyone starts with an unique vision, is exposed to different sequences of data from the environment, and progressively modifies that vision by a personal combination of reaction and reflection.

If cognitive processes happen during any kind of experience, they can be acquired from experience during work and utilised in carrying it out. This is illustrated by the occupational wisdom of several proverbs: 'don't go against the grain'; 'strike while the iron is hot'; 'stick to your last'; 'we learn from our mistakes'. Although knowledge can be acquired during work, its form, content and utilisation are empirical questions; but until recently the study of it has been neglected for lack of appropriate concepts, methods, interest or esteem. The empirical evidence from recent case studies in several countries, summarised elsewhere (Cartier, 1993), demonstrates some of the manifold characteristics of learning from work. The cognitive abilities acquired varied according to the range and levels of ability of a worker and the components of a job. The ability of a worker to interpret a job and to devise his or her own solutions on his or her own initiative and responsibility was manifest in ingenuity in devising a variety of techniques to economise time and effort. Cognitive abilities were progressively extended by work itself, even if the work appeared to remain unchanged; but they developed more extensively when unfamiliar production problems created a marked incongruity between them and prior understanding. In a variety of types of work, without formal instruction or close supervision, workers were to a greater or lesser degree able to learn for themselves from the margin of uncertainty in a job.

Learning and work are inseparable. *Learning* proceeds by an alternation between the *observation of experience* and its *interpretation* in the mind. *Work* can be defined as the use of this knowledge to alter the human environment. This marginal alteration of the environment by work marginally changes a worker's experience, he or she observes and interprets it, so the cycle of learning is continuous. A job and its incumbent develop mutually and simultaneously. Learning is a continuous joint product of work. Coping with the margins of uncertainty in the contract of employment develops a worker's cognitive abilities. Although a worker can make mistakes, they become a source of learning if they are corrected. It is only an apparent paradox that imperfect knowledge is a source of both error and learning during work. Labour is not a source of error to be avoided or corrected by supervision. A worker has the ability, whether potential or realised, to discriminate in the use of understanding, to question it, store it, and decide whether to use it or not, when, and how. The utilisation of continuously corrected cognitive ability enables a worker to resolve dysfunctions, discover more effective solutions to work problems, and help innovate in methods and products. The quality, pace, seriousness, and relevance of effort are internal to each worker.

A degree of thoughtful effort in work is intrinsic in a worker, making quality a variable which complements the quantity of labour. It happens among all workers, and at all levels of work, from the apparently simplest manual work to the highest expertise, which differ in degree, not in kind. Learning-by-doing is not the restricted process often said to be more appropriate to people of limited ability doing lower level work.

The content of cognitive ability, like learning, cannot be directly observed by either a worker or a supervisor, and may only be revealed when a worker successfully draws upon it to cope with a problem beyond the range of previous experience. Cognitive ability is much wider than the common notion of skill. It is the full set of abilities which makes it possible to deal with contingencies. These abilities include substantive technical knowledge and affective behavioural attributes. The latter are: the ability to modify one's own cognitive abilities, acquired sometimes by explicit study, and sometimes by deriving from the experience of solving one problem the ability and confidence to solve others; the ability to discriminate in the use of knowledge for purposive action; and the ability to acquire and use knowledge in conjoint work. These are the reasons why workers are idiosyncratic, labour is heterogeneous, and cognitive ability is a dynamic quality.

5. Mobility and asset specificity

The redeployment of workers within and among firms is a condition of the development of comparative advantage in the long term. For a satisfactory empirical analysis of this question, an appropriate set of concepts of mobility is needed. The focus on observable movements of labour and impediments to them which has dominated the recent debate on economic restructuring and labour market flexibility (e.g. Boyer, 1988; Pollert, 1991) gives little insight into the kind of changes to work or why they are made. Although functional flexibility is recognised it is often treated as mobility among jobs within a firm. A closer analysis of the tasks in a job became necessary for the design of automatic contol systems (e.g. Edwards and Lees, 1974; Moray, 1979; Verdier 1984). Studies of '*les savoir-faire*' have recently been made in France, but do not explain how the ability to work is acquired (Cartier, 1993). However, the dynamic endogenous development of cognitive abilities during work helps explain the nature of mobility.

