

## **An Entrepreneurial Theory of the Firm**

**Mark Casson**

Key words: FIRM, INFORMATION, COORDINATION, INSTITUTION.

Address for correspondence:

Mark Casson  
Department of Economics  
University of Reading  
P.O.Box 218  
Reading RG6 6AA, UK.

Telephone: (0) (44) 118 9318227  
Fax: (0) (44) 118 9750236  
E-mail: [m.c.casson@reading.ac.uk](mailto:m.c.casson@reading.ac.uk)

March 30 1998.

## **Abstract**

The most effective way to synthesise the insights of transactions cost-based and resource-based theories of the firm is to build the theory of the firm around the personality of the entrepreneur. The theory of entrepreneurship has a long tradition, going back to Cantillon and Say, but its contribution to the theory of the firm is often ignored.

An entrepreneur is someone who specialises in identifying new opportunities for coordination. The entrepreneur's personal comparative advantage lies in processing information, and particularly in synthesising information. This requires considerable judgement, because the information collected is sometimes conflicting, and often incomplete.

An entrepreneurial theory of the firm can encompass all the the major issues in current debate on the nature of the firm. The synthesising skills of the entrepreneur are closely linked to the core competencies of the firm. The appropriation of rents from entrepreneurial innovation raises important transaction cost issues, and the way these issues are resolved determines where the boundaries of the firm are drawn. The synthesis of information can be effected in different ways, and different types of synthesis lead to different forms of corporate evolution: organic growth, merger and acquisition, diversification, joint ventures, and so on.

## 1. Introduction

Four main issues are addressed by the modern theory of the firm; they concern

- the boundaries of the firm;
- the internal organisation of the firm;
- the formation, growth and diversification of the firm; and
- the role of the entrepreneur.

These issues are all related, but the relationships between them are not always clear. This paper argues that the the fourth issue - the role of the entrepreneur - is the most fundamental. The theory of the entrepreneur has a distinguished pedigree in the work of Cantillon (1755), Say (1803) and Knight (1921), but its contribution to the theory of the firm has often been ignored. This is unfortunate, because theories that neglect the entrepreneurial dimension can offer only a partial explanation of the behaviour of the firm. Apparent conflicts between some of the more recent theories of the firm - such as transaction cost theories (Williamson, 1985) and resource-based theories (Foss, 1997) - can be traced to a common failure to take full account of the entrepreneur. In providing a modern exposition of the entrepreneurial theory of the firm, therefore, this paper also contributes to the debate over the relative merits of other theories.

Why has the entrepreneurial approach been so neglected? There are three main reasons. The first is the *idiosyncrasy of some influential ideas about the motivation of the entrepreneur*. Both Schumpeter (1934) and Kirzner (1973, 1997) have stressed the distinctive psychology of the entrepreneur, in a manner that suggests a conflict with the mainstream 'rational action' approach to economic behaviour. Most psychological insights into entrepreneurship can be accommodated, however, by a simple modification of the conventional rational action approach. It is sufficient to take account of (i) emotional rewards as well as material rewards and (ii) the costs of information incurred in taking a decision. The limitations of conventional theory are largely a consequence of underestimating the size and significance of emotional rewards and information costs (Casson, 1995). It is possible to take account of the distinctive psychology of the entrepreneur by postulating a strong preference for certain types of emotional reward and an unusually low level of personal information cost. In this way the ideas of Schumpeter and Kirzner can be reconciled with the rational action approach.

Secondly, *most theories of the firm take the existence of markets as given*. But it is more realistic, in many cases, to take the existence of firms as given, and to explain the emergence of markets as a consequence of the entrepreneurial activities of firms. Consumer markets, in the

way that we know them today, would not exist except for innovations by firms. Firms introduce new products and act as hubs of information about them: making contact with customers, quoting prices, and adjusting these prices from time to time in response to change. Firms are not just producers who passively respond to prices set by a Walrasian auctioneer; they are actively involved in creating markets and setting prices in them. Firms act as market-making intermediators, continuously monitoring potential sources of change in demand and supply and adjusting price and output in the light of the information they acquire (Casson, 1997).

In his original discussion of the nature of the firm, Coase (1937) took a holistic view of the economy, which focused on the problem of organising the economic system to maximise social welfare. This concern reflected the debate over planning *versus* prices at the time he wrote. But this approach was soon overshadowed by refinements of Chamberlin's (1933) monopolistic competition theory of the firm, which adopted a more partial equilibrium view. Entrepreneurship can only be fully understood within a holistic approach, and so *neglect of the holistic approach* also contributed to the neglect of the entrepreneur. It is mainly Richardson (1960) and Loasby (1976, 1991) who have kept alive the holistic view, even though neither is a disciple of Coase, and the present paper draws heavily on their work.

The holistic approach has the great advantage of highlighting the role of information, because information processing is so central to the planning problem. Thus it is true to say not only that firms emerge as 'islands of conscious power' within a market system, but also that markets emerge as institutionalised sources of flexibility within a planning system. From the latter point of view, firms represent the legacy of the planning system after power has been decentralised through markets.

The role of information is central to this paper. Information impinges on the design of property rights, because the heterogeneity of property rights makes their codification very costly. The subjectivity of information explains the diversity of opinion between entrepreneurs and hence the nature of competition between them (Hayek, 1937). The costs of gathering information, and the need for suitable incentives to collect it, explain a great deal about the way that entrepreneurs are rewarded.

Once the significance of information is properly appreciated, the entrepreneurial firm emerges quite naturally as an integral part of the market system. Firms and markets are not only substitutes but complements: markets are difficult to organise without firms, and firms

are difficult to operate without markets. The entrepreneurial theory of the firm portrays the firm as an organisation dedicated to the planning of markets. The firm's planning is based upon a synthesis of information. An initial synthesis of information improvised by its entrepreneurial founder is followed by recurrent syntheses effected more routinely by its managerial organisation. This organisation is in turn embedded in an institution framework which defines and secures property rights. This vision accords an even-handed emphasis to planning and markets, and synthesises complementary insights derived from both these approaches to the economy.

The paper is organised into three parts. Part I (sections 2-4) introduces the basic concepts of information, coordination and institution. Part II (sections 5-9) examines the role of information synthesis in effecting coordination within a volatile environment. The advantages of a flexible division of labour in information processing are set out. It is shown how the institutions of a market economy provide flexibility of the kind required. This discussion exhibits the relationship between the firm and the market in a new light. Part III (sections 10-17) applies the general principles of Parts I and II to a range of special issues in the theory of the firm: organisational structure, the employment relation, the financing of growth, and the determinants of the boundaries of the firm. It is shown that entrepreneurship holds the key to the 'core competencies' of the firm, and that the appropriation of rents from entrepreneurial synthesis is crucial to every aspect of the strategy of the firm.

## **Part I: Basic concepts.**

### **2. Information**

The conventional view of the real economy is highly materialistic, and the legacy of this view is still reflected in the theory of the firm. It emphasises the flow of tangible resources such as physical products, rather than intangible resources such as information. It fails to emphasise sufficiently that the coordination of tangible flows require major flows of information to support it.

A substantial proportion of the output of an advanced economy is intangible, and most of this is accounted for by information in its various forms (Eliasson, Folster, Lindberg, Pousette and Taymaz, 1990). Some information is consumed directly - as entertainment, for example - but most of the information used in business has purely instrumental value. It is not a final product, but an intermediate product - for example, technological know-how

**Table 1**

**Key characteristics of information**

---

*Public good.* Access to information can be extended without limit (though communication costs will be incurred in the process - see below).

*Discretionary exclusion.* Unlike some public goods, such as defence, information does not have to be shared with everyone. Thus information can, in principle, be privately appropriated.

*Property rights are difficult to enforce.* Appropriation is difficult because legal rights to the exploitation of information (e.g. patents) are difficult to enforce. This means that markets in information are difficult to operate. Indeed, the concept of a market in information seems strange to many people precisely because so few markets of this type actually exist.

*Quality assurance problems are acute.* Information can be kept secret, but selling secrets is difficult because their quality is difficult to for the purchaser to assess. Information can be either true or false. Corroborative evidence is often difficult to obtain because such evidence would give the seller's secret away.

*Information is very heterogeneous.* A vast amount of information is required to specify all the spatial, temporal and personal details of an economic situation. If every item of information had a separate market then the number of markets would be enormous.

*Communication is costly.* Communication costs increase with geographical distance, psychic distance and the speed required. Some modes have high fixed costs and low variable costs (e.g. broadcasting) while others have low fixed costs and high variable costs (e.g. face-to-face conversation).

*Access to information is unequal.* Primary sources of information are widely dispersed, and people cannot be in several places at the same time. Communication costs impede the dissemination of information from one place to another. Thus everyone tends to have a unique set of information at their disposal.

*Information is produced by observation and interpretation,* rather than by the physical transformation of raw materials. Since the costs of discovery cannot be recovered from direct sales, they have to be appropriated in other ways instead. This appropriation problem is a crucial building block for a comprehensive theory of the firm.

---

generated by R&D. It is often supposed that information is ultimately factual, but this is not so. Theories as well as facts are informative, for without a theory to interpret it a fact has no meaning.

The principal instrumental use of information is for coordination. A high proportion of clerical and administrative workers are involved, directly or indirectly, in coordination. This includes professionals working in the service sector, and managers in manufacturing firms responsible for marketing, purchasing, personnel and accounting. Tangible assets are involved in coordination as well. Offices are used to accommodate clerical workers, paper-based filing systems aid memory, while computers support not only memory, but communication and decision-making too. Industries such as banking, advertising, and travel agency are almost entirely devoted to coordination of one kind or another.

Information resembles an ordinary product in some ways, but is crucially different in others, as indicated in Table 1. Like ordinary products, it flows across space and time along channels created for this purpose. These channels have a network structure. Like an ordinary durable good, information is both a stock (when stored in memory) and a flow (when communicated as news to other people).

The role of information is downgraded by conventional economic theory in numerous ways. The legal framework is taken as given, even though the availability of codified laws and registered titles to property depends on the level of information costs (Clanchy, 1993). It is well known that the expedient of replacing the market process with the Walrasian auctioneer avoids confrontation with the information problems of the market. Information-intensive activities within the firm are bundled into black boxes labelled 'management' or 'marketing', and then handed over to other disciplines for detailed study (Hollander and Rassuli, 1993). Personal skills in handling information are lumped together under the heading 'human capital' (even though this was not the original intention of Schultz, 1971). Important spatial interactions based on information flow are dubbed 'agglomeration economies' and treated as indistinguishable from any other source of economies of scale (despite the important differences mentioned by Marshall, 1919). Product flows of a variable and intermittent nature are treated as continuous and steady, and their leads and lags ignored, so that information regarding the timing of flows can be conveniently suppressed (Winston, 1981).

