

The Legal Nature of the Firm and the Myth of the Firm–Market Hybrid

GEOFFREY M. HODGSON

ABSTRACT *A sharp conceptual distinction used to be drawn between the firm and the market. However, since the 1970s, many economists and sociologists have argued that the boundaries of the firm are indistinct. Ideas emerged of ‘internal markets’ within firms, of the ‘quasifirm’, of ‘hybrid firms’ and of firms as ‘quasi-markets’. This article re-examines the formal, legal conception of the firm. It is argued that there is no good reason to abandon this conception, even in the light of relational contracting, networking, subcontracting and other developments. To avoid confusion, additional terms such as ‘supplier network’ or ‘conglomerate’ should be used to describe these phenomena, rather than abandoning a legally-based definition of the firm. With this clarified definition, ‘internal markets’ and ‘hybrid firms’ disappear from view.*

Key words: Firm; Market; Relational exchange; Internal markets; Hybrids.

JEL classifications: K0, L2, M2.

1. Introduction

This paper argues that the notion of the firm in economics and elsewhere has become dangerously devalued of legal meaning and much institutional substance.¹ It is shown that prominent earlier economists recognised the legal structure of the firm and saw it as important for economic analysis. Subsequently, however, for several reasons, economists lost sight of the essential legal element in the definition of the firm. These legal aspects were neglected when the organizational structure and essential nature of the firm became once again important areas of analysis in the 1970s. A result of this omission was that economists remained unable to agree upon a definition of the firm. Another outcome was the appearance of a notion that there

The author is very grateful to Ana Castro, Simon Deakin, Stephen Dunn, Nicolai Foss, Bart Nooteboom, Susumu Ono, an anonymous referee, and others for their helpful comments on earlier versions of this essay.

Geoffrey M. Hodgson, The Business School, University of Hertfordshire, Mangrove Road, Hertford, Hertfordshire SG13 8QF, UK. Address for correspondence: Malting House, 1 Burton End, West Wickham, Cambridgeshire CB1 6SD, UK; e-mail: g.m.hodgson@herts.ac.uk.

is no clear or essential boundary between the firm and the market. It became popular for both economists and sociologists to argue that the boundaries of the firm were fuzzy and indistinct. Ideas emerged of 'internal markets' within firms (Doeringer and Piore, 1971), of the 'quasifirm' (Eccles, 1981), of 'hybrid firms' or 'hybrid forms' (Cheung, 1983; Williamson, 1985; Ménard, 1995, 1996) and of 'quasi-markets' (Ménard, 1995). As Richard Langlois (1995: 72) has noted: 'much of transaction-cost economics has reached the conclusion that the distinction between firm and market is little more than semantics.' Similarly, most other approaches to the theory of the firm have failed to establish a clear boundary between the firm and the market, or to proclaim a clear definition of the firm.

It is argued in this article that these conceptual developments are largely misconceived and that they result from the lack of a clear, legally-grounded definition of the firm. In particular, the idea of 'markets' within firms is at best a misleading metaphor. Likewise, the popular idea of the pervasive firm-market hybrid is sustainable only if the firm is defined as something different from the legal entity of the firm in the real world. The theoretical value of this redefinition is questioned. Furthermore, the idea that there is no boundary between the firm and the market impairs more adequate definitions of the firm and analyses of its nature.

Evasions over the essential nature and definition of the firm are not confined to transaction cost theory. As Bengt Holmström (1999: 100) notes: 'property rights theory, as articulated in Hart and Moore (1990) and other representative pieces, says very little about the firm. The problem is that there are really no firms in these models, just representative entrepreneurs'

The lack of a consensus concerning the definition of the firm in the literature is a serious problem and there is a widespread reluctance to use legal forms as a basis for defining the firm. It is likely that this reluctance stems in part from a misguided enthusiasm to frame the analysis of the firm and the market in universal, ahistorical and relatively de-institutionalised terms (Hodgson, 2001). Accordingly, the criticism of the idea of the firm-market hybrid leads to other penetrating questions concerning the theoretical analysis of economic institutions.

The central argument in this paper is that there is no good reason for economists to relinquish a legally-grounded definition of the firm or corporation. Legal relations are an essential part of the definition, alongside other necessary factors. This is partly because the legal specifications and frameworks have important economic consequences. This argument is not undermined, but strengthened, by the real growth of other economically significant entities such as business units, conglomerates, *keiretsu*, strategic alliances, supplier networks, relational contracting and so on.² In fact, the growth of a diversity of business and industrial structures makes it imperative to develop clear, distinct definitions of the different entities involved. Making one term cover all these diverse and changing frameworks is a recipe for analytical and theoretical confusion. Consequently, a legally-grounded definition of the firm should be retained, alongside the specification of other definitions for relevant structures that are not firms.