It has been argued *a priori* that mobility is inhibited by asset specificity, implying that the asset is specific to the firm, even to a job (Baumol, 1986), i.e. that costs are sunk and cannot be fully recovered elsewhere. Williamson (1984) compares it to site specificity or to physical assets, claiming that it 'arises in a learning by doing fashion', and that the supplier who wins a contract has an advantage when it is renegotiated. The application of the concept of investment to expenditure on training (e.g. Becker, 1964) or to labour (e.g. Thévenot, 1985) recognises that work helps generate a continuing stream of output and income. Although the term 'investment' is used rhetorically by employers and individuals who think it confers some advantage, employers neither acquire an equity interest in workers, nor calculate profits as a return to labour or to skill: they commonly treat labour as a variable, current cost; wages and training expenditures are charged to the profit and loss account, and workers are hired and fired at short notice.

The concept of labour as a quasi-fixed factor appears as a compromise between treating it as either a capital or a current resource, recognising the evidence of mobility among workers ' . . . who are dismissed rather than resold, and are not bolted down' (Okun, 1981). The observed mobility of workers among activities, occupations and firms suggests that there are few sunk costs by the firm. More important obstacles to mobility are separation costs, such as the firm's hiring and firing costs, or a worker's own investment in housing and family ties. Unskilled workers are to some extent homogenous, and substitutable within the firm and in the external labour market. Middle level managers, craftsmen and para-professional workers are rotated among jobs or establishments to develop their versatility; whilst professional workers, scientists, engineers and technologists are not only highly mobile in a large, even international, labour market, but are often redeployed within the internal labour market. There is extensive anecdotal evidence of workers changing jobs or occupation, often more than once.

When asset-specificity is used to explain immobility and the construction of the internal labour market it refers to worker as an asset of the firm, because of some specific substantive knowledge of the job or of the firm's technology, organisation or practice which is thought to inhibit mobility. The fallacy in this argument is the tactic assumption that learning is the acquisition of specific subject–matter knowledge, and that work is the use of it. If this were so, workers would become increasingly immobile among firms and within them too. Another interpretation of asset specificity is needed because mobility can be observed even when a worker's knowledge is ostensibly highly firm- or job-specific. The production function is more idiosyncratic than is assumed in neo-classical theory; and coping with the margins of uncertainty confers the opportunity to acquire the cognitive abilities that constitute assets of value to the worker as well as to the firm.

Individual understanding is a framework that is continuously reformulated. In each of a sequence of tasks a worker acquires some understanding specific to it which is modified and used in the next. As the sequence continues the understanding becomes progressively detached from the tasks; and if the tasks are done in varying situations the worker's understanding becomes detached from them too. In this way each worker develops a personal corpus of understanding, in the form of a set of rules of thumb, ideas or principles, which becomes increasingly general and can be more widely applied, and also extends the ability to apply it. The process by which a worker develops the job also develops the worker. The idea that knowledge is accumulated (Arrow, 1962) is at variance with the theory that it is continuously reformulated.

Although some of the substantive technical knowledge a worker acquires about the organisation, practices, social codes, technology, markets etc. of one firm may have no

apparent direct utility elsewhere, the ability to acquire it can be useful to other firms. Personality attributes, such as the ability to learn, to discriminate in the use of knowledge, to recognise and solve problems, often learned in the experience of work, can be used in other tasks within a job, or other jobs, and makes it possible for a worker to adapt technical knowledge to new purposes. These acquired abilities cannot be vested in the employer, only hired.

Cognitive abilities are assets specific to each worker which make functional and occupational mobility possible. The concept of a worker as an asset specific to the firm or a sunk cost, is a debatable interpretation of the phenomenon of continuing job attachment. As long as the employment relation continues between a worker and an employer each has a degree of bilateral monopoly, they share the use of cognitive ability, and they earn shared quasi-rents. Moreover, the relationship changes dynamically as cognitive ability develops. The transfer of knowledge from one situation to another is possible if a worker perceives the relationships within each of them which are common, rather than their ostensibly common features. The ability of a worker to acquire and utilise cognitive ability for new purposes in new circumstances develops personal autonomy and mobility and makes the redeployment of labour possible among tasks within a job, jobs within a firm, and workers among firms. This ability to use cognitive ability elsewhere enables a worker to consider alternative employment and prospects and wages in the labour market.

This argument suggests some changes to labour market theory. Mainstream neoclassical theory assumed labour to be homogeneous and mobile. Internal labour market theories rest on the assumption that the specific nature of technology defines jobs. When asset specificity conceives a worker as an asset of the firm who becomes less mobile because knowledge becomes increasingly job- or firm-specific it is the counterpart of the traditional assumption which posits a negative relationship between mobility and learning on the job. But if cognitive abilities become increasingly personal attributes specific to each worker they make workers heterogeneous and mobile. Consequently, technology or 'skill' are less plausible explanations of the internal labour market than internal administration.