All these 'tricks of the trade' serve to eliminate the heterogeneity of real world

economies from economic models. In abstracting from complexity, economic models systematically abstract from information too. Models are distorted because they understate the demand for information. Exclusion of entrepreneurship is only one of several consequences of this. Weak microeconomic foundations of monetary theory are another. The homogeneity of money, which is the foundation of money demand, cannot be contrasted with the heterogeneity of other products when these other products are wrongly described as homogeneous too. Similarly, the role of literacy and numeracy has been understated in economic models of development, because the need to record and analyse information in order to effect coordination has not been taken into account. A related point concerns the concept of private property that underpins the market system itself. Markets involve the exchange of heterogeneous property rights, and the relative costs of handling information on ownership rights of different kinds affects the sort of ownership rights that are recognised by law. In so far as the firm is a response to market failure, and market failures are in turn a consequence of legal failures, the costs of recording legal rights of different kinds impinge significantly on the size and scope of firms.

### **3. Coordination**

Coordination may be defined as an improvement in the allocation of resources. It is assumed that coordination is normally effected by a specialist coordinator. Under certain conditions this specialist coordinator can be identified as an entrepreneur (Casson, 1982). Coordination can be interpreted from either a private or a social point of view. From a social perspective coordination may be defined as a change in allocation which makes at least one person better off without making anyone else worse off. From a private point of view, it is a change that that makes the coordinator themselves better off. In so far as the coordinator is bound to compensate others in order to implement his plan, coordination succeeds when the benefit to the coordinator exceeds the compensation paid. Private and social coordination diverge when the coordinator does not need to compensate some of the people who are harmed, or when the coordinator cannot obtain compensation from some of the people who gain.

If no attempt were made at coordination then the allocation of resources would remain unchanged. Coordination is therefore a 'pro-active' form of decision-making in which the decision-maker rejects the passive option of maintaining the *status quo*. A change in the

use of any one resource will normally require an accommodating change in the use of some other resource. Thus if one resource produces less of one output and more of another, then some other resource will have to substitute in favour of the more abundant resource and against the scarcer one. On this view, coordination consists of inter-related substitutions.

A market is a web of potential linkages between alternative sources of supply and demand. Because of transport costs and transaction costs, these alternatives are not perfect substitutes for each other. Coordination alters the volume of trade along each linkage. By reducing some volumes to zero, and increasing others from zero, it replaces one set of linkages with another. This means that the inter-related substitutions consists of pairs of individuals switching from one linkage to another.

Coordination can raise and lower the volume of trade in a quite dramatic way. New markets are brought into existence by raising volumes to positive levels for the first time, whilst existing markets are eliminated by reducing volumes to zero. Coordination does not therefore merely adjust the volume of trade within a given set of markets; it brings new markets to life through innovation, and eliminates other markets through obsolescence.

#### **4. Institutional embeddedness: law, morality and the state**

A coordinator cannot implement a plan without possessing influence or control over resources. A coordinator needs to know who controls the relevant resources, so that they can either extract a commitment from them to use their resources in a particular way, or acquire this right of control for themselves. Coordination cannot function effectively where property rights are ill-defined, or when it is difficult to find out who possesses them.

Property rights are upheld by institutions. The firm is an institution, and so too is the state. The policies of the state determine the legal forms that firms can take.

According to traditional ‘social contract’ theories of the state, natural rights are ultimately vested in individuals. Because selfish individuals may claim a lot of rights for themselves whilst denying any obligation to respect the claims of others, the state emerges as the final arbiter of rights. To avoid the threat of anarchy, natural rights are converted into formal legal ones (Buchanan, 1978). The state employs its monopoly of legitimate force to uphold these legal rights. This ‘individualistic’ view of society suggests that the state, like all other institutions, is the product of voluntary association between its members. It also suggests that individuals are free to exercise their legal rights for their own benefit, rather

than just for the benefit of the state. The state therefore accepts the pursuit of self-interest as legitimate, within certain bounds.

When these legal rights are freely alienable, people will voluntarily reassign them if they are suitably compensated. The state employs its monopoly of coercion to enforce contracts of this kind. State enforcement of individual rights is costly, however, because it requires the collection and weighing of a large amount of evidence (North, 1981). Self-interested individuals will exploit the limitations of enforcement to infringe each other's rights covertly. This can be solved by requiring people to make an emotional commitment a moral code. An ethic of honesty is established - often backed up by religious belief - and offenders punish themselves with guilt.

An important difference between material sanctions and emotional ones is that material sanctions involve external monitoring of infringements, whereas emotional sanctions depend on people monitoring themselves. Self-monitoring economies on information costs, but requires investment in moral rhetoric to support it (Casson, 1991).

Morals needs to be standardised, so that everyone's behaviour becomes predictable. Morals also need to be legitimated, because otherwise the intensity of emotion will be too low to outweigh the material incentive to cheat. By promoting communication between its members, the state is well-equipped to disseminate a standardised morality as an internal public good.

Individuals can reassign their rights, not only to other individuals, but to other institutions too. Indeed, in an individualistic society people are free to form institutions on their own initiative. Most institutions are corporate - they have several members, although some are singular - such as an one-person firm. The associating parties may include other institutions. The private firm - such as a partnership or a joint stock company - exemplifies a corporate institution comprised of individuals, while an industry trade association exemplifies a corporate institution comprised of other corporate institutions.

The function of an institution is to coordinate the exercise of the rights that are vested in it by constitutional means. The constitution of an institution prescribes where ultimate authority lies - namely with the principal stakeholders, who have a right to participate in all key decisions. Stakeholders have normally contributed some of their rights to the institution. These rights may provide the institution with control over physical or financial resources, or access to valuable information. In some cases (particularly where the state is involved)

special permission for the activity may have been given as well.

The constitution also specifies the intended beneficiaries. These are people who enjoy the right to consume resources provided to them by the institution. In a voluntary association all the members, including the stakeholders, expect to benefit from joining - though not necessarily in a purely selfish way. Indeed, not-for-profit institutions such as charities specifically aim to benefit others. The stakeholders of a charity benefit emotionally through empathy with the beneficiaries. Similarly, trades unions and professional associations further the collective interests of the membership rather than the personal interests of individuals. Institutions such as the family and the state are concerned with child-rearing and long-run group survival, as well as with short-run material objectives. In this respect, the private firm, with its emphasis on shareholder's profit, is the exception rather than the rule. Even here, though, the interest in profit is often diluted by concerns with dynasty, or even with the product as an end in itself.

Institutions vary in their degree of formality. A formal institution normally has a written constitution, and a list of members. An informal institution may be just a social network whose members recognise each other and trust each other.

Firms are institutions that specialise in the implementation of coordination plans. Firms are readily established whenever individual rights are freely alienable and the principal of voluntary association for private profit is recognised by the state. Indeed, the existence of tax privileges, the option to limit liability, and indefinite legal lifetime all the make the 'legal fiction' of the firm extremely attractive.

Institutions, therefore, provide the framework within which coordination is carried out. They identify unambiguously who is responsible for what decisions. They build up trust between their members so that plans can be carried out in confidence. Specialised institutions like firms are normally embedded within a framework established by a more general institution like the state (Granovetter, 1985), because this increases confidence that coordination plans will not be disrupted by unforeseen events.

## **Part II: Information synthesis**

### **5. Volatility**

Why is coordination a continuing activity, rather than just a once-for-all activity required to set the economy up? The answer lies in volatility. In the absence of volatility,

coordination would eventually shift the economy into an equilibrium state, where it would remain indefinitely. No new information would need to be collected because everything would remain exactly the same as before. By contrast, volatility brings change, which means that any allocation of resources will in time become obsolete. Attempts to maintain the *status quo* in the face of changing conditions will cause the system to degenerate into anarchy. Because coordination needs to be continuous, information flow has to be continuous as well.

Because coordination involves inter-related substitutions, a single observation is normally insufficient to identify a coordination opportunity. Several items of information are normally required, one relating to each substitution. These observations need to be synthesised in order to identify a coordination opportunity. Synthesis is more than just aggregation. To effect a meaningful synthesis, some kind of theory or mental model of the environment is required. Without some sort of theory it is difficult to decide what sort of observations need to be made. Since theories require imagination, however, there may sometimes be no theory available for this purpose. In other cases, theories may proliferate. On many topics, the set of historical and experimental observations so far accumulated is insufficient to determine which of several theories is correct. Indeed, the less factual evidence is available, the greater is the scope for proliferation of theory.

Observations are costly. Some observation that could be made may not be made simply because the cost is too high. This has two effects. The first is that some coordination plans may be developed on the basis of very limited information. The benefits of wider synthesis are traded off against savings in observation costs. The second is that observations become less frequent, so that the information used in coordination becomes, on average, more out of date. In addition, the desire to economise on costs of synthesis means that coordination plans are produced less frequently too. Less attempt is made to change the *status quo*.

The information used in coordination is often highly subjective. It is not only theories that can be in error, but observations too. What is usually observed is a symptom of an underlying situation which cannot itself be directly observed. While symptoms are correlated with the underlying situation, the correlation is not perfect, and mistakes can therefore be made. Moreover, since no one can observe the true situation, but only different symptoms of it, disagreements over interpretation of an observation are difficult to resolve.

Not all the errors that are made are unavoidable, though. Some may be the result of the incompetence of the observer. Others may be a consequence of deliberate negligence. These

problems can be tackled through appropriate selection and motivation, but only partial solutions can normally be found.

An individual coordinator can obtain information not only at first-hand, by direct observation, but also at second-hand, through communication with other people. Communication can result in delays and in misunderstanding. It also provides opportunities for deceit. Embedding communication within an institutional framework helps to speed up communication, reduce misunderstandings, and provides moral checks on deceit.

Errors in observation, communication and synthesis show that a coordination plan can be far from perfect. A mistaken coordination plan can make matters worse. In a volatile economy there are thus two main reasons for coordination failure. One is that new coordination plans are developed only intermittently, so that the *status quo* is preserved after changes have occurred. The other is that new coordination plans are imperfect, and may even be so misconceived as to make the situation worse. The slower is the development of plans, and the greater the risk of error in them, the lower is the degree of coordination that prevails, on average, at any given time, and more inefficiently the economy performs as a result.