Section 2 of this article discusses the history and basis of legal definitions of the firm and shows that some leading economists acknowledged these legal realities. Section 3 shows how such legal definitions were eroded and abandoned, along with attempts to provide an alternative definition of the firm. Section 4 discusses some additional literature that has implications for the boundary between the firm and the market. These sections pave the way for the central arguments in the paper: in

section 5 that contains a critique of the idea of an internal market within a firm, and in section 6 that criticises the notion of a firm-market hybrid. Section 7 considers some rare economists who have attempted to bring some order to the confusion, and acknowledge the legal reality of the firm. Section 8 concludes the essay and provides some definitions of the firm and the market.

2. The Legal and Historical Reality

In economics, the ‘firm’, the ‘corporation’ and the ‘company’ are used as virtual synonyms. The modern laws of contract and trade have precedents in the Roman Empire. Under Roman Law it was possible for a merchant or producer to own capital goods, hire workers, buy slaves, and engage in trade. These institutions might be regarded as firms, but the real origins of the terms ‘company’ and ‘corporation’ are medieval. In different countries, firms evolved in different ways. As Avner Greif (1996) explains, the ‘company’ originally developed in medieval Italy when several individuals agglomerated their capital by establishing a permanent partnership with unlimited and joint liability. Each partner invested some capital in the company and took a share of any profits in proportion to his investment. These arrangements were originally between members of the same family, but were later extended to cover non-family partners (*compagnia*). The earliest post-Roman record of such a legal partnership was that of a group of merchants in Genoa in the twelfth century.

In medieval England – in contrast to the Italian *compagnia* – the ‘corporation’ emerged as a term to describe organizations that received a royal charter from the state, involving special privileges and immunities. Typical of such ‘corporations’ were town boroughs, guilds, church institutions, universities or charities. But most of these were not business institutions and they themselves did not evolve into modern business corporations (Hessen, 1979, 1987). However, the legal terms ‘corporation’ and ‘act of incorporation’ date from this time, with the connotation of a grant of rights from the state to a relatively autonomous organization. Elements of this legal framework of ‘incorporation’ were later applied to business firms.

The institutional ancestors of the modern business corporation were the joint-stock companies that flourished in England from the sixteenth century. Under the control of their directors, they raised capital and sold shares. Until the Glorious Revolution of 1688, they had to obtain a royal charter of incorporation, because the freedoms of commerce and travel were highly limited. Subsequently, legal restrictions were often circumvented by the use of mechanisms such as the legal partnership or legal trust. Legally chartered corporations continued to exist alongside unincorporated firms that often adopted a similar organizational structure (Du Bois, 1938).

In Britain, during the Industrial Revolution, the advantages of corporate organisation in raising and investing large masses of capital became apparent. Some restrictive legislation was repealed, and the Company’s Act of 1844 made explicit in law the distinction between a partnership and a joint-stock company. Partners are mutual agents, each able to establish contracts with others that are binding on all partners. In contrast, shareholders are not mutual agents, they cannot themselves act in the name of the firm, and shares in a company are freely transferable. Limited liability became established as a general and legal right of a legal corporation in 1855 (Du Bois, 1938; Hunt, 1935a, 1935b).

In America, the Connecticut legislation of 1837 established a simple registration procedure for the creation of a legal corporation. This legislation was quickly copied

by the other American states (Handlin and Handlin, 1945; Hessen, 1979, 1987). However, corporate legislation became so permissive in the US that in the early twentieth century it was widely complained that the state and federal authorities had abdicated their original responsibility to protect the public from abuses of corporate power (Berle and Means, 1930; Abbott, 1936). Consequently, the legal framework of corporate activity has been an ongoing matter for legislation throughout the twentieth century.

The concept of the firm as a 'legal fiction' was thus established in law. By a legal fiction it was meant that a specified organization would be treated in law as a single individual. The 'fictional' nature of this construct was that contractual considerations that applied to individuals would also apply to partnerships or incorporated firms. It did not mean that the application itself was imaginary or false. On the contrary, the law had to establish analogous tests of contractual offer and agreement to apply to firms as well as individuals.