6. The role of labour in production and growth

Recognising the simultaneous, mutual and dynamic development of their attributes of workers and their jobs also fills a *lacuna* in existing theories and empirical studies of production and growth. Production is inadequately represented in theories which assume a fixed technical production possibility set, and a determinate equilibrium conditional on rational actors maximising utility. Theories of human capital in the neoclassical tradition disaggregate labour by levels of educational attainment which take no account of the effects of experience on the development and use of individual cognition. Such simplified and static approaches are incompatible with observed variations in the incidence, forms and rates of growth. The net transaction benefits of the contract arise from the symbiotic relation between labour and production, endogenous in the experience of identifying and resolving problems and dysfunctions which progressively transforms the cognitive abilities of a worker and the content of his or her job. This makes the capacity of the firm unknown and open-ended. Rather than a given set of quantities of labour, fixed assets and technology, it can only be defined with respect to the abilities of workers to use and modify understanding, which makes labour and capital complementary, not substitutes.

If the attributes of workers and jobs are indeterminate and develop dynamically, a more open and flexible conceptual framework is needed to permit a microeconomic analysis at the level of jobs and workers. More appropriate is the notion of a production system using varied resources whose attributes are incompletely known and which yield an indeterminate outcome (Mallet, 1980; Wilkinson, 1983). The constraints of rationality and utility maximisation are not relevant. Human beings, none of whom is representative, make a difference to production. Leibenstein's proposition (1966) that inputs, whether fixed or variable, yield an unknown performance can be taken further. Labour has an unknown specification and yields an unknown performance which comprises increments in production and in cognitive ability.

Persistent human intervention of varied and uncertain quality is both cause and consequence of economic growth. This helps reformulate the conceptual foundations of theories of endogenous growth. Although the germ of the idea of endogenous growth has lain dormant for over two centuries, the dynamic role of workers' cognitive abilities has been overlooked in explanations of increasing returns. Young (1928) ascribed increasing returns to new resources and uses, science and population growth, but considered labour an obstacle, 'resistant to change'. Although Kaldor suggested learning-by-doing he made no use of it in his explanation of increasing returns and cumulative causation (Kaldor, 1960, 1972). Arrow suggested that the product of learning was accumulated knowledge embodied in capital, not in workers' cognitive abilities (Arrow, 1962). A large volume of empirical work on manufacturing progress functions (learning curves) has focussed on their statistical regularity, offering only weak hypotheses about why costs fall with experience; but it is plausible to interpret these data as evidence of the reiterative process in which understanding is generated during the experience of production in the interaction between a worker and a job (Cartier, 1993).

If learning from work were subject to diminishing returns (Smith, 1776; Arrow, 1962; Rosen, 1972) it would have no force as an engine of growth. If this were general it would dismiss learning from experience as insignificant and the quality of labour as a factor to be reckoned. However, the fallacy is the assumption that technology is given subject-matter knowledge. That might not be plausible even for highly prescribed jobs in well-defined systems, for even apparently simple jobs are sets of complex tasks carried out in contexts that can vary, and so be opportunities for understanding to develop. The process by which cognitive abilities and jobs are transformed by learning from experience suggests that it is reasonable to assume that learning is subject to at least constant returns.

One reason why growth is endogenous is that the expansion of output itself reveals new kinds of contingencies in the content and circumstances of production, extending the margins of uncertainty in contracts of employment. The more problems and dysfunctions appear beyond the range of prior experience the more can returns increase and cognitive development become a powerful engine of continuous growth. Under such conditions, the bias towards technical progress which arises in the original experience of dealing with new problems and dysfunctions is reinforced. The very presence of error is a source of persistent learning which advances, even though it remains incomplete. If learning is endogenous in the individual and in the firm, the development of knowledge is endogenous in the nature of work: improved solutions are retained in the mind of a worker, or in the memory of a firm in its procedures, the design of capital assets, or the specification of materials and products. The continuity of interaction in problem-solving between the experience of work and its interpretation by workers can progressively reduce error within a current technological paradigm, and generate innovation and a

continuing reduction of unit costs. It is possible to hypothesise that the current rate of growth determines future growth. It is also conceivable that it is the *breadth* of new experience generated by growth which creates the conditions for periods of sustained growth that is dynamic, irregular and non-equilibrium, without necessarily determining a given rate.