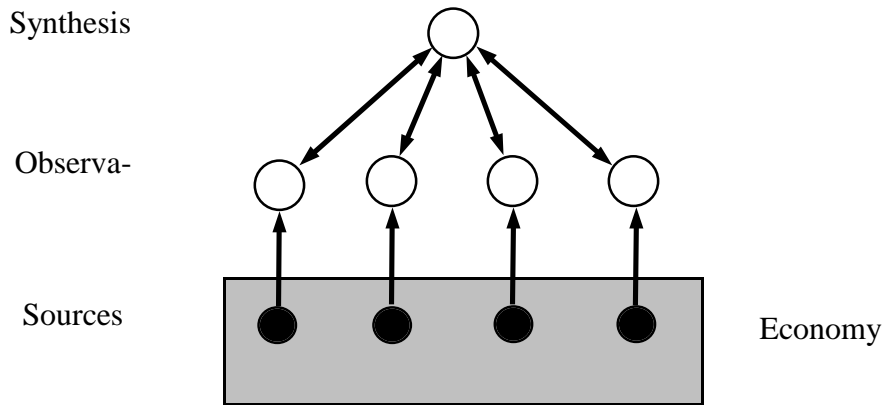
## **6. Observation and synthesis: a flexible division of labour applied to coordination.**

The trade-off between information cost and coordination error can be improved by implementing a division of labour in which people specialise in those aspects of information processing which they do best. Thus some people are more accurate observers than others, whilst synthesising ability is often quite scarce. Specialisation incurs additional communication costs, however. Figure 1 illustrates a combination of a vertical specialisation between observation and synthesis and horizontal specialisation between different observers. Each observer monitors a different source of information. The sources are embedded in the economy, and are differentiated by location, subject matter, and so on. There is two-way communication between the observers and the synthesiser - the observers report their findings to the synthesiser and the synthesiser passes down instructions about how to implement his coordination plan (see the discussion of authority relations in section 10 below).

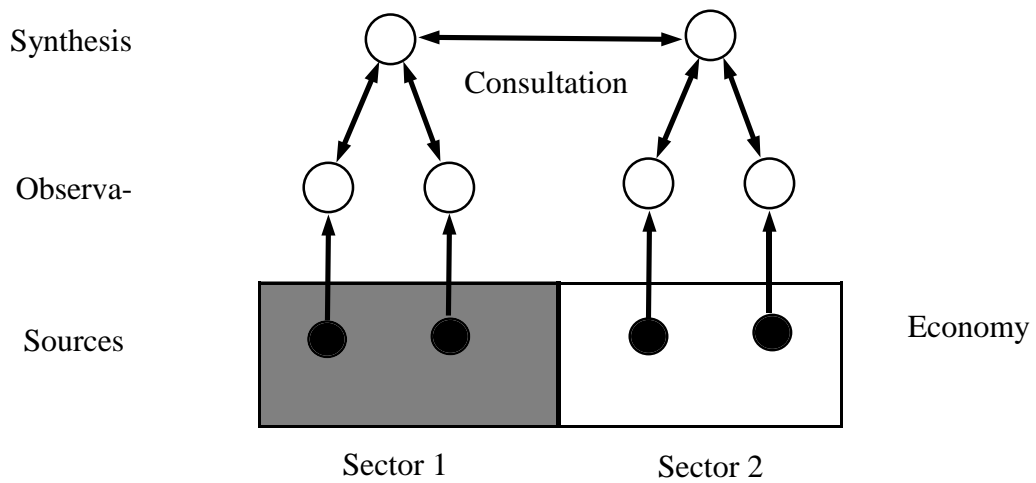
The centralisation of decision-making on a single synthesiser creates problems, however. Communication costs are very high, and the synthesiser is likely to be 'overloaded' by a mass of detail, and unable to filter out the really important information.

A division of labour can also be implemented between synthesisers. This can be

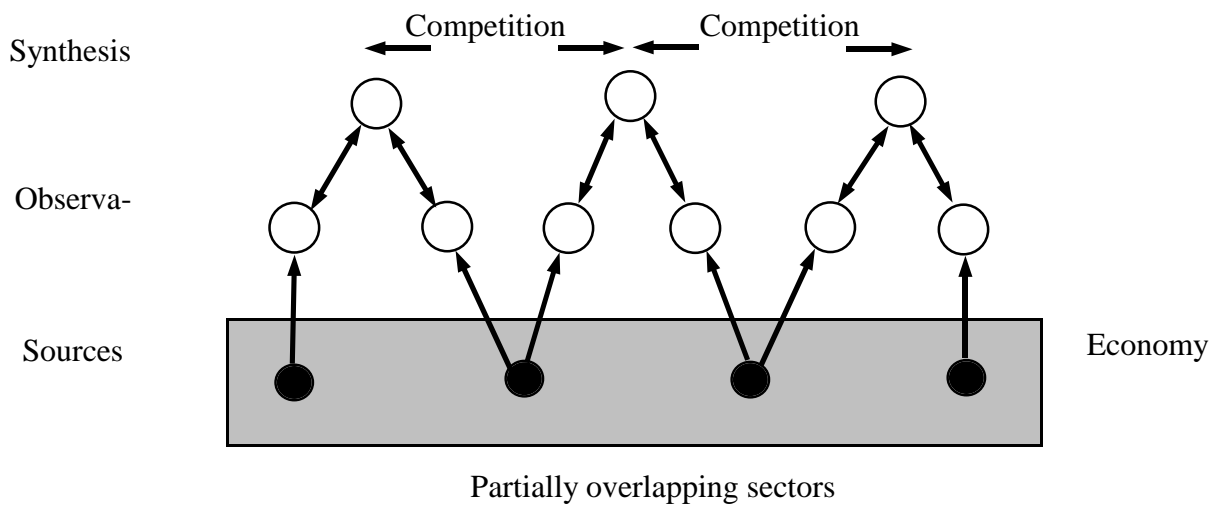
**Figure 1**  
**Centralised synthesis with vertical division of labour between observation and synthesis**  
**and horizontal division of labour in observation**



**Figure 2**  
**Rigid sectoral division of labour in synthesis**



**Figure 3**  
**Flexible decentralisation**



effected in either a rigid or flexible way. A rigid system of specialisation is illustrated in Figure 2. Each synthesiser is given responsibility for a particular sector of the economy. On this basis, each monitors separate sources of information. Each synthesiser uses local information sources, and so avoids the use of the distant sources which a single synthesiser must employ.

Decentralisation creates its own set of problems. The different sectors of the economy are not totally isolated from each other, but are linked by product flows, so that each sectoral coordination plan has knock-on effects on the other. One solution is to buffer the different sectors against each other by investing in inventories and idle capacity. Buffers do not separate different sectors on a permanent basis, however; they only 'buy time' before the 'knock on' effects of a change in any given sector are felt elsewhere. The alternative solution is for the synthesisers to share provisional plans with one another, as indicated in the figure. These plans are reconciled through a process of revision. However, the plans communicated between the synthesisers are complex, and so communication costs may quickly escalate as the revision process iterates.

While a combination of buffering and revision goes a long way to eliminating inconsistencies between different partial plans, it is inefficient to try to eliminate them altogether. Inconsistencies will remain, which create future opportunities for coordination. The process of coordination is therefore myopic, but in the long run it is myopic to an optimal degree. It trades off the risk of coordination failure against inventory costs, capacity costs and the costs of communication.

Partial planning does not have to be implemented rigidly on a sectoral basis, however. Partial plans can be allowed to overlap. Indeed, rival plans can be produced for the same sector. The obvious weakness is that a large amount of duplication of synthesis is tolerated. A synthesis is a bundle of information, and information, as noted in Table 1, is a public good. A single synthesis can therefore be shared without limit. This suggests that replication of synthesis is wasteful. Replication of synthesis would seem to be warranted only when distance-related communication costs are high, and even then only if the replicators are suitably dispersed.

This objection loses much of its force, however, when synthesis is liable to error. Not every synthesiser has access to the same information, and so each synthesis is likely to be different. Even where the information is the same, the interpretation of it may be different because different synthesisers use different theories and no one knows which theory is right.

This is the situation illustrated in Figure 3. Each synthesiser uses a distinctive information set, which is incomplete. He must make a judgement about what the missing information would be like. Different synthesisers formulate different plans because they use different information, and because they hold different theories about what the information they are missing is likely to be. The coordination plan that they devise will normally apply to some subset of the economy. The same synthesiser may develop plans relating to different sectors, or different groups of activities, on different occasions, depending upon what his information sources suggest to him.

## **7. Recruitment and motivation of synthesisers**

Synthesis requires considerable time and effort, but the potential rewards - both social and private - are high. Attracting the right sort of person to synthesis is crucial for the performance of the economy.

Under rigid decentralisation, each synthesiser is assigned a monopoly of coordination within a certain sector by the state authorities. It is undesirable to allow him to appropriate large monopoly rents for himself. He is therefore likely to be paid a fixed salary. Potential candidates compete for these posts. The higher the salary, the larger the field is likely to be. The number of competent applicants is likely to be higher as well, but screening them out from the others is difficult because the relevant qualities are often hidden. It requires considerable judgement to pick the best applicant.

Under flexible decentralisation, synthesisers are not selected by other people and appointed to predetermined roles. Rather, synthesisers volunteer for their role. Indeed, they are free to set up and define their role for themselves. Each synthesiser can back their own judgement if they wish. Each individual has an initial endowment of resources, and if these are not sufficient then they can borrow the use of more resources from other people. They exchange these resources for the resources required to implement their coordination plan.

If different synthesisers plan to use the same resources in different ways then they must compete. Prices will be bid up until all but one of the synthesisers drops out because their plan is no longer profitable. The maximum reward that can be appropriated by the successful coordinator is equal to the excess of the value of his plan over the value of the next best plan being promoted at the same time. (This refers, of course, to the value of the plan as anticipated by the coordinator, which may be different from what the value actually turns out to be.) If no new plans are being promoted, then this is equal to the excess of the value of the plan over the

value of the *status quo*. The profit accruing to the successful synthesiser provide a natural method of reward. This is the basis if the free enterprise system.

Flexible decentralisation works better than rigid decentralisation when there is a high probability that any given coordination plan will be mistaken. When there is general recognition that mistakes are likely then it is best to have several alternatives to choose from. The entrepreneurs who promote rival plans each believe that they are right and the others are wrong. Other things being equal, it is the most optimistic and self-confident entrepreneur who will prevail. This entrepreneur takes a more favourable view of the economic environment than others, and is therefore prepared to pay more for the resources for which they are competing. Being self-confident, the entrepreneur does not discount these valuations for the possibility that he may be wrong. Because his firm actually pays for the resources it uses, the successful entrepreneur insures the less optimistic against his own misjudgement, for if he is wrong then his firm bears the loss. Competition therefore forces entrepreneurs to gamble on their own optimistic judgements by insuring others against the possibility that they are wrong.