Despite the periodic circumventions and liberalizations of corporate law in the past, the law itself has never entirely disappeared from the picture. The thrust of nineteenth century corporate law reform in Britain, America and elsewhere was not only to facilitate the formation and operation of business corporations but also to provide them with the right and protection of limited liability. The liberalization of corporate law did not mean that the state withdrew from the scene. On the contrary, the state provided legal rights and protections for corporate activity, necessarily involving a formal mechanism for the legal recognition of the corporation itself.

The word 'firm' derives from the Latin adjective *firmus*, meaning strong, powerful, durable and lasting. As a noun, the word went on to acquire the significant meaning of (legally binding) 'signature', and with this important connotation it survives today in some form in several Romance languages. It is reasonable to use the broad term 'firm' to include a both partnership and a corporation, as well as a single person employing others in production or trade. This is in accord with the *Concise Oxford English Dictionary*, which defines the firm rather loosely as 'partners carrying on business; group of persons working together'. However, even with this broad description, the firm still remains an organizational and legally circumscribed entity. Given that the etymology of 'the firm' connotes a legal entity, and a broadly defined but singular legal person, it is reasonable to regard partnerships, cooperatives, companies and corporations all as firms. But, of course, these different types of firms all have important differences of structure, rationale and possible behaviour. Nevertheless, they can all be placed within a single taxonomic class. Within these broad and inclusive terms, a more precise definition of the firm will be attempted later below.

Leading economists such as Adam Smith and Alfred Marshall were acutely aware of the importance of the legal framework in which the firm operated. In the 1784 edition of the *Wealth of Nations*, Smith inserted a new section where he compared the corporations and the minimally regulated joint-stock companies of his time, and criticized monopolistic and other adverse practices (Smith, 1976: 741–7; Anderson and Tollinson, 1982). Clearly, Smith was sensitive to the impact of legal arrangements on the nature and practices of the firm. For Smith and other economists, the firm was a legal entity, whose boundaries were prescribed in law. Accordingly, in his *Industry and Trade*, Marshall (1919) discussed the legal structure of the joint-stock company and argued that nineteenth century legislation had facilitated its capacity to obtain investment capital. Early American institutional economists, such as John Commons (1924) also discussed the legal foundations of

the firm. Until quite recently it was taken for granted by economists that the firm was a legal entity, even if the legal realities and structures were not discussed as much as they were by Smith, Marshall or Commons.

3. From Consensus to Confusion

As a result, economists in the early twentieth-century took it for granted that the firm and the market were different things. Dennis Robertson (1923: 85) referred to firms as ‘islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk’. However, at this time, the distinction between the firm and the market was more often tacitly sustained in the elaboration of economic theory than by explicit reference to legal structures. The distinction between the firm and the market was sustained loosely by the distinction between production and the market. The former was represented by production and cost functions, the latter as the sphere where supply interacted with demand.

After Marshall’s death there was a growing effort to analyse economic structures and activities in terms of mathematical functions. Hence the tendency was to treat the firm simply as a production function. The ‘theory of the firm’, as evidenced by Fritz Machlup’s (1967) influential review on the topic, became a theory of how production and cost functions interacted with demand on the market. As long as the firm was treated principally as a production function, many questions of its definition, identity, internal structure and external boundaries could be elided, even if a distinction between the market and the firm was maintained.

Subtly, the treatment of the firm as a production function changed the idea of the firm from a legal structure to a functional locus of production. The notion of the firm as a production function shifted the meaning of the term ‘firm’ towards a coarse functional description of a production process. From the 1920s to the 1970s, the original, legal conception of the firm was abandoned in a context where institutional structures were largely ignored.

The great exception, of course, was Ronald Coase. His 1937 article was one of the rare studies after Marshall to treat the firm as an organization, rather than merely a production function. Coase (1937: 388–90) demarcated the market from the firm, with the ‘price mechanism’ on the one hand and its ‘supersession’ on the other. For Coase the alternative to the firm was the coordination of self-employed individual producers by the market, each being his or her ‘own master’. Coase (1937) thus offered a very rare discussion of the internal structure of the firm.³

While most other micro-economic theorists turned the real world firm into a production function, others dealt with the empirical problem of classifying firms and measuring such phenomena as industrial concentration. These applied economists dealt with the statistical and taxonomic problems with the industrial data. A single firm could have several separate plants. Furthermore, a single factory might house quite different production processes. Considerations of statistical analysis could be at odds with the notion of the firm as a production function.

Consequently, some applied economists tried to establish a different definition. For example, Sargent Florence (1957: 244) argued: ‘To economists, more directly interested in degrees of competition and monopoly than in industrial location, the firm, the unit of control, is more important than the plant, the physical unit.’ Florence placed more emphasis on the ‘unit of control’ than the legally defined unit of the firm. The problem was how to deal statistically with subsidiary companies

that legally were separate firms but were managed and controlled by their parent companies, as if parent and subsidiaries were a single unit.