The possession of a stock of cognitive abilities, even if invisible, can enable the first entrant to an industry, or the first to introduce a new technology, to begin to accumulate a competitive advantage, progressively improve its market position, and remain a market leader. This may encourage the spread of industrial concentration. When one firm can retain it, the private benefit of investment is greater than the social, and total investment is less than optimal. However, it is unlikely that one firm can retain the entirety of this knowledge: it can leak to suppliers when subcontractors are provided with specifications, to competitors if workers move among firms, and to customers if it is visible in the product.

Economies due to the endogenous acquisition and use of cognitive ability can be distinguished from economies of scale, and evidence that is as yet fragmentary suggests they may be relatively more important (Cartier, 1993). Unit labour costs can still fall while investment, the scale of production, and factor prices remain unchanged. Moreover, changes in resource utilisation not only reduce unit costs: they often improve quality. The neglect of endogenous learning by workers can lead to mistaken explanations of endogenous growth which ignore its roots in microeconomic behaviour.

Although technical progress is promoted by learning in resolving dysfunctions, there are several constraints on efficiency. Interpreting experience during the contract of employment has complex transaction costs and benefits. Because individual and collective rationality are bounded, learning is imperfect and knowledge is incomplete. Prior experience is only a partial guide to action when problems arise outside the scope of it, especially when technology advances quickly. Knowledge derived from production may be invisible, and may or may not be used by the worker or the firm. Managers may hesitate to allow workers to develop or use their cognitive abilities for fear of uncertainty or loss of their authority. There is unlikely to be a social optimum in the volume and distribution of knowledge. Labour is an unknown resource with elastic but finite abilities, so it is inevitable that neither workers nor the firm can maximise utility. The theoretical conditions of an equilibrium level of output are not met. There is no reason to suppose that growth is smooth or optimal. If performance is sub-optimal the experience gained during it is an incomplete guide to future action; but it is the only guide there is.

7. Concluding remarks

Early steps in the economics of institutions attempted to explain their fundamental nature as sets of habitual practices, norms and laws governing economic activities. Recent developments in the analysis of transaction costs (e.g. Williamson *et al.*, 1975; Williamson, 1985; Hodgson, 1988; Eggertsson, 1990) have been useful in examining property rights, the relation between law and economics, the existence of institutions like firms and markets, and the variety of organisational forms. Changes in transaction costs can help explain changes in the boundaries of the firm and in its share of market production. Transaction analysis helps understand the nature of the classic problem any firm or institution faces of whether to make or buy, but implicitly assumes that the benefits to the firm of an activity are equal whether carried on inside or outside its

boundaries. The choice of make or buy requires comparing the costs and benefits of the alternatives, but as long as transaction analysis is framed in terms of economising costs and neglects the benefits which are the primary reason for making a transaction, it is static and insufficient. The tactic assumption that behaviour is opportunistic lacks objectivity and is misleading. The implied need for management by command or inducement has been subtly reinforced by research in other social sciences into extrinsic influences on behaviour at work which place undue weight on behaviouralist psychology. Although labour is hired at current wage rates, the benefits to the firm and to the workers arise in the process of production, develop dynamically, and accrue in the future. They are influenced by the ways in which workers develop and use their abilities in the course of producing. Individual economic behaviour cannot be explained by the constructs of economic theory alone, and will remain imperfectly understood as long as the findings of theoretical and empirical work in cognitive psychology are overlooked.

This article proposes the concept of the transaction benefits of the incomplete contract of labour, derived by applying the findings of cognitive psychology to the acquisition and use of cognitive abilities during work. This has several implications for the treatment of individual human behaviour in economic analysis. It allows understanding to be distinguished from specific technology, or knowledge *qua* information. It explains why labour is heterogeneous. It permits the incomplete contract of employment to be analysed objectively. It indicates the falsity of distinctions between manual and cerebral work. It makes analysis dynamic, and suggests a basic coherence between the economics of labour and of production and growth, and between microeconomic behaviour and macroeconomic phenomena. It helps explain how the response of workers to problems, errors and dysfunctions endogenously generates improved understanding, whilst at the same time understanding remains incomplete and suboptimal performance is inevitable.