#### **8. Prospecting for opportunities under flexible decentralisation.**

To avoid wasteful duplication of discovery, and losing out through competition, different synthesisers need to focus their search for coordination opportunities in different regions of 'information space' (Boisot, 1995). Unfortunately, however, this space is not 'mapped out' in the same way as geographical space. This means that agreements to divide up the space between different searchers are difficult to specify (let alone enforce - see the discussion of patents above). The best that can be expected is for individuals to recognise their neighbours and 'keep their distance' from them. If opportunities are equally likely to be discovered anywhere in information space then this generates a tendency to an equilibrium in which searchers are uniformly distributed over this space. In practice, there are likely to be widely-held beliefs that potential discoveries are particularly dense in certain areas. Beliefs of this kind, which are common to most members of a group, are part of its culture. The most successful searchers are likely to be competent 'deviants' who correctly recognise mistakes in these beliefs, and search in other areas instead. The majority believe that the deviants know less than they do about where discoveries are likely to be made, and this acts as a natural 'barrier to entry' into the deviants' fields of search. In fact, the competent deviants they know

more than the average person does. Once the deviants succeed, and their mistake has become obvious to the average people, the deviants have more difficulty maintaining their lead because they have become leaders of the new majority, and face competition from others as a result.

The threat of competition intensifies by the fact that the field is likely to be too crowded in the first place. This is a special case of the ‘tragedy of the commons’ applied to undiscovered opportunities as an unappropriated resource. The expected number of opportunities discovered in any period increases with the number of prospectors, but at a diminishing rate. Marginal returns to the number of prospectors eventually falls to zero. But average returns remain positive at this point. When a new entrant goes ‘fishing’ for a synthesis, the incumbents crowd together more closely in order to keep the same distance from the incumbent as they do from each other. Even if the new entrant adds nothing to the expected number of discoveries, therefore, he still receives a positive return. Because he is treated the same as everyone else, he receives the average return. His entry is subsidised by the redistribution of territory from incumbent searchers.

The problem is not quite so serious as this suggests because searchers differ in their comparative and absolute advantage. Many people stay out because their expected earnings are higher in other work, so the average return never falls to zero. Those who are very good at search snap up opportunities before others, and earn a reward that is well above the average. The ablest synthesisers continue to be attracted to the field, therefore, but there may be a surplus of less able synthesisers who are not very good at anything at all.

## **9. Transaction costs and market institutions**

The system of flexible decentralisation described above relies heavily on voluntary transactions. Transactions need to be easy for this system to work well. But in practice, transactions can be costly. An efficient economy minimises the transaction costs of supporting a given set of transactions. Devices for reducing transaction costs are a prominent feature of the market economy, and a significant influence on the behaviour of the firms within it. Transaction costs are of two main kinds: information cost and appropriation cost. Appropriation costs are incurred as a consequence of strategic behaviour by the other party - notably, bluffing in negotiations before the contract is made, and default in delivery after the contract has been made. Bluffing results in the other party appropriating more of the

‘gains from trade’ than would otherwise be the case. Default is a form of theft, whereby the other party fails to honour their side of the deal.

Information costs are incurred in making contact with the other party to begin with, specifying the property rights to be exchanged, carrying on the negotiations and arranging the logistics of delivery. Additional costs of investigation may be incurred in order to reduce appropriation costs. Thus alternative trading opportunities available to the other party may be investigated in order to assess whether they are bluffing, and the reputation of the other party may be investigated to discover whether they are likely to default.

Communication costs can be reduced by the coordinator acting as an information hub. Because many different transactors are involved in the implementation of a typical coordination plan, it is uneconomic to set up the plan through multilateral negotiations involving all the parties. The coordinator transacts separately with each party instead. Bilateral negotiations economise on communication costs. A firm may be established for this purpose. It becomes a a nexus of contracts. The contracts can be negotiated sequentially, to avoid an information overload for the coordinator, although this runs the risk that a deal negotiated early on may be regretted later. Use of a firm as a legal shell means that the nexus does not disintegrate when the coordinator dies.

If contracts involve barter then the coordinator will be involved in handling many different goods in order to achieve the requisite ‘double coincidence of wants’ with each party. Money is therefore useful as a medium of exchange, and the state can help by making providing a stable supply of currency in suitable denominations. The coordinator can therefore concentrate on handling only those goods in which he has an information advantage. He can also reduce information costs by using money as a unit of account as well.

The bundling of rights is another important strategy for reducing transaction costs. Because coordination consists of inter-related substitutions, a coordination plan creates a set of demands and supplies for specific services. The specificity of the services expresses the definiteness of the plan proposed. The degree of specificity is increased by the fact that most coordination plans involve a significant time dimension, so that all the services involved must be dated. Indeed, most coordination plans might be better described as ‘coordination projects’. Lags in production and distribution mean that inputs usually occur before outputs are generated, so that even the simplest plan may involve an element of inter-temporal substitution.

In principle, the natural way to set up a coordination project in a market economy is using forward contracts. A comprehensive set of forward contracts would telescope all decisions about the future into the present, binding all parties to a predetermined course of action. Because of volatility, contingent forward contracts are required - an ordinary contract would remove the opportunity to respond to information received later on. When contracts are contingent, an organisation is required to monitor the environment in order to decide what to do at each stage. The decision rules embodied in the organisational procedures are, however, entirely dictated by the contractual arrangements made at the time the organisation was set up. Thus although decision-making is continuous, contracting is not. Contracting is a once-for-all affair. Each transaction consists of two separate parts: the negotiation of the terms at the outset, and its implementation and enforcement later on.

The problem with this approach is that the transaction costs are prohibitive. Each contract must specify the service, the time of its delivery, and the circumstances under which delivery has to be made. The further forward a contract runs, the greater its memory cost. Memory is more than storage; it involves making records, storing them in the proper order and retrieving them at the appropriate time. It is only memory that links the negotiation of the contract to its subsequent enforcement, and there is scope for disagreement if the memory of one or other of the parties is flawed. Because of these problems, the state is reluctant to invest in enforcement mechanisms for contracts of this kind. This means that for individual transactors, the risk of default is increased.

To tackle this problem, the element of specificity must be reduced. One approach is to replace rights to specific services with general rights to goods instead. The simplest general right to specify is one that includes all possible specific rights. Instead of being specified by inclusion - listing all the specific rights involved - it is specified by exclusion - namely by what is left out - which is nothing at all. This is the classic form in which the 'ownership' of a good is expressed. Replacing specific rights with comprehensive general rights like 'ownership' effects enormous economies in communication costs.

Since a single general right replaces an entire set of specific rights, a small number of markets in general rights can replace an enormous number of markets in specific rights. Many transactions costs are fixed costs, specific to the market and independent of the number of trades. For this reason, reducing the number of markets effects an additional economy in transactions costs.

When ownership is traded, the allocation of resources takes place in two stages rather than one. To acquire a service, an individual first acquires the good and then allocates it to the appropriate use. The seller does not therefore have to be told the specific use to which the good is to be put. Indeed no one else is told at all - this is entirely private information. This arrangement incurs no penalty so long as no one else cares how the good is used. But if others do care, because they experience negative 'externalities' such as pollution, for example, then it may be more efficient to revert to trading specific rights instead. The benefit of the ownership system is greatest when the use of a resource affects only the owner himself. The ownership system is therefore particularly suitable for mobile resources that are utilised inside buildings, because the owner can carry the goods off and use them in private, without disturbing anyone else. It is less suitable for immobile resources utilised out-of-doors, where others are more likely to be unavoidably affected by the way in which the goods are used.

Ownership also has implications for the time dimension. Because services are perishable, they have to be dated. Thus an intertemporal coordination plan must specify 'appointments' when deliveries of particular services are to be made. But goods are more durable than services, and it is therefore more natural to define rights over them in perpetuity. 'Ownership' of a good therefore confers a perpetual right. This right is particularly easy to enforce because it corresponds to the bundle of *de facto* rights conferred by uninterrupted physical possession of the good.

The perpetual nature of ownership means that the ownership of durable goods can be used as a substitute for forward contracts in specific services. Instead of acquiring a dated claim for the delivery of a service, an individual can purchase ownership of the good which generates the service instead. This extends the principle of multi-stage allocation of resources, explained above, from the present to the future.

Durable goods are sometimes rented out. Not all coordination plans require the perpetual use of the same assets, and it is often economic to limit commitments to only a short period of time, with an option to negotiate a renewal later on. Like ownership, rental affords a substantial degree of control. But it does not include the right to consume or sell of the good. Rented goods must be returned to the owner in the same condition they were received in, after normal allowance for wear and tear. Rental agreements are more costly to make than outright sales because of the additional information on dates that is required.

Nevertheless, rental is extremely useful for the utilisation of specialised but versatile machinery. Because it is specialised, most users require the machinery only on an intermittent basis, and not in perpetuity. Because it is versatile, many different people can use it, and each of them may wish to put it to a different use. It is uneconomic to specify every use in detail. It is therefore more efficient to rent it out than to sell the specific services it provides.

The contingent aspect of a contract also contributes materially to transaction cost. This component of cost can be reduced through the standardisation of contracts on simple contingencies that apply to everyone. By choosing contingencies that can be objectively specified and publicly observed, the costs of enforcement can be reduced. These are the principles on which the insurance industry operates.

Insurance can also be effected indirectly, by investing in the ownership of a very versatile goods instead of more inflexible ones. The more different uses there are for the good, the more likely that an adequate response can be made to unfavourable circumstances. In a market economy the ultimate versatile good is money. Money can be quickly and easily converted into any other good through trade. Holding a stock of money therefore supports a flexible multi-stage resource allocation mechanism, in which money is first converted into ownership of goods, and the goods are then used to generate the services required. The intermediating role of money in this multi-stage strategy means that access to money is crucial to the implementation of coordination plans.

### **Part III: The organisation of the entrepreneurial firm**

#### **10. Transitory and persistent change**

Some of the factors driving change are transitory and others are persistent. The role of coordination in response to transitory change is to restore the original situation as quickly as possible. Persistent changes cannot be reversed, however, and so coordination adjusts other factors to accommodate them. The factors driving transitory change tend to small in number, but they change fairly frequently. By contrast, the factors driving persistent change are more varied, but change only intermittently.