Florence approved of the practice of the UK Census of Production, who used the term ‘business unit’ to refer to a parent company together with any subsidiary companies of which it owned more than 50 percent. Hence a ‘business unit’ was not the same as ‘a single company in the legal sense’. Neither would it include subsidiaries in which the interest of the parent company was in practice sufficient for managerial control, but did not amount to as much as 50 per cent. Florence (1957: 245 n.b.) lamented the lack of discussion of this problem, in the context of several studies ‘comparing degrees of concentration in British and American industries without troubling about exactly what it is that is concentrating.’ But above all he made it clear that: ‘Neither country actually uses the word “firm” in their official statistics on which these economists base their calculations. America uses “producers,” Britain “business unit.”’ (ibid.)

Hence Florence made a case for an additional term, ‘producer’ or ‘business unit’, to be used in empirical studies of the concentration of corporate power in an industry. Note, however, that he did not describe these units as firms, or promote the Census of Production definition as an alternative definition of the firm. Florence did not abandon the legal conception of the firm. He simply approved of the additional concept of ‘business unit’ to be used alongside it.

However, Machlup adopted a very different strategy in his seminal article. Machlup (1967: 26) listed no less than ten ‘concepts of the firm employed in the literature of business and economics’ and suggested that even more were in use. He argued that the firm had been regarded as an organization, a decision-making system, a collection of assets and liabilities, a juridical person, a business unit under a single management, and much more besides. Machlup (1967: 28) then concluded:

This exercise should have succeeded in showing how ludicrous the effort of some writers are to attempt *one* definition of the firm in economic analysis . . . I hope that there will be no argument about which concept of the firm is the most important or the most useful. Since they serve different purposes, such an argument would be pointless.

But this clearly involves a *non sequitur*. The fact that several different conceptions of the firm exist does not imply that the formulation and promotion of one definition should be abandoned. On the contrary, it could be argued that the very confusion itself calls out for a single, commonly accepted definition. Machlup rightly pointed out that these different concepts ‘serve different purposes’. But it did not occur to him that these different definitions might also point to different things. A (legal) firm is not the same thing as a business unit, or a collection of assets, or a production plant. These are not simply different *concepts*: they are also different real *entities*. Accordingly, we require a plurality of concepts to refer to a plurality of real structures. Contrary to Machlup, the existence of multiple definitions of the firm cannot be tolerated simply because they ‘serve different purposes’. Definitions are not merely convenient instruments of thought but attempts to carve reality at the joints so that it can be understood and explained. In reality we face several different types of entity and we require several different concepts to differentiate them.

Such confusion was prevalent when Oliver Williamson revived transaction cost analysis in his *Markets and Hierarchies* (1975). On the positive side, Williamson

inherited a Coasean emphasis on legal and other institutional structures. He has taken seriously the differences in legal framework that govern transactions and greatly expanded the disciplinary interface between economics and law. But these efforts were not enough to overcome the prevailing conceptual bewilderment. Williamson himself prioritized the vague term 'hierarchy' instead of any attempt to define the firm.⁴ Sadly, as the 'black box' of the firm was opened, there was not sufficient agreement on the terms to describe the box or its contents. Unfortunately, matters have not improved since. The definitional imprecision in the literature has been ill-equipped to deal with the explosion of interest in the study of the firm as an institution that has occurred since 1975.

The central argument here is that no good reason has ever been presented to abandon the original, legal concept of the firm. At the same time, however, it is vital to use other concepts such as business unit, production plant or conglomerate, to describe other important structures in the business world. Multiple concepts are required to describe multiple real items. In principle, an argument could be made that one such entity was more important in economic terms, and hence some concepts were more useful than others. But no one has ever made such an argument. Instead it is repeatedly proclaimed that all attempts to define 'the firm' are fruitless. This licensed imprecision promotes habits of conceptual vagueness and terminological sloppiness that have prevailed in the business and economics literature ever since.

The outcome of this confusion is that a clear, single boundary between 'the firm' and the market cannot be established. It has become commonplace to use the concept of 'the firm' to refer to a multiplicity of different structures, so it is not surprising that no single boundary can be discerned. Undefined symbols and inferior conceptual instruments mean that the original boundary between the firm and the market – that so many past economists had taken for granted – lies undetected. The definitional maps have been degraded. This process is discussed in more detail in sections 5 and 6 below.

