Part II

Social enterprises
A theoretical perspective
16 Social enterprises as incentive structures

An economic analysis

Alberto Bacchiega and Carlo Borzaga

Introduction

In recent years, the third or non-profit sector (as opposed to the private for-profit and public sectors) has attracted increasing attention from policy-makers and public opinion. The reasons for this are twofold. On the one hand, the organisations that constitute the third sector have operated successfully in industries (health, cultural, leisure and welfare) which have experienced considerable growth in recent years. On the other hand, these organisational forms have undergone a constant process of evolution, which has enabled them to meet part of the growing demand for social, community care and collective services. Since the 1980s, third-sector organisations have gradually acquired a productive and entrepreneurial role. Productive third-sector organisations providing services have become, especially in Europe, increasingly active in areas previously dominated by governmental or charitable intervention, such as social services and the work integration of disadvantaged people. In countries where the provision of these services was scarce and mainly undertaken by public institutions (e.g. Italy and Spain), the emergence of third-sector entrepreneurial organisations is an almost entirely new phenomenon. In other countries (e.g. France and Belgium), where private third-sector organisations were already involved in the provision of welfare services, the trend has been towards their greater autonomy from the public authorities.

The term 'social enterprise' has come into use to distinguish the new entrepreneurial forms from more traditional third-sector or non-profit organisations. The distinction underlines the growing involvement of the new organisations in the production of services and it is this which differentiates them from traditional charities. It is also necessary, because the non-profit distribution constraint – the characteristic of third-sector organisations most stressed in the economic literature – does not seem to be crucial to distinguish them from the private for-profit enterprises. Indeed, although all social enterprises pursue objectives other than profit maximisation, not all of them are subject to a non-profit distribution constraint.

The key feature of social enterprises seems to be their ability to strengthen the fiduciary relationship within and around the organisation, and to mobilise resources from individuals and from the local community (social capital). They do
so using institutional and organisational mechanisms that rely, inter alia, on the forceful and broader representation of the interests of stakeholders, on a participatory and democratic governance system, and on the use of volunteer labour. As a consequence, the simple dichotomy between for-profit and non-profit, so widely used in the economic literature, does not suffice to explain the emerging organisational differentiation within the third sector, and particularly the evolution of social enterprises as described above. In our view, this dichotomy stems from the emphasis in this literature on a very specific organisational form – the foundation – and its insufficient use of the insights offered by the institutional approach to the study of organisations. We believe that there should be reconsideration of the nature of third-sector organisations, bearing in mind that the sector is made up of a plurality of organisations far richer and more complex than the homogeneous entities described by the non-profit distribution constraint alone.

We will argue that social enterprises implement a number of institutional arrangements whereby the objectives of the organisation (to serve the community or a specific group of people, and to promote social responsibility at the local level) give rise to original incentive and relational systems involving donors, consumers, public authorities, volunteers and employees. Relations of trust with the local community, the users, volunteers and workers are guaranteed by the involvement of different stakeholders, and by democratic management.

We first develop a general framework for the analysis of social enterprises from an institutional point of view, and then concentrate on a specific aspect that has been neglected by existing studies, viz. the place of workers in the organisation. We will argue that, in social enterprises, incentives for workers are not based exclusively on monetary rewards; rather, they derive mainly from workers’ involvement in shaping and sharing the organisation’s goals and mission. As a consequence, opportunistic behaviour may be reduced and the functioning of the organisation improved.

The chapter is organised as follows. A brief introduction to the institutional approach to organisations is presented in section 1. The purpose of this section is to summarise the main elements of the theory of institutions. The shortcomings of existing theories in analysing the recent evolution of the third sector are considered in section 2. The peculiarities of the markets in which social enterprises operate are highlighted in section 3. In section 4, the institutional specificities of social enterprises are analysed, and in section 5, a view of social enterprises as original incentive systems is proposed, and the main strengths and weaknesses of these organisations are highlighted.

1 The firm: production function versus co-ordination mechanism

The neo-classical and institutional approaches

In this subsection we briefly compare the neo-classical view of the firm and the view based on the (neo-) institutional approach. Neo-classical theory seeks to explain resource allocation as a process guided by prices (Demsetz 1977). The relevance of the firm to the understanding of the economy is marginal; once the available technology and the prices of inputs and outputs are given, there is no discretion in the production process. With perfect competition, the entrepreneur needs only to choose the best technology among well-known alternatives, go to the market to buy the machines (whose price is known), hire the workers (at the wage established in the labour market), and sell the goods produced (at the price determined on the market). Technology is also the determinant of the firm’s size.

Whatever happens inside the firm is of no interest because there are no particular decisions to be taken, and any rational agent would act in the same way. The firm is a ‘black box’, a production function in which resources go in and goods come out. There is no need to open the black box, because nothing of interest is inside; all the interesting things happen in the market. Indeed, it is in the market that all relevant contracts are stipulated and that transactions take place.

The neo-classical view has been challenged by those theorists who consider the firm to be more than a mere production function. Firms involve more than the transformation of inputs into outputs alone. They are institutions which coordinate and manage costly transactions. Firms rely only partially on price signals and contracting, and base their internal relations on other co-ordination mechanisms (hierarchy, communication, ownership, and so on). Firms may therefore constitute a more efficient institutional setting than markets for the conduct of some economic activities.