There are two main kinds of response to volatility. One is to anticipate change and plan in advance how to deal with it. This is the appropriate way of handling transitory change. The other is not to anticipate change, but simply to deal with it as and when it occurs. This is the appropriate way of handling persistent change.

Because of the small number of factors involved, the situations generated by transitory changes tend to be of a limited number of types. Because each factor changes frequently, the same situation tends to recur regularly. The relevant sources of information are fairly easy to identify in advance. It is therefore feasible to employ specialists to monitor each of these sources on a regular basis. This means that it is economic to plan in advance how to deal with these situations. The set-up costs of planning can be spread over a large number of subsequent situations. Situations are dealt with using rules which relates the action to be taken to the observations that have been made. These rules are codified, and routinely applied whenever changes occur.

There is often a division of labour between rule-making and rule-implementation. Rule-making is an entrepreneurial activity which calls for some improvisation, whereas rule-implementation is routine, and is characteristic of purely managerial and administrative work. The division of labour can be applied to the implementation of rules as well. Some people collect information, some communicate it, others store it and retrieve it and some use it to take decisions.

The rules that are used in an organisation typically have a procedural quality to them. This is because the processing of information normally benefits from being carried out in a sequential manner. A sequential process allows the processing of information at later stages to be conditioned on the outcomes of earlier stages. This applies to the processing of information on both persistent and transitory factors, but its implications are greater for transitory factors because of the recurrent nature of decision-making. The advantage of sequential information processing is that the later stages of a procedure can be modified in the light of the results obtained at earlier stages. If the early steps yield decisive results then the later steps need not be carried out. Indeed, the different stages can be sequenced to maximise the probability that a decisive result is obtained early on. This last step is a refinement of the 'optimal stopping rules' employed in the conventional theory of search (Lippman and McCall, 1979).

Procedures need to be fairly simple to reduce the cost of memorising them. However, if using a procedure 'refreshes' the memory, so that its memory cost falls the more frequently it is used, then quite complex procedures may be used to handle the more common situations (Nelson and Winter, 1982).

By contrast, the proliferation of factors, and the intermittent nature of change, means

that the situations created by persistent factors recur very infrequently indeed. For any given situation, the average lag between recurrence is very long. This makes it uneconomic to plan in advance how to deal with a persistent change. It is better to wait until the change occurs and improvise a response instead. It may not be worth remembering this response either, because it is cheaper to repeat the improvisation later. Such improvisation is characteristic of the entrepreneur.

It is difficult to identify in advance the relevant sources of information on persistent changes. People who specialise in the identification of persistent change may therefore rove around, working on the basis of hunches as to where the relevant information is likely to be found. This makes it difficult for them to employ specialists on a regular basis, although they may do in a casual way.

Because transitory situations recur regularly, it is easy to learn how to deal with them, and so over time fewer mistakes are likely to be made in responding to them. It is unnecessary to routinely seek a 'second opinion' on each situation, because the symptoms associated with each possible situation are well understood. This does not apply to persistent changes, however, which only infrequently recur. Even if they do recur, they may impinge upon a later generation that has forgotten 'the lessons of history'. Mistakes are much more likely in the response to persistent change. What is even more significant, opinions may well differ about what the appropriate response should be. Opinions may even differ about what change has occurred, whether it is persistent or transitory, or whether any change has in fact occurred at all. This means that multiple responses, based on different interpretations of events, may well be efficient. Indeed, the greater the risks involved, the more important it is to encourage diversity of response, so that the probability of experimenting with the right response is as high as possible.

A rigid division of labour is inappropriate for identifying and responding to changes of this kind. A market system which provides freedom of entry into prospecting for information tends to work better than a formal organisation which rigidly demarcates areas of search and allocates exclusive rights in each of them to particular people. Because the sources of transitory changes are more predictable, coordination of response benefits more from organisation instead. Thus persistent changes favour market coordination and transitory changes favour organisation instead.

An organisation is different from an institution. An organisation is dedicated to the processing of information through the routine application of official procedures. An

organisation must be part of an institution, because otherwise it cannot make the implementation of its decision secure. Most institutions have an organisation, because there are certain routine matters that need to be decided about the resources under its control. But institutions do not have to have organisations - thus the market is an institution, but not an organisation, in the sense defined above.

### **11. Organisational innovation in response to persistent change: the market-making firm**

A persistent change may change the pattern of transitory volatility. For example, a new cult of 'healthy living' may increase the demand for medicines on those transitory occasions when a person falls ill. A change in the pattern of transitory volatility may in turn lead to a new type of coordination opportunity occurring on a regular basis - for example, the opportunity to sell a new kind of medical drug.

Changes can be general or specific. Very general changes affect the entire economy, but most changes are specific to particular industries, particular locations, or particular people. Firms specialise in responding to product-specific changes. Some of these may be specific to customers too (where products are 'bespoke') but many - such as fashion - are common to everyone.

A particularly important type of opportunity is to establish a market for an entirely new type of good. This kind of opportunity is often associated with the foundation of a new firm. The geographical scope of the firm will depend on the nature of the persistent factor that is driving the change. It might be local, regional, or national, although if it is based on a new technology then it is likely to be global instead. The recognition of such opportunities requires a distinctive type of information synthesis. The sources of information tend to be widely dispersed, and a sophisticated mental model may be required to interpret the information. Improvisation based on outstanding judgement is often needed to effect a synthesis of this kind. By contrast, one-off opportunities can often be identified by repeatedly scanning the same small set of local information sources. Synthesis can be effected without judgement just by following simple procedures.

The establishment of a market-making firm thus typically involves two distinct types of synthesis. The first identifies a persistent change and the second involves recurrent syntheses relating to transitory changes. The impossibility of organising the first type of

synthesis makes it a highly individual task. The second type of synthesis, by contrast, normally benefits from systematic organisation.

A firm contributes to the continuing smooth operation of a market through the supply of intermediating services. These services reduce the transaction costs encounters by its customers.

- *Contact-making.* In the absence of intermediation, consumers may be unable to identify suitable sources of supply. This can be addressed by locating retail premises at a suitable central place, and by advertising the product through the media.
- *Specification.* Consumers may be unsure of how the product addresses their specific needs. Advertising entices customers to visit showrooms which display the product, and where sales people demonstrate how it performs. Where goods need to be customised, the firm can establish a protocol whereby consumer and producer can share the knowledge in order to compromise on a suitable specification.
- *Negotiation.* The intermediary can avoid protracted negotiations by publicising a non-negotiable price. The fact that he sells to many people, and that a concession to one customer would soon be discovered by others, gives credibility to this stance. Also by agglomerating together to facilitate comparison shopping, intermediators can invoke local competition to back their claim that they are setting the lowest possible price.
- *Enforcement.* the intermediary can use his reputation to assure customers about quality, and to assure producers that they will get paid.

An intermediating firm can simplify the administration and the logistics of the physical distribution of the product too. To economise on memory costs, orders may be fulfilled on a spot basis. Because consumers normally demand products in smaller batches than producers manufacture them, inventory is held for break-bulk purposes. To minimise the risk of stock-out, precautionary inventory is also held. This means that orders to replenish inventory are placed before levels get too low. To minimise transport costs, the warehouse is sited at the hub of a freight transport network. Speed of transport is determined by the trade-off between energy costs and the cost of financing stock in transit. Processing of payment is streamlined by substituting cash for credit wherever possible, though credit cards are accepted if the increased information cost is offset by the greater convenience and security afforded to the customer.

To cover his costs, the intermediary needs to extract a reward for this this service.

This can be done in two main ways. The first is to charge a fee, and the second is to buy and resell the product. The first approach may be identified with brokerage. It is most commonly used for trading in heterogeneous secondhand durable goods. Brokerage is not entirely risk-free, because most brokers operate on a no sale - no fee basis. It involves less risk than buying and reselling the product, however, because there is normally a lag between purchase and resale, during which the price may change unexpectedly.

## **12. Pyramids and authority relations**

Organisations are sometimes referred to as 'hierarchies', suggesting that their structure is like a pyramid, and that the relation between adjacent levels is one of authority and subordination. However, it can be questioned how far organisations really are like this. Recently it has been suggested that organisations can dispense with the hierarchical principle altogether, and substitute internal 'networking' instead (Hedlund, 1993).

The logic of volatility and information costs suggests that most large organisations will tend to have a pyramid form, and that many internal relationships will be based on authority, but that there is plenty of scope for variation. Unusual environments may call for unusual organisations.

Consider the pyramid principle first. Transitory changes are initially monitored at the lowest level of the organisation, and reported to the next level above where a preliminary synthesis takes place. The most commonly occurring situations are dealt with at this level using simple procedures, whilst the 'exceptions' are passed up to a still higher level. The specialists at this level know procedures which allow them to distinguish between different types of unusual situations, and to resolve each type. Extremely rare situations are referred to the top where a solution is improvised.

While this explains the emergence of a pyramid of competence, it does not directly predict a hierarchy of authority. It is also necessary to explain why the person who has the knowledge of what needs to be done tells someone else what to do. If a person is in the best position to implement his own solution then no authority relation is involved. The person who receives instructions is often the person who collected the information in the first place. This is because observation is often cheapest when it is a by-product of implementation. Thus a production worker is in the best position to observe the production conditions, and a sales person is in the best position to observe local demand. It is because the observer was

implementing the previous decision taken by the synthesiser that he was able to collect the information that has been fed into the current synthesis.

The synthesiser could, of course, simply pass back his synthesis to the observers, leaving them to work out the strategy for themselves. Because the synthesis is 'tacit', whilst instructions are explicit, however, communication costs are lower when the synthesiser acts, not as advisor, but as 'the boss'.

This explanation of hierarchy differs radically 'principal-agent' theory, which emphasises supervision instead (Milgrom and Roberts, 1992). Hierarchies exist, it is said, because superiors are needed to monitor subordinates. According to this view, the ratio between the numbers of people at successive layers of the pyramid is governed by the 'span of control'. The problem with this approach is that it does not really explain why superiors need to give orders, since it does not even consider the kind of orders that need to be given. The only information collected by superiors, according to this view, is about how their subordinates behave. The obvious objection to this view is that since the superiors do not collect or process information about the firm's environment, they cannot know what orders need to be given.

Once it has been established that a hierarchy of competence exists, however, it is easy to show that the synthesisers who take the most difficult decisions will normally take responsibility for recruiting and motivating the observers as well. The structure of supervision can only be properly understood as an adjunct to organisation of synthesis.