4. The 'Inside' and the 'Outside' of the Firm

It is convenient at this stage to discuss some additions to the literature in the 1970s and 1980s that in some way addressed the boundary or relationship between the firm and the market. Exceptionally, in their famous article on the theory of the firm, Michael Jensen and William Meckling (1976) argued that the legal aspect of the firm was of some importance in determining the behaviour of individuals in the economy. Jensen and Meckling (1976: 311 n.) emphasized

the important role which the legal system and the law play in social organizations, especially, the organization of economic activity. Statutory laws sets bounds on the kind of contracts into which individuals and organizations may enter without risking criminal prosecution. The police powers of the state are available and used to enforce performance of contracts or to enforce the collection of damages for non-performance.

This was a rare acknowledgement of the economic and real significance of the legal aspect of the firm. However, Jensen and Meckling misused the legal specification of the firm as a 'legal fiction' to draw an inappropriate conclusion. They suggested that the 'legal fiction' that allows some organizations to be treated legally in some

respects as individuals meant that firms in fact *were* individuals. They thus concluded:

Viewed this way, it makes little or no sense to try to distinguish those things which are ‘inside’ the firm (or any other organization) from those things that are ‘outside’ of it. There is in a very real sense only a multitude of complex relationships (i.e., contracts) between the legal fiction (the firm) and the owners of labor, material and capital inputs and the consumers of output. (Jensen and Meckling, 1976: 311)

The irony here is that a rare acknowledgement of the legal nature of the firm was used in an attempt to close down any discussion of its internal organizational structure. Instead, the firm was treated as a personal singularity. But their argument is invalid. The fact that the law treats the firm in some respects as a person does not mean that the firm is a person. The law imputes some individual capacities of contractual agreement, some individual rights of ownership, and some individual obligations to organizations. But it does not treat the organization and the person as identical. For example, limited liability does not apply to individuals in the same way as it does to corporations. Bankrupt organizations and bankrupt persons are also treated differently under the law. Furthermore, the law does not ignore the organizational character of the firm. Legislation has been developed to cover many relationships inside the firm, including the conduct of existing and agreed employment contracts, the rights and duties of employers and employees, and so on (Masten, 1991). Reference to the legal aspect of the firm does not give licence to treat the firm simply as an individual person, even if the firm is an individual organization. Even in legal terms, the firm is a structured set of relations between persons.⁵

In what sense, with due deference to the legal reality, can one identify the ‘inside’ or the ‘outside’ of the firm? The boundary of the firm is made up of points in time at which the ‘legal person’ of the firm concludes legal (written or unwritten) contracts with individuals or other legal persons. When an employee is hired or an asset becomes the property of the firm, then those persons or assets cross the boundary to the inside of the firm. However, an employee (unlike a slave) is only ‘inside’ the firm during the periods, and under the circumstances, specified in the employment contract. Hence on the inside of the firm we find the firm’s owned assets, plus anything that is hired and within its period of hire, including employees during their contracted time of employment.

What lies ‘outside’ the firm? Up to now we have discussed the boundary of the firm as if it is a boundary between the firm and the market. This is a chronic oversimplification. It is important to rectify matters at this stage.

By definition, markets involve multiple exchanges, multiple buyers and multiple sellers, and thereby a degree of competition. *A market is an institution in which a significant number of commodities of a particular, reasonably well-defined type are regularly exchanged.* Exchanges themselves involve property rights and take place in a framework of law (Commons, 1924; Sened, 1997). Hence the market also has also a legal aspect. Markets contain legal and other rules and structures that pattern these exchange negotiations and transactions. Markets, where they exist, help to structure, organize and legitimate exchange transactions. They involve pricing and trading routines that help to establish a consensus over prices, and often help by communicating information regarding products, prices, quantities, potential buyers

or possible sellers. Markets, in short, are organized and institutionalized exchange (Hodgson, 1988).⁶

In three classic articles, George Richardson (1972), Victor Goldberg (1980) and Ronald Dore (1983) argued that the relationship between a large corporation and its subcontractors is often more durable and intensive than a typical market relationship. They referred to the reality of ‘relational contracting’ in which firms develop longstanding ties with other contractors. In some accounts, these relational contracts are infused with considerations of mutual understanding and trust that are not typical of open market trading. Instead of resorting to the market, the firm will often rely on a trusted supplier, especially when their relationship involves shared knowledge and mutual, interactive learning.