Moreover, firms may enable transactions to take place when price mechanisms do not convey the actual value of goods and inputs (i.e., when markets fail). The market institution is costly insofar as there are costs involved in finding a suitable partner for the transaction, in drawing up the contract, and in specifying the relevant clauses. The costs of using the price system are greatly increased if information is asymmetrical and if the future is only imperfectly known. In the latter case, it is extremely difficult to devise contracts that foresee all possible eventualities, and the co-ordination of transactions within the firm becomes comparatively cheap.

Firms may be able to reduce the costs of using the market, or the costs of contracting. This is because relations within firms do not have to foresee all possible eventualities; they need only define the scope of the co-ordination to take place between agents. For example, a labour contract would contain the maximum number of daily working hours, the wage and a rough description of the tasks to be performed, leaving the entrepreneur free to determine the specific content of the job.

However, relations within firms involve costs as well. There are the costs of setting up and running the organisation, of delegating tasks and controlling subordinates, of collective decision-making (when, as is often the case, more than one agent shares authority), and of bearing the risk involved in the enterprise. These costs can be labelled ‘co-ordination costs’ or ‘costs of ownership’. As Hansmann (1996) puts it, the institutional choice between markets and firms for
the relevant transactions depends on the relative size of the 'costs of contracting' versus the 'costs of ownership'.

It is worth noting that the firm as a co-ordination mechanism is a much more flexible concept than the firm as a production function. While the latter leaves no room for diversity (all firms, given prices and technology, are the same), one can imagine a great number of possible co-ordination structures. The entrepreneur making decisions within the contractual limits is a way of co-ordinating transactions, but this does not exhaust the range of options.

We may, accordingly, say that institutional theory is better able than neo-classical theory to explain the variety of organisational forms that exist in reality. In fact there is no such thing as 'the firm'. Rather, there are a number of institutional arrangements that substitute for the price system in governing transactions (e.g. investor-owned firms, customer-owned firms, worker-owned firms, and non-profit firms). The organisational form that most efficiently performs a given type of transaction in a given environment is the one that is able to minimise transaction costs, i.e. the sum of contracting and ownership costs (Hansmann 1996). From this point of view, each organisational form is a particular institutional answer to a situation in which the markets do not guarantee an optimal solution, with its own costs and benefits. According to Arrow (Arrow 1963: 947), 'when market fails to achieve an optimal state, society will, to some extent at least, recognise the gap, and non-market social institutions will arise attempting to bridge it'.

The idea of the firm as an institutional alternative to the market is closely bound up with the idea that the main and most pervasive market failure is due to the imperfect and asymmetric distribution of information. This applies both to knowledge about technology and the price system and to the division of labour within firms. Thus, the co-ordination role attributed to the entrepreneur has been justified in terms of his/her ability to deal with imperfect information. Aghion and Demsetz (1972) claim that the role of the entrepreneur is to control the effort of the other contractual parties when the contribution of each to the final output cannot be attributed clearly. In order to motivate the entrepreneur in this endeavour, he/she is given ownership of the firm, including the right to co-ordinate all contractual relations with other agents, and the right to retain the residual earnings after all contracts have been honoured.

The firm as a nexus of contracts

Co-ordination within the firm is not per se an alternative to market contracting. The firm may be viewed as a nexus of contracts where the agents voluntarily accept the co-ordination role performed by the owner, but give him/her no authority or special rights apart from those explicitly stated in contracts (Demsetz 1997). Under this approach (sometimes called the new-institutional theory), the rationale for the existence of the firm is not the superiority of internal co-ordination with respect to the price system, but the need to have someone to monitor the behaviour of numerous different agents in complex production processes. One assumption of this approach is that all agents are able to optimally process all the available (asymmetrically distributed) information at the time when they write contracts. Contracts are thus optimal, in the sense that given the available information, the (second) best outcome is achieved.

This new institutional theory views the firm not as an antagonist to the price system but as an organisation which fully incorporates the price system into its contracts, in an environment characterised by information asymmetries among agents. From this point of view, new-institutional theory bridges the gap with neo-classical theory. The entrepreneur, however, still has an important role to play in co-ordinating contracts. The diversity of organisational forms can be explained under this approach in terms of the diversity of the systems of contractual arrangements within firms. Owners may exercise their right to co-ordinate contractual arrangements personally, or they may appoint other agents (the managers) to act on their behalf. An organisation may pursue a variety of objectives, and it may use different types of contract to motivate agents to take part in production.

The firm as an alternative co-ordination mechanism to the market

The contractual view of the firm considered in the previous section has been criticised by those economists who view the nexus of contracts as nothing more than the reintroduction of the market system into production processes. According to the institutional approach, organisations cannot be reduced to a web of prices and contracts; they are fully fledged alternatives to the market. Organisations are defined not only by the co-ordination of different agents and by the rewards that they obtain from participation in production (price mechanism); they also entail authority and trust relations among agents, and the identification of agents with the organisation and its goals (co-ordination mechanism).

It is possible to identify two main interrelated objections against the view of organisations as simple nexuses of contracts. The first is based on bounded rationality, and the second on the incompleteness of contracts. These will now be considered in turn.

Bounded rationality challenges the idea that agents in firms write optimal contracts. Empirical observation shows that actual contracts within firms are much simpler than predicted by optimal contracting theory. The bounded rationality approach maintains that this is because, in reality, agents do not have the computational and rational abilities that the new-institutional theory requires in order to derive optimal contracts. As a consequence, relations within firms cannot be understood solely by relying on self-interested interactions among perfectly rational agents. In particular, the participation and motivation of workers in organisations cannot be studied within the usual principal-agent framework. They depend on a larger set of variables, which Simon (1991)
summarises as authority, reward, identification and co-ordination. While authority, reward and co-ordination can be explained through optimal contract theory, identification with the goal of the organisation does not seem compatible with the assumption of maximising, self-interested individuals.