This account of hierarchy explains, in turn, the employment relation that is so characteristic of the firm. A firm typically 'rents' the services of its employees in the same way that it rents the services of machinery. The main difference is that employees care about the tasks to which they are assigned whereas machines do not. This applies both to information processing and to the implementation of decisions. Employees usually need to be in place before the information becomes available that determines what they have to do, so they cannot be hired to perform specific tasks. They could be compensated, using a contingent contract, for being asked to perform unpleasant tasks, but the information costs of formalising this arrangement are very high. Employees must therefore take a risk when accepting employment. This is a calculated risk, however, so long as they can estimate the probability that they will be assigned to any given task. When the organisation is driven by standard procedures which respond to volatility in a predetermined way then it is possible, in

principle, to calculate the probability that an employee in any given role will be assigned to a given task. Even if the employee cannot do this, the employer can. Provided that the employee can trust the employer to give an honest description of the job, he can make a fully rational choice of occupation. Thus the contract of employment emerges as a rational response to the problems of running an organisation. To make it work successfully, however, it is necessary for law, society and morals to encourage the employer to describe each job in an honest way.

### **13. The market for information: the appropriation problem.**

The formation of a market-making firm requires a combination of skills in identifying a market opportunity and then establishing an organisation to exploit it. People who are good at identifying changes may not be good at building an organisation, and *vice versa*. It is therefore advantageous, in principle, to establish a division of labour between the discovery of a market opportunity on the one hand and its exploitation on the other. Given the advantage of flexibility, this division of labour would be effected by a market in entrepreneurial ideas, in which opportunities for coordination discovered by one person could be sold to another for exploitation.

In practice, this arrangement is rarely observed. The explanation lies in one of the most important features of the market economy - namely that the transaction costs of trading information about coordination opportunities are exceptionally high (Casson, 1982). To begin with, the market system rewards the coordinator only if his synthesis is unique. This is related to the public good property of information emphasised in Table 1. People will not pay for information they can discover easily for themselves or obtain more cheaply elsewhere. Competition from a rival synthesiser selling the same idea will bid down the payment towards zero, because once a synthesis has been effected, the costs are sunk. Unless they collude, the synthesisers will lose out. Even if there is no current competitor, prospective buyers of the information may believe that competition will shortly emerge and so delay their purchase.

If the initial synthesiser could obtain a patent then he would be better protected against competition. But patents are difficult for the state to define and enforce. The complexity and specificity of a synthesis makes the cost of communication with prospective buyers very high. Evidence of priority may be difficult to supply to the patent authority.

Simple variations may be patented by rivals.

There is a bargaining problem too. If one of resources required to exploit the patent is monopolised, then the patent may alert the monopolist to the enhanced value of the resource he controls. This may shift the outcome of bilateral monopoly bargaining against the patentee.

An alternative is to have recourse to secrecy. Secrecy is a substitute exclusion mechanism: the ignorance of others replaces an unworkable patent right. Secrets are difficult to sell, however. When a secret is advertised for sale, the buyer needs to know what it can do for him in order to value it. But in describing what it can do, the seller may give the secret away. In any case, if the buyer does know what the secret is, he cannot know that he does not know it already. Neither can he check out whether it is true. Checking is advisable because, given that a synthesis can be very valuable, there is an incentive for those who have no information to make it up and pretend that it is true. But if the buyer knows the secret before he pays for it, then he does not need to pay in order to obtain it. He simply remembers what he has already been told.

If the coordinator cannot sell his synthesis then he has to appropriate returns in some other way. Two main alternatives are available. One is to take up a speculative position, and the other is to exploit the idea himself. A speculator's strategy is to analyse the consequences of his plan being adopted by other people. Suppose that he published his plan and in consequence a number of people competed to implement it. Are there certain resources which would increase in value? Any resource which would be used more intensively as a result of the plan, and whose supply is less than perfectly elastic, would appreciate in value. The coordinator can therefore benefit by acquiring all such resources. Once the purchases have been made, he announces his secret, and he sells out to the imitators. Because of competition between them, and his own monopoly position, the coordinator reaps all the profit. If 'short sales' are possible (because their transaction costs are not too high) then the coordinator also profit from resources that will fall in value. He can sell them forward and buy them more cheaply later. Indeed, if there is anyone whose judgement differs from his then he can profit by betting with them on those aspects of his plan where their differences lie. The transactions costs involved in this ploy are normally prohibitive, though. Indeed, the whole speculative approach, though ingenious, is rather limited in its potential. It requires an enormous outlay of funds, far in excess of what the coordinator is likely to have at his disposal (see below). It also depends on key resources being in inelastic supply, which certainly cannot be

guaranteed, especially when the scope of the plan is relatively small.

The most promising approach is the second - for the coordinator to implement the plan himself. He internalises the market for the right to exploit the opportunity. To do this effectively, though, he must pre-empt others in the field. As soon as he begins to trade, his secret may leak out. One way to avoid this problem is to tie in all the customers and suppliers using forward contracts. The difficulties of doing this have already been explained. It has also been explained that a simple alternative is to acquire the ownership of assets instead. In this sense speculation can come to his aid. If he buys up the assets required by his plan then he loses nothing if he sells their services back to himself, but he gains if he sells them to imitators that he has failed to stop.

It is interesting to compare the strategic significance of ownership and rental in this context. Physical implementation of a coordination plan normally requires only the rental of the assets required. Rental gives sufficient control to reallocate the assets to an alternative use, while the ability to renew rental agreements fairly easily provides a useful element of flexibility. By contrast, ownership of assets is most valuable as part of an appropriation strategy. Using secret information about the value of the plan to purchase undervalued assets appropriates profit in a pre-emptive way. On this view, the only link between ownership and physical operations is that ownership may sometimes be a cheaper way of obtaining control because it is more easily specified and more readily enforced.

Practical confirmation of the significance of this point comes from the way that many firms do not use the assets they own, but rent them out, often to potential rivals. Conversely, these same firms often rent assets of the same type that they own, because they are in a more convenient location for their use. A prime example concerns integration by trading companies into shipping, where the company's ships are chartered out, and the company's goods are carried on independent vessels.

There are other means of pre-emption too. The exploitation of an opportunity usually benefits from organisation, as noted above. An organisation represents an investment, in which fixed costs are incurred at the outset in order to reduce variable costs later on through the application of a routine. By sinking substantial funds in an organisation, the coordinator raises the stakes for imitators, who must make similar investments if they are to match his variable costs. If they do not match them then they can expect that if they enter he will initiate a price war to drive them out. It is important, though, to avoid the use of fixed interest bonds

to finance the investment, because otherwise he could be forced into bankruptcy by a fall in price, and potential entrants would be encouraged by this prospect.

The main problem with this second approach is that the synthesiser has to become an organiser too. He can, of course, hire an organiser, but then he has to select the right person and motivate them too. By requiring a synthesiser to be an organiser, the effective supply of synthesisers is reduced. The problem should not be exaggerated, however. In many cases the information gained through the process of synthesis is crucial to the design of an efficient organisation. This information may be tacit, in the sense that it is costly to communicate to other people. Just as before, high communication costs discourage a division of labour in information processing, and in this case encourage the synthesiser to give up his role when he makes a successful synthesis and turn his attention to organisation instead. Only when appropriate procedures have been designed, and routines are working smoothly, does he return to synthesis again.

#### **14. Internalisation**

This emphasis on the internalisation of information runs counter to much of the transaction cost literature on the theory of the firm, which focuses instead on vertical integration based on material flow (Williamson, 1975). Many transaction cost issues are, of course, common to both types of flow, and are simply more acute in the case of information. Thus the question of whether secret information is genuine or not has a parallel in the way that quality control applied to components and raw materials stimulates backward integration (Casson, 1987).

On the other hand, the concept of ‘asset specificity’ (Klein, Crawford and Alchian, 1978; Williamson, 1985) translates rather poorly from the material domain to the information domain. At an intuitive level, there is nothing specific about the use of a public good like information - particularly information about a product innovation opportunity that can be exploited in any part of the world (Kay, 1993). Indeed, many of the assets used to exploit a product innovation, such as distribution facilities and computer information systems, are extremely versatile and not specific to the product at all. The only relevance of asset-specificity lies in the fact that the costs of discovering an opportunity are sunk at the time the opportunity comes to be exploited, so that if there were just a single potential licensee or franchisee then they would be able to bid down the price very low.

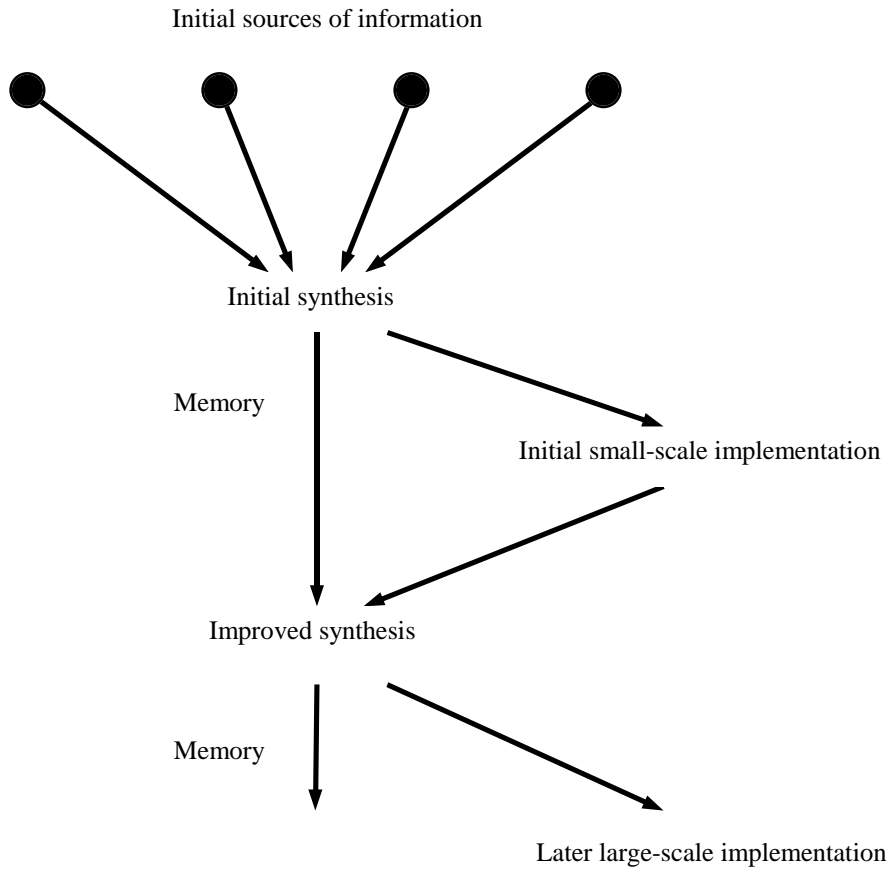
The question of whether transaction costs apply first and foremost to information flow or material flow has important implications for the theory of the firm. Emphasis on the transactions costs of trading information leads directly to a theory of the multinational enterprise in a way that theories of asset-specificity do not. The concentration of multinationality in high-technology advertising intensive industries is most naturally explained by the costs of licensing product innovations (Buckley and Casson, 1976). Extending this idea to the exploitation of more general coordination opportunities suggests that the boundaries of the firm are determined by the scope of the coordination plan that has been developed by the entrepreneur (Casson, 1985). The industrial scope of a firm's operations reflects the range of vision of the entrepreneur. An entrepreneur who plans a radical innovation which, he believes, will rationalise and restructure an entire industrial sector, has a strong incentive to acquire control of resources in all the industries within this sector in order to maximise the appropriation of rents, for the reasons given above. Similarly, an entrepreneur whose vision of an industry encompasses every activity between raw material input and the distribution of the finished product has a strong incentive to integrate all the stages of production, in comparison to an entrepreneur with more limited vision, who may integrate just a single vertical linkage instead.