Bibliography

Alchian, A. A. and Demsetz, H. 1972. Production, information costs and economic organisation, *American Economic Review* vol. 62, 777–795

Argyris, C. 1964. *Integrating the Individual and the Organisation*, New York, Wiley

Arrow, K. J. 1962. The economic implications of learning by doing, *Review of Economic Studies*, vol. xxix (3), no. 80, 155–173

Arrow, K. J. 1974. *The Limits of Organisation*, New York, Norton

Baumol, W. J. 1986. Williamson's economic institutions of capitalism, *Rand Journal of Economics*, vol. 17, no. 2

Becker, G. 1964. *Human Capital*, New York, NBER

Blaug, M. 1983. 'Where are We Now in the Economics of Education?' Special Professorial Lecture, University of London Institute of Education

Boyer, R. (ed.) 1988. *The Search for Labour Market Flexibility*, Oxford, Clarendon Press

Brown, W. 1973. *Piece-Work Bargaining*, London, Heinemann

Brown, W. and Nolan, P. 1988. Wages and labour productivity, *British Journal of Industrial Relations*, vol. 26, no. 3

Bruner, J. S. 1966. *Towards A Theory of Instruction*, Cambridge, MA, Harvard University Press

Burchell, B., Elliot, J., Rubery, J. and Wilkinson, F. 1993. Job content from managers' and employees' perspectives, in R. Penn, M. Rose and J. Rubery (eds.), *Patterns of Skill Change*, Oxford University Press, forthcoming

Cartier, K. 1993. 'The Quality of Labour: The Economics of Discretion and Learning During Work', mimeo, Marshall Library, Cambridge

Child, J. 1969. *The Business Enterprise in Modern Industrial Society*, London, Collier-Macmillan

Coase, R. 1937. The nature of the firm, *Economica*, New Series, vol. 4, 386–405

Coase, R. 1987. 'The Nature of the Firm', Yale Conference, mimeo

Duru, P. 1986. *La recherche en sciences économiques et en gestion*, Paris, CNRS
Edwards, E. and Lees, F. P. (eds), 1974. *The Human Operator in Process Control*, London, Taylor and Francis
Eggertsson, T. 1990. *Economic Behaviour and Institutions*, Cambridge, CUP
Fama, E. F. 1980. Agency problems and the theory of the firm, *Journal of Political Economy*, vol. 88, 288–307
Fox, A. 1974. *Beyond Contract*, London, Faber and Faber
Garnier, O. 1984. 'La Théorie Néo-Classique Face au Contrat de Travail', Table Ronde, INSEE, Paris
Handy, C. B. 1985. *Understanding Organisations*, Harmondsworth, Penguin
Hart, O. and Moore, J. 1985. 'Property Rights and the Nature of the Firm', LSE Discussion Paper
Hart, O. 1987. 'Incomplete Contracts' in Eatwell, J. *et al.* (eds) *The New Palgrave*, London, Macmillan
Hart, O. and Moore, J. 1988. 'Incomplete Contracts and Renegotiation, LSE Discussion Paper
Hirschman, A. 1970. *Exit, Voice and Loyalty*, Cambridge, MA, Harvard University Press
Hobbes, T. 1969 [1651]. *Leviathan*, Menston, Scolar Press
Hodgson, G. M. 1988. *Economics and Institutions*, Cambridge, Polity Press
Jaques, E. 1956. *The Measurement of Responsibility*, London, Tavistock
Jaques, E. 1976. *A General Theory of Bureaucracy*, Oxford, Heinemann
Kaldor, N. 1960. *Essays on Value and Distribution*, London, Duckworth
Kaldor, N. 1972. The irrelevance of equilibrium economics, *Economic Journal*, vol. 86
Kanter, R. M. 1983. *The Change Masters*, London, George Allen and Unwin
Knight, F. H. 1926. *Risk, Uncertainly and Profit*, LSE reissue, 1964
Kuhn, T. S. 1968. *The Structure of Scientific Revolutions*, Chicago, IL, University of Chicago Press
Lakatos, I. and Musgrave, A. 1970. *Criticism and the Growth of Knowledge*, Cambridge, CUP
Leibenstein, H. 1966. Allocative efficiency vs. "X"-efficiency, *American Economic Review*, vol. 56, June
Leibenstein, H. 1975. Aspects of the X-efficiency theory of the firm, *Bell Journal of Economics*, vol. 6, no. 2
Leibenstein, H. 1982. The Prisoners' Dilemma in the Invisible Hand, *American Economic Review*, Papers and Proceedings, vol. 72, no. 2
Lindbeck, A. and Snower, D. 1985. Explanations of unemployment, *Oxford Review of Economic Policy*, vol. 1, no. 2
Locke, J. 1967, [1689]. *Two Treatises on Government*, Cambridge, CUP
Malcomson, J. M. 1984. Efficient labour organisation, incentives, power and transaction costs, in Stephen, F. H. (ed.), 1984
Mallet, L. 1980. *Le Marché Local du Travail*, Toulouse, CNRS
Marsden, D. 1986. *The End of Economic Man?*, Brighton, Wheatsheaf
Marshall, A. 1920. *Principles of Economics*, 8th ed., London, Macmillan
Marx, K. 1865. *Wages, Prices and Value Added, collected Works of Karl Marx and Frederick Engels*, Vol. 16, London, Lawrence and Wishart
Moray, N. (ed.) 1979. *Mental Workload: its Theory and Management*, New York, Plenum Press
North, D. C. 1984. Transactions costs, institutions and economic history, *Zeitschrift für die Gesamte Staatswissenschaft*, no. 140, 7–17
Okun, A. 1981. *Prices and Quantities*, Oxford, Blackwell
Oswald, A. and Turnbull, P. 1985. Pay and employment determination in Britain, *Oxford Review of Economic Policy*, vol. 1, no. 2
Peter, L. J. and Hull, R. 1969. *The Peter Principle*, Morrow, Bantam Books
Paiget, J. 1969. *Psychologie et Pédagogie*, Paris, Denoel
Polanyi, M. 1967. *The Tacit Dimension*, Routledge and Kegan Paul
Pollert, A. (ed.) 1991. *Farewell to Flexibility*, Oxford, Basil Blackwell
Pugh, D. S. (ed.), 1971. *Organisation Theory*, Harmondsworth, Penguin Books
Rivero, J. and Savatier, J. 1981. *Droit du Travail*, Paris, Presses Universitaires de France
Rose, M. 1988. *Industrial Behaviour*, 2nd edn., Harmondworth, Penguin Books
Rosen, S. 1972. Learning by experience as joint production, *Quarterly Journal of Economics*, vol. LXXXVI
Rosen, S. 1985. *Implicit contracts: a survey, Journal of Economic Literature*, vol. 23, no. 2