As Douglass North (1977: 710) observed: ‘most exchanges do not take place in markets.’ In all exchanges there is an implicit or explicit legal contract, leading to the transfer of property rights. But not all exchanges involve the type of competitive and transitory relationships that we find on markets (Hodgson, 1988). As a result, instead of just firms and markets, there are three main possibilities: the firm, market exchange and non-market exchange. Markets are but one possible type of commodity exchange ‘outside’ the firm.

As shown below, some theorists ignore the possibility of relational exchange. A possible confusion then arises. When they observe enduring contractual relations of the type discussed by Richardson, Goldberg and Dore they are then faced with a problem. The real-world ensemble of such interactive relations is neither a firm nor a market so – according to the false dichotomy between firms and markets – it must assume the ‘strange’ form of a ‘hybrid’ or a ‘quasi-market’. The error here lies in the adoption of a false dichotomy, ignoring the third possibility of non-market contractual exchange.

Other economists regard the existence of relational exchange as grounds for abandoning the distinction between the firm and the market. But the existence of a third possibility is no excuse for fudging the distinction. Indeed, it points not to two but to at least three types of institutional relationship. An enduring relationship between a dominant firm and a subcontractor is not an open market relationship, but it is still one of commodity exchange, involving the legal transfer of property rights, between two distinct firms. It is not evidence of commodity exchange, nor evidence of a ‘market’ within a single, encompassing firm. Firms, markets, and relational exchanges are different things.

5. The Myth of the Internal Market

With the exception of some slight ambiguities, Coase (1937) drew a clear distinction between the firm and the market.⁷ Benjamin Klein (1983: 373) later commented on Coase’s early views:

Coase mistakenly made a sharp distinction between intrafirm and interfirm transactions, claiming that while the latter represented market contracts the former represented planned direction. Economists now recognize that such a sharp distinction does not exist and it is useful to consider also transactions occurring within the firm as representing market (contractual) relationships. The question what is the essential characteristic of the firm now appears to be unimportant.

Despite the citation of the authority of unnamed ‘economists’, researchers will search in vain in the literature of the social sciences for any previous, adequate and scholarly argument against the ‘sharp distinction’ between the firm and the market. There is no explanation in the literature why it would be ‘useful’ to consider transactions within the firm as ‘market’ relationships. No scholarly argument will be found for abandoning the question of ‘the essential characteristic of the firm’. The danger is that once we believe that the firm has no clear boundaries or identity then it becomes a conceptual sponge ready to soak up anything put in contact with it.

Yet Klein’s view proved to be contagious. A few years later, Coase (1988: 27) himself declared a revised view. He wrote: ‘I have come across numerous examples of markets found *within* firms, but one which amused me was the discovery of a kind of market operating in the heart of a nationalized industry in England, the electricity supply industry.’ Coase then quoted from a 1961 lecture given by an official of the Central Electricity Generating Board: ‘the National Control Room becomes in effect an auction room, with a National Control Engineer asking the Regional Centres to quote the price at which they could supply a certain number of kilowatts at specified periods during the following day. . . . Wherever possible he accepts the lowest bid’. Coase continued: ‘An analogous situation may, of course, be found within a privately owned firm in which separate departments or divisions may supply one another as a result . . . of what are essentially market transactions between them.’

Once we apply clear legal and other criteria, however, Coase’s argument begins to disintegrate. The UK Central Electricity Generating Board (CEGB) was itself a singular legal person. The Regional Centres were not legally separate firms but internal divisions of the CEGB. The Regional Centres neither owned nor sold electricity to the CEGB. Any semblance of contract between the Regional Centres and the CEGB would not have been recognized in law as such. It was not an exchange of property rights because ownership of the electricity remained in the hands of the CEGB throughout. Instead, the bidding and ‘selling’ of electricity was an internal management mechanism to reduce costs and encourage increases in productivity.

It is true that there are often internal negotiations and transfers of resources between divisions of the modern firm. These divisions may have their own accounts and profit targets. Most firms use price indicators for internal accounting. But are there ‘internal markets’ within firms? Again a key test is whether or not these divisions have separate legal status, and are recognized as ‘legal persons’. Internal transfers within the firm do not involve the exchange of legal property rights. The objects of ‘exchange’ remain the property of the firm. These ‘exchanges’ are not legally enforceable contracts of trade: they are internal transfers. If a division of the firm is delegated the power to enter into contracts with outside bodies, then the firm as a whole is legally the party to the contract. The division is merely exercising delegated powers: it acts ‘in the name’ of the corporation, and the corporation as a whole is legally responsible for its liabilities under the agreed contract. Because the firm is a singular legal entity, trading and contract *within* a firm are highly limited.