This point re-emphasises the importance of the holistic perspective introduced at the outset. Modern transaction cost theories tend to work with a partial equilibrium view of the firm, which emphasises just a single linkage involving material flow, whereas the holistic approach affords a vision of an entire networks of such flows, coordinated by another network of information flow. As Schumpeter (1934) suggests, elite entrepreneurs promoting radical innovations in a leading sector have a vision of an industrial network they plan to create, rather than just a single linkage they plan to make, and this is reflected in the scope of the coordination plan that they propose. The economics of internalising any given linkage will certainly influence their plans, but the importance of internalising all the activities encompassed by their vision is the main driving force. It is, therefore, in the appropriation of rents from entrepreneurial vision that the key to the boundaries of the entrepreneurial firm are to be found.

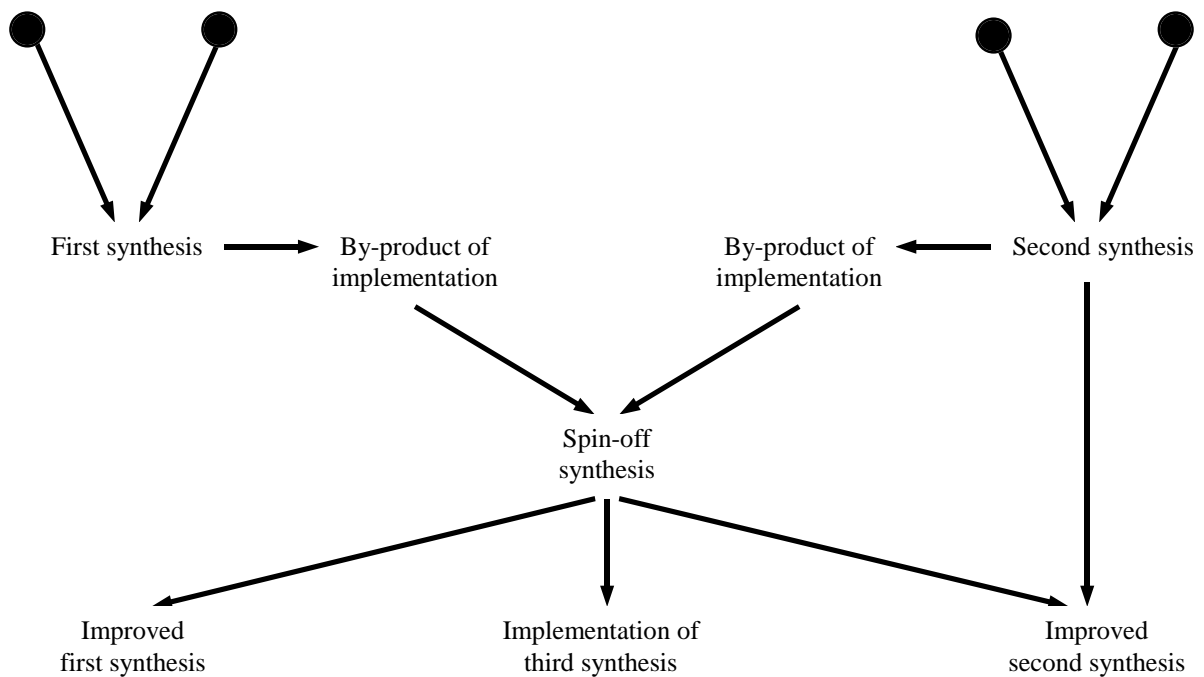
## **15. Diversification and the evolution of the firm**

It may seem improbable that a large-scale sectoral coordination plan could spring

**Figure 4**  
**Incremental synthesis through learning from experience**



**Figure 5**  
**A distributed information synthesis with two-way information spill-overs**



from nowhere into the mind of an entrepreneur. Ambitious plans of this kind are more likely to unworkable schemes devised by a populist politicians to be carried out at public expense. It is possible, though, that such a plan could evolve from the accumulated practical experience of a successful entrepreneur. By continually feeding back information from his recent operations the entrepreneur may be able to augment the scope of his synthesis, as indicated in Figure 4. Because he is accumulating information on persistent changes, his initial information is unlikely to obsolesce whilst additional information is being collected. He can therefore evolve a large scheme from a sequence of increasingly ambitious smaller schemes.

There is a problem, however. As his schemes are progressively refined, the organisational procedures required to implement them will become increasingly complex. As the organisational division of labour becomes more elaborate, the links between individual managerial performance and individual managerial reward are weakened. In addition, jobs are increasingly delegated to people who are not personally known to the entrepreneur. Furthermore, every modification of procedures disrupts existing ones. Continuous growth means that no sooner has one set of procedures been stabilised than it is time to change them again. New recruits have to be trained as the organisation expands, increasing the pressures on experienced employees (Penrose, 1959) Naturally, growth brings opportunities for promotion for those who are able to 'grow with the job'. But, on the other hand, long-serving staff who are unwilling or unable to take greater responsibility may become resentful as they 'get left behind in the promotion stakes'.

One solution is to structure the organisation in a modular fashion. The procedures in each module can remain largely unchanged whilst new modules are added elsewhere. Each module is led by an internal entrepreneur, who 'networks' with other internal entrepreneurs to negotiate inter-modular relations. The effects of change are focused as far as possible on the inter-modular relations, leaving intra-modular relations stable for as long as possible. This goes some way towards introducing the principle of flexible specialisation to the organisation, making the organisation less intrinsically rigid than it was before. It is similar to the process of divisionalisation, but not identical to it, since unlike divisionalisation, it is specifically geared to the problem of maintaining organisational flexibility in a fast-growing firm. (The same principle, incidentally, can be applied to smooth out the running down of a firm through modular divestment, but this issue cannot be pursued further here.)

**Table 2**

**Some types of information synthesis and their implications for the boundary of the firm**

<b>Type of synthesis</b>	<b>Probable institutional response</b>
<b>Singular synthesis</b>	New firm formation
<b>Distributed synthesis</b>	Partnership formation
<b>Combinatorial synthesis</b>	Merger
<b>Incremental synthesis</b>	Diversification
<b>Improved synthesis</b>	Growth through learning
<b>Spin-off synthesis</b>	Inter-firm joint venture

It could, of course, be argued that it is a mistake to incorporate new activities within an existing organisation, and that it is better to establish entirely new firms for this purpose. The advantages of integrating the new activities into an existing organisation are greatest when the same administrative procedures are applicable to both (Teece, Rumelt, Dosi and Winter, 1994). Where the activities face the same pattern of transitory volatility, there are economies of scope in applying existing procedures to the new activity. If the demands for two different consumer products are driven by similar kinds of fashion, for example, then it may pay for them to share the same market research and distribution procedures. Conversely, if the patterns of transitory volatility are unrelated, it will normally pay the entrepreneur to establish a new organisation instead.

Not all incremental synthesis follows the pattern represented in Figure 4, however. Figure 5 illustrates a more complex situation, exemplified by ‘technological convergence’, in which two distinct activities generate information spill-overs for each other. These spill-overs not only afford improvement to the existing activities, but offer scope for the innovation of a new activity too. If both of the original entrepreneurs recognise the new opportunity simultaneously then it will pay them to collaborate rather than to compete in its exploitation (Langlois and Robertson, 1995). Although one entrepreneur could, in principle, buy out the other, they could not be certain that the other would not later re-enter the field. Since the new activity is likely to benefit from information feedback from the original activities, economies of internalisation mean that the speed of improvement of the new activity is maximised when both the entrepreneurs to hold a continuing equity stake .

Figure 5 represents just one of the many forms that incremental information synthesis can take. Some other possibilities are listed in Table 2. This table indicates in the right-hand column the type of corporate growth and diversification strategy that is favoured by each type of synthesis. The correspondence indicated by the two columns underlines the importance of the logic of information synthesis as a factor in the evolution of the boundaries of the firm.

## **16. Financing the firm**

It was noted earlier that inputs normally precede outputs in a coordination plan. It has been shown that the entrepreneur must normally assume responsibility for implementing his own plan. He must obtain inputs by owning or renting the relevant resources. Such rights are acquired through the payment of money. If his vision is very broad then a lot of resources

may be required. Conversely, services are often supplied to consumers bundled up in the form of the goods which are sold in return for money. It follows immediately that coordination requires financing in terms of money.

It is possible that the entrepreneur has sufficient funds of his own for this purpose, particularly if he possesses good collateral (such as a house) to secure a loan. If he does not, then he is dependent on others. He cannot guarantee the repayment of additional loans because he cannot sell his output forward. The only way to guarantee such loans is to obtain insurance for them. This insurance is provided by equity shareholders. By accepting an uncertain and potentially volatile income stream, the equity-holders guarantees a certain and stable income to other parties - not only to fixed-interest lenders, but to waged and salaried employees as well. The simplicity of the equity contract arises from the fact that the equity holder supplies insurance against every kind of risk. It is specified, not by the contingencies it includes, but by those it excludes, which is basically none at all.