The identification with organisational goals is, according to Simon, closely related to the idea that agents do not behave as maximising individuals; instead, they display bounded rationality. Imperfectly rational agents internalise the objectives and values of the organisations in which they work. Simon (1993) shows that boundedly rational agents with a certain degree of 'docility' (i.e. they are ready to accept the values of the organisation as their own, even if these values do not affect their utility function) are, from an evolutionary point of view, more resistant than purely self-interested individuals.

The contract incompleteness approach to the study of organisations is relatively recent.\(^7\) Like the bounded rationality approach, it starts from the empirical observation that actual contracts are usually strikingly simple - contrary to the predictions of optimal contract theory. Incomplete contract theory suggests that this is so because there are prohibitive costs to writing optimal contracts. Even when it is possible to envisage all future events and their associated probabilities, giving exact contractual specification to all of them entails costs. Even assuming that these costs are not particularly high, it may still happen that a third party (e.g. a court that must decide whether the contract has been fulfilled) is unable to verify important events that can be observed by the parties involved in the contract. For this reason, contracts are often incomplete insofar as they do not specify all possible events but instead specify who has the authority to take decisions (i.e. who possesses 'control') in situations not covered by the contract. The incomplete contract approach therefore regards the firm as a truly alternative co-ordination mechanism with respect to the market institution. At the same time, it leaves room for diversity among organisational forms in the ways in which authority can be distributed within an organisation.

As noted above, the rationale for choosing a particular organisational form resides in its capacity to reduce transaction costs. Both bounded rationality and contract incompleteness provide a theoretical foundation for the comparative study of organisations as institutions that seek to minimise the costs of contracting (i.e. the costs of carrying out transactions in the marketplace) and the costs of ownership (i.e. the costs of carrying out transactions within the firm). In the sections that follow, we shall mainly use the conceptual categories of incomplete contract theory (assuming that workers are rational, although not purely self-interested) to analyse the existence, the evolution, and the diversification of third-sector institutions. However, it should be noted that incomplete contracting and bounded rationality are, to some extent, complementary in the theory of organisations.\(^8\) Very similar conclusions to ours could be reached starting from an approach based on bounded rationality.

2 The institutional approach and third-sector organisations: existing theories and their limitations

Economists have drawn largely on institutional theory to explain the existence of third-sector organisations and, more generally, of organisations and enterprises owned by stakeholders different from investors. A very brief review of the economic literature on third-sector organisations may help to illustrate this point.\(^9\) This literature has focused principally on four explanations for the existence of such organisations:

- Hansmann (1980) considers mainly non-profit organisations, and interprets them as less liable to exploit consumers than their for-profit counterparts when producers have an informational advantage on the quality of the product;
- Ben-Ner (1986), extending his analysis to include other third-sector organisations, stresses the role of consumer control on output production in overcoming the asymmetric information problem;\(^10\)
- Weisbrod (1977) considers third-sector production to be the private response to a public failure to satisfy heterogeneous demand for public goods. This theory has since been integrated with Hansmann's (Weisbrod 1988);
- the role of entrepreneurs and managers in shaping the objectives of third-sector organisations is stressed by Young (1985, 1997) with regard to entrepreneurial preferences, James (1989) who considers religious and political beliefs, Rose-Ackerman (1987) who analyses cross subsidisation of preferred non-profit activities through market production, Glaeser and Shleifer (1998) who describe the choice of a non-profit organisation as a rational entrepreneur's commitment to soft incentives.

To a large extent, these interpretations of the economic role of third-sector organisations are compatible (Anheier and Ben-Ner 1997; Krashinsky 1997), and they have often been used jointly. For example, Hansmann and Weisbrod explain why consumers may want to buy from non-profit organisations, while Ben-Ner and the 'entrepreneurial approach' explain why people may decide to set up a third-sector organisation. Not all the theories consider the non-profit distribution constraint to be crucial for understanding the phenomenon. However, the most influential explanation, which is almost universally accepted, is the one proposed by Hansmann. As a consequence, the 'non-profit-distribution constraint' has become the most important characteristic used to define the organisational typology.

In Hansmann's theory, the non-profit distribution constraint coincides with the absence of agents formally identified as owners of the organisation. The importance of this characteristic can be better understood if we recall the two principal formal rights associated with ownership, (1) the right to select, sign and co-ordinate contracts with the agents that establish a relation with the organisation, either as suppliers of inputs or as purchasers of the goods or services.
produced, and (2) the right to appropriate the residual income of the organisation, once all contractual relations have been honoured.11

The fact that these are formal rights obviously means that it is not necessary for them to be directly exercised by owners. In several organisational forms some ownership rights are delegated to some other agents. In particular, the right to decide on contracts to be signed is often delegated to managers, who thus have effective control over the organisation. The owners retain a right to exercise control over the actions of the managers.13

In his early works, Hansmann (1980) focused mainly on the right to the residual income of the organisation, disregarding the issue of the allocation of control rights. The essence of the non-profit distribution constraint is, in fact, the formal recognition that nobody is entitled to the residual income of the organisation. This view was expressed by Hansmann in the following way: 'a non-profit organisation is, in essence, an organisation that is barred from distributing its net earnings, if any, to individuals who exercise control over it' (Hansmann 1980: 830). Since the organisation cannot distribute profits, the owners gain no advantage from exploiting any superior information that they may possess on product quality in order to maximise profit. Consumers are therefore somewhat protected against contractual failure.