Rousseau, J-J. 1762. *Treatise du Contrat Social*, Paris, Garnier
Rubery, J. C and Wilkinson, F. 1981. Outwork and segmented labour markets, in Wilkinson (ed.) *The Dynamics of Labour Market Segmentation*, London, Academic Press
Simon, H. A. 1962. New Developments in the theory of the firm, *American Economic Review, Papers and Proceedings*, vol. 52
Simon, H. A. 1965. *The Shape of Information for Men and Management*, New York, Harper and Row
Smith, A. 1956 (1763). *Lectures on Justice, Police, Revenue and Arms*, E. Cannan (ed.), New York, NY, Kelley and Millman
Smith, A. 1950 (1776). *An Inquiry into the Nature and the Causes of the Wealth of Nations*, 6th edn., London, Methuen
Stephen, F. H. (ed.) 1984. *Firms, Organisation and Labour*, London, Macmillan
Stiglitz, J. E. 1975. Incentives, risk and information: notes towards a theory of hierarchy, *Bell Journal of Economics*, vol. 6, no. 2
Tanguy, L. (ed.) 1986. *L'introuvable relation formation/emploi*, Paris, La Documentation Française
Thévenot, L. 1985. *Les Investissements de Forme*, Cahiers du Centre d'Etudes d'Emploi, Paris, PUF
Verdier, E. (ed.) 1984. *Changements techniques et qualification: vers une nouvelle productivité?*, Paris, La Documentation Française
Vroom, V. H. and Deci, E. L. (eds) 1970. *Management and Motivation*, Harmondsworth, Penguin Books
Whitehead, J. 1976. *Personality and Learning*, London, Hodder and Stoughton, Open University Press
Whyte, W. H. 1955. *Money and Motivation: A Study of Incentives in Industry*, New York, Harper and Row
Wilkinson, F. 1983. Productive systems, *Cambridge Journal of Economics*, vol. 7, no. 4, December
Williamson, O. E., Wachter, M. L. and Harris, J. E. 1975. Understanding the employment relation, *Bell Journal of Economics*, vol. 6, no. 1
Williamson, O. E. 1984. Efficient labour organisation, in Stephen (ed.) 1984
Williamson, O. E. 1985. *The Economic Institutions of Capitalism*, New York, Free Press
Young, A. 1928. Increasing returns and economic progress, *Economic Journal*, vol. 28