The question is why anyone would become an equity shareholder in a firm that is run by another entrepreneur. The problem is not that the firm's environment is uncertain, because the insurance principle can deal with this. If different firms face different local market environments, then the risks can be pooled using a diversified equity portfolio. Furthermore, because the entrepreneur cannot influence the environment, moral hazard is unimportant in this respect. The problems relate entirely to the integrity and competence of the entrepreneur. The entrepreneur may pay less attention to important information than he would if he bore all the risks himself. This can be dealt with, to some extent, by ensuring the entrepreneur remains significantly underinsured - in other words, that he holds a significant proportion of the equity himself. He needs to hold sufficient equity that loss of value will have a major adverse effect on his consumption plans. Material consumption is not everything to the entrepreneur, however, as noted earlier. An ethic of stewardship, and pride in his reputation, may be more important in helping the entrepreneur to keep other people's interests in mind.

Competence is the more difficult issue. Entrepreneurial competence is difficult to screen for, because it is often acquired through experience rather than formal education. It is revealed mainly by performance on the job. The entrepreneur's true competence will be slowly (and only partially) revealed as the coordination project proceeds. Escalating costs, or disappointing sales, may reflect misjudgement of the entrepreneur. But they could be explained by an adverse environment too. Diagnosing the cause of failure is itself an

entrepreneurial task. Indeed, a difference of opinion is likely to emerge over the cause of failure, with the self-confident entrepreneur blaming the environment, while some of the equity holders blame the entrepreneur.

A successful equity holder requires knowledge of entrepreneurs, and their track records, to decide whether to back them or not. They also need to monitor the performance of the firms in which they invest, and interpret the causes of success and failure correctly. As they refine their own judgement in this way, they increase their investments in firms that they believe are run by competent entrepreneurs, and run down their investments in others.

The ability to hire and fire entrepreneurs is also important. A specialist equity holder can operate a market in entrepreneurs - or a market in management control, as it is sometimes called - 'head-hunting entrepreneurs from successful firms in order to 'turn around' failing ones. Indeed, specialised institutions such as investment banks are often formed for this sort of purpose. It is sometimes suggested that entrepreneurs who play an active role of this kind cross-subsidise the more passive entrepreneurs, and this is true - up to a point. There is plenty of scope, however, for these entrepreneurs to appropriate speculative profit from their superior judgement, by secretly buying up shares in the firms that they plan to turn around, and selling them off at a profit later. Although the appropriation may be incomplete, it is not obvious that it is any less complete than in other areas of entrepreneurial activity (see below).

This situation is by no means unique to the firm, however. Most organisations have an elite of members- the stakeholders - who have contributed resources and in return demand a share of the control over the way they are used. They have the greatest authority because, generally speaking, they bear the brunt of the risk that a bad decision may be made. They therefore have the power to replace the head of the organisation.

## **17. Conclusion**

This paper has embedded the theory of the firm within a holistic view of the economy, in which information flow is just as important as material flow. Coordination in a volatile economy depends on the continuing synthesis of new information. Processing this information is costly, and requires a division of labour. There is a vertical division of labour between observers and synthesisers, and a horizontal division of labour between different observers. The latter is particularly significant in dealing with regular changes in transitory factors, where an organisation may hire full-time observers to submit routine reports. These

observers also handle the implementation of the decisions that are made through a synthesis of their various reports, generating a pyramid of authority.

The pattern of variation in transitory factors depends in turn on persistent factors. It is difficult to set up an organisation to monitor persistent factors, because the relevant sources of information cannot be identified in advance. The synthesis of information about changes in persistent factors requires improvisation rather than routine: it requires the judgmental abilities of the entrepreneur.

It is difficult to be sure that the information used to identify persistent change is correct. Different entrepreneurs synthesise information from different sources, and even when they use the same sources they may interpret them differently. To encourage the most able entrepreneurs to specialise in monitoring persistent changes, and to give each individual as much scope as possible, the division of labour between synthesisers is structured in a highly flexible way. Flexibility is achieved through a market system, whose institutions facilitate the transactions by which entrepreneurs gain access to the real resources they require. The formal specification of all the property rights that an entrepreneur may require is a formidable task. To economise on information costs, a comprehensive right of ownership, and an allied concept of rental, has therefore been developed to replace a wide spectrum of other more specific rights. Another institution is the use of money - a homogenous stock that circulates in order to simplify transactions. The firm, too, falls into the same category; it is a specialised institution designed to act as a nexus of contracts, and as such is endowed with fiscal privileges and unlimited life.

Embedded within this institutional framework, entrepreneurs implement their coordination projects by borrowing money and exchanging it for ownership or rental of productive resources, using their firm as a legal shell. In negotiating for command over resources, entrepreneurs are careful not to give away crucial information on which their synthesis is based. Entrepreneurs are organisers too. Because they cannot patent their own ideas, and cannot sell them as secrets either, they are obliged to internalise the market for their ideas. This involves them in the design of organisational structures - the most notable example being the organisation of intermediation effected by the market-making firm.

Successful entrepreneurs must be optimistic and self-confident in order to compete resources away from rival entrepreneurs, and to live with the risk that their judgement may turn out to be wrong. To obtain the widest possible synthesis of the latest information they

cultivate networks of social contacts who feed them the information they require. They are persuasive in raising finance, and astute in bargaining with customers and suppliers.

The high cost of information makes legal enforcement of business contracts problematic, and so trust engineered through moral values is very important in providing a climate of confidence within which risky coordination projects can take place. Traditional values such as honesty and hard work are particularly important for entrepreneurs in motivating their employees. While entrepreneurship thrives on plurality of opinion about coordination opportunities, it requires a degree of consensus on moral issues. The entrepreneurial theory of the firm suggests that a market economy driven by a combination of moral unanimity and pluralism of practical judgement is likely to prove the most successful.

## REFERENCES

- Boisot, M.H. (1995) *Information Space: A Framework for Learning in Organizations, Institutions and Culture*, London: Routledge
- Buchanan, James M. (1978) *Freedom in Constitutional Contract*, Texas A&M University Press.
- Buckley, Peter J. and Mark C. Casson (1976) *The Future of the Multinational Enterprise*, London: Macmillan, 2nd. ed. 1991.
- Casson, Mark C. (1982) *The Entrepreneur: An Economic Theory*, Oxford: Martin Robertson, 2nd. ed. Edward Elgar, 1999.
- Casson, Mark C. (1985) Entrepreneurship and the Dynamics of Foreign Direct Investment, in P.J. Buckley and M.C. Casson, *The Economic Theory of the Multinational Enterprise*, London: Macmillan, 172-191.
- Cantillon, Richard (1755) *Essai sur la Nature du Commerce en Generale* (ed. H. Higgs) London: Macmillan, 1931
- Casson, Mark C. (1987) *The Firm and the Market*, Cambridge, MA: MIT Press
- Casson, Mark C. (1991) *Economics of Business Culture: Game Theory, Transaction Costs and Economic Welfare*, Oxford: University Press
- Casson, Mark C. (1995) *Studies in the Economics of Trust*, Cheltenham: Edward Elgar
- Casson, Mark C. (1997) *Information and Organization: A New Perspective on the Theory of the Firm*, Oxford: Clarendon Press
- Clanchy, Michael T. (1993) *From Memory to Written Record: England 1066-1307*, 2nd. ed., Oxford: Blackwell
- Eliasson, Gunnar, Stefan Folster, Thomas Lindberg, Tomas Pousette and Erol Taymas (1990) *The Knowledge-Based Information Economy*, Stockholm: almqvist and Wicksell

- Foss, Nicolai J. (1997) *Resources, Firms and Strategies: A Reader in the Resource-Based Theory of the Firm*, Oxford: Oxford University Press.
- Granovetter, Mark (1985) Economic Action and Social Structure: The Problem of Embeddedness, *American Journal of Sociology*, **91**, 481-510.
- Harper, David A. (1996) *Entrepreneurship and the Market Process*, London: Routledge
- Hayek, Friedrich A. von (1937) Economics and Knowledge, *Economica* (New series) **4**, 33-54.
- Hedlund, Gunnar (1993) Assumptions of Hierarchy and Heterarchy: An Application to the Multinational Corporation, in S. Ghoshal and E. Westney (eds.), *Organization Theory and the Multinational Corporation*, London: Macmillan, 211-236.
- Hollander, Stanley C. and Kathleen M. Rassuli (eds.) (1993) *Marketing*, Aldershot: Edward Elgar
- Kay, Neil M. (1993) Markets. False Hierarchies, and the Role of Asset Specificity, in C. Pitelis (ed.), *Transaction Costs, Markets and Hierarchies*, Oxford: Blackwell, 242-261.
- Kirzner, Israel M. (1973) *Competition and Entrepreneurship*, Chicago: University of Chicago Press.
- Kirzner, Israel M. (1997) Rationality, Entrepreneurship and Economic 'Imperialism', in S.C.Dow and P.E.Earl (eds.) *Conference to Celebrate Brian Loasby's Work at Stirling University, 1967-97*, 1-21.
- Klein, Burton, R.G.Crawford and Armen A. Alchian (1978) Vertical Integration, Appropriable Rents and the Competitive Contracting Process, *Journal of Law and Economics*, **21**, 297-326.
- Knight, Frank H. (1921) *Risk, Uncertainty and Profit*, Boston, MA: Houghton Mifflin
- Lippman, S.A. and J.J.McCall (eds.) (1979) *Studies in the Economics of Search*, Amsterdam: North-Holland
- Langlois, Richard N. and Paul Pobertson (1995) *Firms, Markets and Economic Change*, London: Routledge
- Loasby, Brian (1976) *Choice, Complexity and Ignorance*, Cambridge: Cambridge University Press
- Loasby, Brian (1991) *Equilibrium and Evolution: An Exploration of Connecting Principles in Economics*, Manchester: Manchester University Press
- Marshall, Alfred (1919) *Industry and Trade*, London: Macmillan
- Milgrom, Paul R. and John Roberts (1992) *Economics of Organisation and Management*, Englewood Cliffs, NJ: Prentice-Hall
- Nelson, Richard and Siney G. Winter (1982) *An Evolutionary Theory of Economic Change*, Cambridge, Mass: Harvard University Press
- North, Douglass C. (1981) *Structure and Change in Economic History*, New York: W.W.Norton
- Penrose, Edith T. (1959) *The Theory of the Growth of the Firm*, Oxford: Blackwell
- Putterman, Louis (1990) *Division of Labour and Social Welfare*, Oxford: Oxford University Press
- Say, Jean-Baptiste (1803) *A Treatise on Political Economy*, New York: Augustus M.Kelley, 1964.
- Schultz, Theodore W. (1971) *Investment in Human Capital: The Role of Education and of Research*, New York: Free Press
- Schumpeter, Joseph A. (1934) *The Theory of Economic Development* (trans. R. Opie) Cambridge, MA: Harvard University Press
- Shackle, George L.S. (1979) *Imagination and the Nature of Choice*, Edinburgh: Edinburgh