In this approach, third-sector organisations have very often been identified solely with this highly specific allocation of residual income: the non-profit distribution constraint. As a consequence, the complexity of the tertiary sector has been reduced to a simpler set of organisations, namely non-profits. Many of the other interesting features of third-sector organisations have been neglected as a result of this reliance on the non-profit distribution constraint in defining the sector.14 This is particularly the case in many European countries, where the non-profit distribution constraint is not the main characteristic of most third-sector organisations, as the contributions to this book show.

More recently, Hansmann (1996) argued that non-profits arise when both the costs of contracting and the costs of ownership are too high for some category of patrons. In this case, defining specific owners for the organisations is inefficient and the control role is substituted by a fiduciary relationship with the managers, who possess effective control over the organisation, so that 'the non-profit form abandons any benefit of full ownership in favour of stricter fiduciary constraints on managers' (Hansmann 1996: 228).

However, it is not clear what the origin of the fiduciary relationship is, nor how it actually takes place, nor what factors guarantee that it will be respected. Implicitly, a great deal of importance is once again given to the non-profit distribution constraint, as the mechanism that prevents opportunistic behaviour by managers. The theory rests on two main assumptions. The first is that the absence of agents formally entitled to the residual income effectively protects consumers against opportunistic behaviour. The second is that for-profit firms do not have an incentive to establish a reputation as being non-exploitative of consumers. However, neither assumption seems particularly robust (Ortmann 1996). In fact, the non-profit distribution constraint has proved highly ineffective in preventing the onset of opportunistic behaviour and the exploitation of consumers. If the fiduciary relationship cannot be enforced, and in the absence of explicit residual rights of control, organisations seem unable to prevent the exploitation of their resources by the agents who hold effective control, they therefore fail to operate efficiently.

Moreover, the economic literature has focused mostly on the relationship between organisations on the one hand, and the donors or the customers on the other. The relations among agents within organisations have not been thoroughly analysed. For example, little attempt has been made to interpret the motives that induce workers and managers to supply their labour to third-sector organisations and to behave in accordance with the goals of their founders. In other words, attention has concentrated on the capacity of third-sector organisations to cope with failures in the product market, while their ability to remedy failures in the control of managers and workers (agency costs) has been neglected.

This theoretical weakness is all the more evident when the theory seeks to explain the more recent forms of third-sector organisations, those that engage in the production and sale of social services to government agencies or directly to the public. In many of these services, in fact, both consumers and government agencies are able to exert some control over quality, while the for-profit provider is able to overcome information problems by investing in reputation (Ortmann 1996). Indeed, Hansmann himself recognises that his theory is mainly a theory of donative non-profits. He admits that, when non-profits produce private goods or services, the problems of asymmetric information 'are inadequate to explain why investor-owned firms do not have an even larger market share in these industries than they already do. The non-profit form is a very crude consumer protection device' (Hansmann 1996: 234–5). Hansmann's conclusion is that the main strength of non-profits which produce goods or services is often their being well-established incumbents in the market, since they do not gain significant comparative advantages from mitigating the cost of contracting. This conclusion highlights the difficulties that beset the view of third-sector organisations as based on the non-profit distribution constraint when it seeks to explain the complexity of the phenomenon and its recent evolution.

For a more general theory, it should be borne in mind that third-sector organisations can assume a plurality of roles including: advocacy for a particular group of citizens or for certain rights; redistribution of resources among individuals, groups, or activities; and the stable and continuous production of social and collective services. In a given period of time, an organisation may more actively engage in one of these activities and, in many cases, entirely specialise in it. The contributions to this book clearly identify a trend in third-sector organisations towards the production of services, with an increasingly entrepreneurial orientation in the marketplace. Often, social enterprises have developed out of already existing consumer or volunteer associations that focus their activity on advocacy. This is so in the case of the Italian social co-operatives. The shift from advocacy
toward direct production has been a consequence of the decreasing role of the public administration in social services provision.\textsuperscript{15}

The different roles played by third-sector organisations may explain the co-existence of a variety of organisational and legal forms in the sector. The social, legal and economic context, the evolution of needs, changes in the demand for services, and the development of alternative forms of service provision (public and for-profit) may help in explaining the shift of third-sector organisations from one role to another. Social enterprises are part of this evolution. In particular, they represent an entrepreneurial evolution of third-sector organisations. Understanding this evolution requires that one must first specify the distinctive features and the market failures of the industries in which they operate. One must then determine how the institutional characteristics of social enterprises can exert a relative advantage over other organisational forms in those industries.

### 3 The peculiarities of personal and collective services

A striking characteristic of the evolution of the personal and collective services sectors over the last twenty years is that supply has not been able to match, either in quantitative or qualitative terms, the rapidly growing potential demand for what are commonly considered to be merit goods.\textsuperscript{10} In other words, the traditional producers of these services have grown increasingly unable to satisfy new and recognised needs arising from society. There are two reasons for this trend. Firstly, the socially efficient level of production may not be achieved because, at that level, consumers' capacity to pay is below the price charged by traditional suppliers. In this case, the efficient level can be achieved only if the provider is able to mobilise private or public resources so that a reduced fee can be charged to consumers. In other words, the organisation must be capable of mobilising and redistributing resources in favour of the beneficiaries of its services, thus shifting their demand to the efficient level. Secondly, the socially efficient level of production may not be achieved because it is impossible or excessively costly to organise a stable supply to serve an existing and willing-to-pay demand. This may be the consequence of market or organisational failures, whereby traditional providers cannot satisfy the needs arising from society. If a new organisational arrangement is able to overcome the causes of the failure, it will yield an efficiency gain in providing the service. Both elements seem relevant in explaining the provision of social and collective services by social enterprises, in which both income redistribution and innovative production organisation play an important role.