Accordingly, attempts to simulate markets within firms raise all the questions and difficulties concerning the establishment of pseudo-markets within a centrally planned economy. As critics such as Ludwig von Mises and Friedrich Hayek have pointed out, these are not true markets because they do not devolve to local managers, *without the possibility of central interference*, the right to set prices and make

allocative decisions, at least within some limits (Hayek, 1935). Just as true markets could not exist in an economy that was entirely owned by and planned from the centre, similarly true markets cannot exist within firms.

Of course, there are examples of modern conglomerates made up of legally separate units. Typically, multi-national enterprises will exist as separate legal bodies in different countries, all owned by the headquarters corporation. These are international conglomerates. That means that we must have several terms, not one: 'firm', 'conglomerate' and 'multinational conglomerate'. It is important not to confuse these, as they refer to different structures. The response may be that these arguments are merely scholastic or legalistic, and the 'real' situation is quite different from these legal 'formalities'. However, if this were the case then the onus is on economists and others to provide concepts and criteria to discern the 'true' reality 'behind' the legal 'forms'. But it seems, according to Klein at least, that this very task has been abandoned.

Of course, the formal legal status of any organization tells us no more than a small part of the whole story. Furthermore, legal formalities can sometimes have an almost fictional status, masking a different reality. For example, a conglomerate of different firms may in practice act like a single firm, because control of the conglomerate is concentrated in the hands of a single group. But in both law and economic reality, the conglomerate is not the same as the firm. They are different structures, with the implication, for example, that the dissolution of a conglomerate is not the same kind of processes as the disintegration of a single firm into separate legal units.

Consider the widely adopted suggestion that 'internal labour markets' exist inside the firm. Employment contracts may be renegotiated, for example. However, these would be cases of the renegotiation of contracts for *inputs* of goods or services *into* the firm. Even if a new employment position is advertised by using the firm's internal communication network, this is an invitation for existing employees to apply for a position which would involve them terminating aspects of their existing contract and re-entering the firm under revised contractual conditions. It would be neither wholly internal nor a market proper. Even the pioneers of the idea, Peter Doeringer and Michael Piore (1971: 1–2) admitted that 'internal labor markets' are not governed primarily by the price mechanism but by 'a set of administrative rules and procedures'. David Marsden (1986: 162) went further: 'internal labour markets offer quite different transaction arrangements, and there is some doubt as to whether they fulfil the role of markets'. What Doeringer and Piore pointed to was a degree of competitive fluidity of labour within the organization. They did not show that a true market existed within the firm.

Would a competition between existing employees for an advertised post within the firm be an 'internal labour market'? Not strictly, because this would not constitute regularized and organized exchange but a series of one-off competitions for advertised internal posts. These would be periodic instances of contract reallocation, rather than a true labour market. The regular, repeated and institutionalized exchange that characterizes true markets would be lacking.

As another example, work on the modern multi-divisional form suggests that each division within a firm competes in the firm's 'internal capital market' for the budget allocated by the firm's head office (Williamson, 1975). To be sure, some substantial competition and rivalry is involved, but it is not competition on a market. It is completely different from the true capital market where shares in individual firms are bought and sold. The division of the firm has no independently

owned share capital to use as a lever to obtain its own loans from outside agencies. The internal competition to which Williamson refers is not market competition but a struggle over power and resources between different parts of the corporate bureaucracy.

Unfortunately, a sloppy use of the term 'market' pervades economics today. In terms of genuine, regular and organized exchanges of goods or services, 'markets' are rarely, if ever, found *within* the firm. This pervasive confusion over the nature of markets and exchange allows free-market and other economists to ignore the reality of non-market organization in capitalist firms and to understand everything in 'market' terms.

6. The Myth of the Firm-Market Hybrid

Just as the misleading idea of 'internal labour markets' was developed by sociologists and uncritically appropriated by economists, a similar story applies to 'hybrid' concepts such as the 'quasifirm'. For example, Robert Eccles (1981: 339–40) considered cases in the construction industry where

relations between the general contractor and his subcontractors are stable and continuous over fairly long periods of time and only infrequently established through competitive bidding. This type of 'quasi-integration' results in what I call the 'quasifirm'.

Notice that Eccles ignores the possibility that these 'stable and continuous' contracting relationships may be examples of relational exchange along the lines of Richardson, Goldberg and Dore. In fact, by inventing his version of the dubious concept of the quasifirm, Eccles makes the concept of relational exchange redundant. Eccles (1981: 342–3) then went on to consider the divisions within the M-form firm:

These autonomous divisions function in many ways like independent firms. Often they compete successfully with firms in the market for business with each other. Thus internal transactions in these firms can have similar characteristics to market transactions. In contrast, the inside contracting system is a set of market transactions with similar characteristics to hierarchical transactions. Most generally, pure markets and pure hierarchies are at opposite ends of a continuum of contracting modes.