Regarding redistribution, the need arises when initial endowments of income and abilities are perceived to be unjust. Even when this need is recognised by substantial groups in society, government policies may be unable or unwilling to guarantee the desired redistribution of the resources. Private individuals, or groups of people, may then take action to alleviate this failure through organisations based at least partly on donations and volunteer work. The activity of these organisations may take different forms, such as: advocacy for more effective government policies, the direct redistribution of wealth, or the production of specific social services to meet the needs of disadvantaged groups. While third-sector organisations are active in all these areas, social enterprises operate in the latter category, providing services to groups unable to pay a price that fully covers the costs and for which public authorities are unable or unwilling to intervene. This is the case for many new social services, for which a private organisation is able to operate only if its stakeholders accept price discrimination (if they are consumers), or if they agree to provide the necessary financial resources (if they are donors) or production factors (if they are workers or volunteers) free of charge or at a price below market levels.

The market and organisational failures that affect social services often depend on the multi-dimensional nature of either the service or its provision. Typically, social services comprise qualitative dimensions that can be evaluated differently by different users. The multi-dimensionality of a service does not per se constitute a source of market failure, nor is it incompatible with provision by traditional entrepreneurial forms. It may indeed be possible to devise contractual forms, which specify the desired level of each dimension of the service as constraints in the maximisation problem, then leave the organisation free to maximise profits. A mechanism of this kind, however, is not available if some of the objectives are non-identifiable or non-measurable. In this case, the dimensions that cannot be clearly inserted as constraints in the maximisation problem will inevitably be ignored or overlooked by the agents who possess effective control over the organisation (Holmström and Milgrom 1991).

The problem of non-verifiability is particularly relevant to services with a high relational component,\textsuperscript{17} for which some dimensions are easily verifiable (for example, unit cost or number of clients served) but others less so (for example, the quality of the service, the degree of client satisfaction, or the relational effort made by the workers). In this case the dimensions which cannot easily be monitored and verified by the parties directly involved (producers and consumers), as well as by third parties, are crucial for the assessment of the service's quality. Indeed, the quality of the service depends on the informal relationship between provider and consumer, and it is therefore particularly hard to measure. This may lead to relations characterised by asymmetric information of various types, which we shall now discuss.

The first type of information failure, and the one most frequently mentioned in the non-profit literature, is the informational advantage of producers over consumers concerning the service provided. If consumers are able to pay, but unable to control the behaviour of producers efficiently without incurring substantial costs, they tend to be exploited by profit-maximising producers who enjoy an informational advantage. Third-sector organisations, on the other hand, may be less willing to exploit consumers because they cannot distribute profits or, as in the case of many social enterprises, because they involve consumers or volunteers in management and control of the organisations. These institutional characteristics reduce consumer exploitation, so that third-sector
organisations may thus help to reduce the costs of contracting when consumers are weak. This view was considered in section 2 and will not be discussed further.

A second type of information failure characteristic of the social services sector is the informational advantage possessed by consumers with regard to their willingness to pay for the services that they desire. This gives rise to the well-known ‘free-rider’ problem. Third-sector organisations and social enterprises may attenuate this problem insofar as they induce consumers to reveal their true preferences and allow for price discrimination. Some authors show that co-operatives and other third-sector organisations may help overcome the free-rider problem (Grillo 1999; Bilodeau and Silvinski 1994). Ben-Ner (1986) considers the case in which consumers take control of the organisation, thereby at least partly eliminating the incentive for free-riding. In this case the high costs of contracting (due to producer-consumer and consumer-producer asymmetric information) induce substitution of the market institution with an organisation controlled fully or partially by consumers.

A third type of information failure arising in the production of social services stems from the incompleteness of labour contracts and relates to the costs of ownership. Managers and workers have better information than their principals do on the effort that they put into their work. This applies to any type of production, and generates the much-studied agency relationships within organisations. In the case of personal and community services, however, the problem of effort monitoring is particularly intractable, given the difficulties in organising well-functioning labour relations. \(^{18}\) It follows that an organisation able to cope efficiently with this problem may be more successful in delivering such services.

However, the latter issue has not received a great deal of attention in the economic literature on non-profit third-sector organisations, with the exception of the relation between the organisation and its managers. \(^{15}\) The experience of social enterprises shows the importance of the relationship between workers and organisation, as a determinant of the cost and quality of the services produced. For this reason, the remainder of this chapter is devoted to this point in particular, analysing how the institutional features of social enterprises are able to alleviate the effort-monitoring problem.

4 The institutional specificity of third-sector organisations and social enterprises

In most European countries, in the last twenty years, third-sector organisations and social enterprises have been the protagonists of a shift toward a more productive and entrepreneurial stance. Indeed, they offer a stable, continuous, and to some extent autonomous production of a fairly well-defined and limited range of services, characterised by one or more of the following dimensions: labour intensive technology, difficult effort monitoring, and the inability of customers to pay the full cost of production. These services, or the method of their delivery, are often innovative, especially when compared with existing private and public production. The relative importance of some dimensions of the services compared with others translates into the different organisational and legal forms (co-operatives, associations) that social enterprises may assume.

Given the complexity of services, and the variety of social enterprises that have evolved to deliver them, analysis of the phenomenon requires full deployment of the tools offered by institutional theory, paying special attention to the allocation of control rights within organisations. In social enterprises, the allocation of control rights is no less important than the destination of residual income. This is because one of the control rights assigned to the agents entitled to exercise them is the crucial right to determine the objectives of the organisation. This right is exercised through the ultimate control over all contractual relationships, which influences the incentive system within organisations and thus their internal structure.

Like many third-sector organisations, social enterprises do not possess clear-cut mechanisms with which to define those entitled to control rights, and consequently to determine the objectives of organisations (as opposed to for-profit organisations, where it is the contribution of risk capital that matters). \(^{20}\) However, in the absence of predetermined control, right assignment does not imply the absence of a control structure (Alchian and Demsetz 1972; Hausmann 1996). The control structure has a different, and potentially more variable, characterisation. The variety of systems for the allocation of control rights reflects the differentiation of the third sector and of social enterprises. The literature has already recognised this feature. For example, some stress the role of donors as the providers of the financial resources in donor-based non-profit organisations, and thus justify their entitlement to exercise the residual rights of control (Fama and Jensen 1983); others argue that there is no clear-cut way to identify who is entitled to control rights, and the allocation follows informal patterns (Ben-Ner, Montias and Neuberger 1993); while Ben-Ner and van Hoomissen (1991) in discussing mainly consumer co-operatives and associations, simply entrust consumers with ultimate control rights. Often, as many of the studies in this book show, control rights are entrusted to more than one category of stakeholders (Borzaga and Mittone 1997). Control over the organisation and over the definition of its objectives gives these stakeholders an incentive to establish and efficiently manage the organisation in the pursuit of its chosen objectives.

In social enterprises, the identification of the controlling stakeholders is a consequence of the type of redistributional needs and market failures that characterise the production of a specific good or service. However, one cannot expect social enterprises to have the same type of control that is observed when the residual income claimants are fully and formally identified. This is due not only to the different institutional mechanisms possessed by the organisation; it is also related to the fact that the objectives of the organisation are not as clear and well established as they are in its for-profit counterpart. On the one hand, the separation of rights of control from rights over residual income strongly suggests that the for-profit motive does not figure importantly among the organisation's formal objectives. \(^{21}\) On the other hand, the alternative formal objectives cannot be
stated independently of the conditions that characterise the market in which the organisation operates.

If a line of reasoning applied to government institutions (Tirole 1994) is extended to social enterprises, it is likely that the profit maximisation objective will be substituted, not by a single, alternative goal for the organisation but by a plurality of potentially conflicting objectives. The presence of multiple objectives is clearly more likely when control rights are shared among different categories of stakeholders. These objectives are often intrinsically very difficult to quantify and codify because they concern qualitative dimensions of the activity or general principles (e.g. democracy in the organisation, customer well-being, a certain degree of resource redistribution, and so forth), the translation of which in precise guidelines is extremely difficult. A related problem is that, even when the objectives can be identified, it is difficult to weigh them in order to obtain a stable and well-defined objective function to maximise. ²²

Social enterprises are thus highly complex and diverse organisations, and the national chapters in this book give a broad picture of this complexity and diversity. Consequently, it is difficult, and sometimes simplistic, to apply a single organisational model when seeking to explain them. The degree of effective control and the category of patrons controlling the organisation may vary according to the type of organisation. Goals and objectives are not always clearly codified even within organisations, and stem instead from the history of the organisation and the shared values of its stakeholders. However, this is very different from assuming that their characterising elements can be identified with the non-distribution constraint alone. The patterns of allocation of control rights, together with a consistent governance structure, are of particular importance in productive and entrepreneurial third-sector organisations like social enterprises, given the amount of potential conflict that may arise between the objectives related to the production process and those related to the distribution of the benefits that it generates. At the same time, the allocation of control rights is the main institutional instrument with which a social enterprise can build an incentive structure within the organisation, which is consistent with its objectives.

5 Social enterprises as incentive structures

The actual allocation of control rights in social enterprises can be explained in the light of the two main difficulties that arise in the production of social and collective services, viz. the need to mobilise resources for redistribution and the presence of market and organisational failures. When the redistributive component is very substantial, and social enterprises need a high level of support in the form of donations and volunteer labour, control rights are allocated mainly to volunteers, trustees and local community representatives. When, on the other hand, the redistributive component of the organisation’s activity is less pressing, for example when the public authority recognises it by subsidising productive social enterprises, control rights are allocated primarily in order to cope with market and organisational failures. In the latter case, control of the organisation is assigned to clients, workers and managers. Moreover, workers and managers often accept monetary compensations which are lower than those that they would receive for comparable tasks in other organisations, thereby accepting a partial redistribution of resources from labour suppliers to consumers.

The picture arising from these considerations is quite complex and dynamic, as the history and the experiences of social enterprises testify. However, since the problems related to the redistribution aspects have been widely analysed by the literature on non-profit organisations, this section is devoted in particular to the ability of social enterprises to mitigate the negative consequences of failures in workers’ contractual relations.

The non-standardised nature of social and collective services gives rise to difficulties in effort monitoring and in the use of traditional contractual instruments. Agency problems within the organisation cannot be solved by linking workers’ remuneration to the dimensions of their performance that can be monitored. In other words, incentives in the organisations that produce these services tend to be low-powered. This engenders incentive structures in which relatively little weight is given to the more easily measurable and verifiable dimensions of production in order not to jeopardise the pursuit of less verifiable objectives.

One consequence of low-powered incentives is that agents have limited motivation to behave efficiently. Since their remuneration is only loosely related to performance, low commitment on the part of workers with no monetary motivations is to be expected. At the same time, opportunistic behaviour may affect output in such a way that production is not viable. This is especially the case when the price determined by public or private demand for the services only partially pays for their cost, leaving the organisation with a redistributive commitment in favour of its clients. In this case, the organisation must mobilise human and financial resources and select workers willing to exert effort not only in exchange for monetary compensation.

The organisational structure of social enterprises seems better able to avert the danger of opportunistic behaviour, when compared not only with for-profit enterprises and public agencies but also with traditional third-sector organisations (e.g. foundations). This is so because the organisational structure can give rise to incentive systems, which are consistent with organisational objectives. In particular, the nature of the organisation influences the type of incentive system offered to workers and managers. The choice of a particular institutional form thus signals to stakeholders the kind of objectives that the organisation will pursue, and consequently the incentives that it will offer.

The characteristics of social enterprises that seem crucial for definition of their incentive structure are closely related to their social dimension. Particularly relevant are: the existence of an explicit social aim to serve the community or part of it; the direct involvement of a group of people belonging to the community or sharing common needs; limited profit distribution; and democratic and open management. These characteristics are closely related. We will consider them in turn, highlighting their common features.
The existence of an explicit social aim in social enterprises is an important signal of the organisation's nature to both workers and other stakeholders, and also to the outside world. As a consequence, both existing and potential workers and volunteers are able to measure their abilities and expectations against the organisational goals that they will be asked to pursue. This can be seen, for example, in work-integration services, where actual and prospective workers are made aware of the fact that part of the added value they contribute will be used to reward and improve the human capital of disadvantaged workers. It is also evident in social services production, where the organisations commit the realised profits to favour the development of new activities or to broaden the number of beneficiaries. Moreover, explicit social goals and monetary objectives are to a large extent mutually exclusive. A clear signal that no monetary objectives are pursued seems to be consistent with the production of services with non-measurable or non-verifiable dimensions. As noted above, contracts for the production of such services tend to be low-powered, and remuneration unrelated to observable performance.

In fact, social enterprises tend to use a mix of monetary and non-monetary components in a remuneration package in which the wage component is less substantial than in for-profit organisations operating in the same industries. The monetary dimension is not the only, or even the main, dimension of the exchange as much for the agents as for the principal. In other words, social enterprises seem able to motivate workers by using non-wage incentives. Of course, the demand for labour expressed by social enterprises can find a matching offer only if there are workers willing to exchange extrinsic (wage-related) rewards for intrinsic motivations. Both experimental (Frey 1997) and empirical evidence suggests that this is the case. A recent empirical study on employment conditions in the Italian social services sector shows that for the large majority of workers in social enterprises the choice of organisation and sector of activity was a consequence of a specific interest in the content of the job, and not dictated by lack of alternatives or attractive monetary rewards (Borzaga 2000). At the same time, work satisfaction was higher for the workers who have joined the organisation in the pursuit of intrinsic motivations.

Indeed, a constant characteristic of social enterprises is more wage moderation and lower wage differentiation within the organisation than in for-profit and public organisations. Moreover, their workers seem to be satisfied with their jobs and less prone to leave the organisation. As a consequence, the worker satisfaction and on-the-job effort crucial for the production of relational services can be achieved in conjunction with cost containment. This point is an important one because wage moderation in social enterprises is sometimes mistakenly taken as an indicator of work exploitation, rather than as an aspect of a peculiar remuneration package. The requirement that the agents who establish and run social enterprises should have a stake in their activities creates trust relations, and strengthens and gives credibility to the goals of the enterprise. Indeed, the opportunistic risk related to low-powered incentives can be at least partially averted if the objectives of the organisation and those possessing control rights coincide to some extent.

In general, the higher degree of freedom for workers to determine the content of the work awarded through low-powered incentives can be used to achieve personal objectives different from monetary profit (although not necessarily coincident with the institutional objectives of the organisation). However, the direct involvement of stakeholders most interested in achieving the social goals of the organisation greatly reduces this risk. Cohesiveness within the group carrying out the activity also increases the degree of awareness of, and control over, the effort exerted by fellow workers. This is especially important when goals have multiple dimensions.

Not much needs to be said about limited profit distribution, which is to be interpreted in relation to the elements discussed above. We have already noted that the existence of social goals is hardly compatible with profit maximisation. The role of a constraint on profit distribution should therefore be only marginal. In social enterprises, the limit on the distribution of profits originates as an institutional choice taken within the organisation, rather than as a characteristic imposed exogenously (by law or by the policy-maker). Obviously, monetary variables are still important when running the organisation, for example in guaranteeing its financial strength and its survival in the market (since these organisations face a significant level of economic risk). However, in this case the monetary dimension acts as a constraint (e.g. a certain return must be achieved in order to reward the capital invested) to the pursuit of the chosen objectives, instead of being a maxim in itself. Nevertheless, the limited profit distribution criterion is important for social enterprises, for two main reasons. Firstly, it is a powerful, although partial, signal for the stakeholders and the outside world, of the real objectives of the organisation. Secondly, it limits the possibility that other, profit-maximising organisations will label themselves social enterprises in order to obtain unwarranted fiscal and reputational advantages. Clearly, this also applies to social enterprises that initially pursue non-monetary goals but then shift towards more traditional profit-oriented behaviour.

Finally, democratic and open management is another major characteristic of social enterprises, and it is consistent with the features discussed above. It is consistent in particular with the need for the constant involvement of workers in affirming and defining the goals of the organisation, and also with the representation and participation of other stakeholders, such as clients and customers. Moreover, democratic management is often correlated with a wage schedule that is relatively flatter than in for-profit organisations. A democratic structure does not sharply differentiate between hierarchical levels when it comes to deciding monetary remuneration. This also seems to correspond with the lower importance of monetary dimensions in social enterprises, and is consistent with the use of low-powered contracts. If contractual relations are necessarily loosely defined, democratic and open management has the important role of redirecting strategies so that they are consistent with organisational goals.

The foregoing analysis supports our initial claim that there is a link between the characteristics of social enterprises and the goods that they produce. Social enterprises emerge as institutional arrangements, which seek to reduce the
difference between the goals of the organisation as such, and those of the agents that take part in production. Insofar as they limit opportunistic behaviour, they are successful in reducing the costs of ownership as such other organisational forms.

Clearly, even if agency problems and opportunistic behaviour may be reduced in social enterprises, they do not disappear completely. Moreover, social enterprises offer an organisational model that is at once very fragile and very unstable. One of the main difficulties faced by social enterprises is achieving a constant balance among organisational goals, the allocation of control rights, the incentive structure, and the characteristics of the markets in which they operate. This is a daunting task for these organisations, especially because there seems to be a lack of points of reference in managerial theory and experience, and because existing incorporation forms significantly limit the stabilisation of their governance structure.

Moreover, social enterprises are extremely sensitive to changes in public policy, especially regarding the types of services eligible for public subsidies. Understandably, social enterprises can be negatively affected by a decrease in public support for the production of social services. But this is not the only example of fragility. Even the decision of public authorities to finance the production of a service previously provided using private resources could require delicate institutional changes, and thus constitute a risk. If the public authority decides to take up the redistributive aspect of a given service, this implies not only contractual change for social enterprises, but also radical change in organisational arrangements, the allocation of control rights, and the incentive structure.

Although these weaknesses are a major obstacle to the development of a social enterprise production model, the studies in this book show that social enterprises may be better able to organise the efficient production of particular goods than are for-profit enterprises, public agencies, and traditional non-profit organisations.

**Conclusions**

In this chapter we proposed an institutional interpretation of social enterprises as a peculiar incentive system. Social enterprises constitute an organisational form that is able to perform well in the production of personal and collective services that cannot be efficiently produced by for-profit or public organisations for two main reasons. Firstly, these services may involve market and contractual failures (often related to the relational component of the service). Secondly, a certain degree of redistribution from financiers or producers to consumers may be required for production to take place.

The institutional characteristics of social enterprises form an original incentive system that helps to overcome the problems associated with production of personal and collective services. Particularly relevant, from this point of view, are the explicit social aim, the proximity between producers on the one hand and consumers and the local community on the other, and an internal organisation that values open and democratic management over monetary remuneration for workers' effort.

Social enterprises strike a difficult balance between organisational goals, allocation of control rights, incentive structure and characteristics of the markets in which they operate. While these elements allow social enterprises to adapt their internal organisation to the specific problems posed by the production of personal and collective services, they also make the social enterprise a fragile organisational model, extremely sensitive to changes in market conditions and public policies.

**Notes**

1 As a result, the relative weight of the third sector has grown steadily over the past two decades in most Western economies. According to Salamon and Anheier (1994), the third sector accounted for 12.8 per cent of the new jobs created in the United States between 1980 and 1990 (its relative weight in the economy being 6.8 per cent in 1990), 11 per cent in Germany (3.7 per cent) and 15.8 per cent in France (4.2 per cent). Borzaga (1991) estimates the growth of employment in the non-profit sector in Italy at 39 per cent during the 1980s, while total employment grew by 7.6 per cent. Only Employment growth continued at a sustained rate in the period 1990–1995 (Salamon and Anheier 1997). The national chapters in this book testify to the development of numerous traditional and new third-sector organisations.

2 As the chapters of this book show, third-sector organisations are virtually non-existent in some industries. 'They are concentrated in the production of what could, broadly speaking, be called collective services' (the term used by Ben-Ner and van Hoomissen 1991) or 'communal services' (OECD definition). In this chapter, we will sometimes use the term 'social services'. When third-sector organisations operate in the agricultural and manufacturing sectors, the core activity is almost invariably service provision, i.e., the creation of employment opportunities for disadvantaged workers.

3 For example, some social enterprises are incorporated as co-operatives, and are therefore allowed to distribute some of their profits.

4 Starting with Coase (1937).

5 Furthermore, the nature of market contracting implies the replication of these costs for repeated transactions.

6 Each of them, in turn, enables further organisational options.

7 The seminal studies are those by Grossman and Hart (1986) and Hart and Moore (1990).

8 Whether bounded rationality is necessary for a theory of incomplete contracts is a matter of debate. It is sometimes argued that there may be a limited degree of bounded rationality embedded in the fact that third parties are unable to verify some clauses of written contracts (Hart 1990). However, the incomplete contracting approach does not generally assume bounded rationality on the part of the contracting agents. The interpretation of unbounded rationality in incomplete contracting has been recently criticised, most notably by Maskin and Tirole (1999).

9 For a more comprehensive survey of the literature, see Anheier and Ben-Ner (1997).

10 See also Ben-Ner and van Hoomissen (1991).

11 Similar concepts can be found in, e.g., Alchian and Demsetz (1972), Fama and Jensen (1983) and Ben-Ner and Jones (1995).

12 The terms 'control rights' and 'effective control' are taken from Burkart, Grönb and Panunzi (1997). A similar idea is put forward by Aghion and Tirole (1997), who distinguish between 'formal' and 'real' authority.
...