It may be true in a sense that component enterprises 'compete' with each other, or with other firms, for business on the market. Consider a firm *A*, with subdivisions *A1* and *A2*, that 'compete' for business on the open market. Assume that *A1* successfully negotiates a deal with another firm *B*. The crucial point is that the legal contract is not between *A1* and *B* but between *A* and *B*. Whatever they do, in legal terms both *A1* and *A2* act in the name of *A*. In particular, the present author is not aware of any real world examples of competition between separate divisions of a single firm for custom with another company. Such a strategy is not impossible but would seem to be largely self-defeating. It would seem to be against the interests of the firm as a whole to allow its divisions to beat each other down on price.

We may ask, if 'pure markets and pure hierarchies are at opposite ends of a continuum of contracting modes' then what lies in between the two extremes? My

answer would not be the quasifirm, but different degrees of relational exchange, including the possibility of networks. However, these intermediate categories are not some kind of taxonomic halfway house between markets and firms, but quite different types of non-market relationship among separate firms (Hage and Alter, 1997).⁸

Like many others, Steven Cheung (1983: 11) had some difficulty in discerning the boundaries of the firm. He discussed this example:

A landlord, who wants to build a high-rise finds a building contractor. This contractor subcontracts with a hardwood floor contractor on an agreed price per square foot – a piece count. The subcontractor, who imports the wood materials and adds finishing work to the wood on a piece-rate basis, in turn finds a sub-subcontractor, provides him wood, and offers him a price per square foot laid. Finally, the sub-subcontractor hires workers and again pays them per square foot laid.

Such a complex integration of contracts and subcontracts is very common in reality. But it offers no insurmountable taxonomic difficulty. With his repeated stress on payments per square foot, Cheung seemed to think that piece-rate payments imply the existence of ‘a “market”’ (p. 10). But there is no reason why one should imply the other. Cheung (p. 17) also claimed that ‘economists may well argue that because they are all vertically integrated by contracts, with transfer pricing, only one firm exists.’ However, contrary to Cheung, being ‘vertically integrated by contracts’ is not the same as vertical integration *within* a firm. Blind to this important distinction, Cheung suggested that the hardwood floor example is something with the characteristics of *both* ‘a “market”’ and ‘one firm.’ The myth of the firm-market hybrid was born.

In truth, the hardwood floor example involves (market or relational) exchange, between not one but four ‘legal persons’: the landlord, the contractor, the subcontractor, and the sub-subcontracted firm with its employees. By the legal criterion, four firms or legal persons were clearly present, not one. Considering another example, Cheung (1983: 16–17) wrote:

If an apple orchard owner contracts with a beekeeper to pollinate his fruits, is the result one firm or two firms? The question has no clear answer. The contract involved may be a hive-rental contract, a wage contract, a contract sharing the apple yield, or, in principle, some combination of these and still other arrangements. In each case the beekeeper receives a remuneration for his service, and the orders he expects from the orchard owner vary with the form of contract.

The confusion here is that Cheung addressed not one possible arrangement but several. The answer to the question ‘is there one firm or two?’ itself depends on whether the contract involves hive rental, wage labour, or sharecropping. In fact, Cheung (p. 17) admits this when he wrote: ‘Most economists would probably opt for only one firm if the beekeeper is hired on a wage contract but for two if the hives are rented.’ This is a reasonable answer to his own question. The problem here for Cheung was not that he lacks a good answer, but that he was reluctant to endorse it and to acknowledge a real difference between the types of contract that he listed.

In contrast, Cheung argued that the definition of the firm is arbitrary. He wrote: ‘according to one’s view a “firm” may be as small as a relationship between two input owners or, if the chain of contracts is allowed to spread, as big as the whole economy’ (p. 17). He then jumped to the conclusion: ‘Thus it is futile to press the issue of what is or is not a firm’ (p. 18). Cheung’s argument was essentially that: different definitions of X are possible, therefore it is futile to define X . But this is another *non sequitur*. The problem here is in the eye of the beholder. The taxonomic problem diminishes greatly once we define the firm as a legal and non-market entity.

Despite its manifest weaknesses, Cheung’s argument persuaded many. Against his earlier distinction between markets and hierarchies, Williamson (1985: 83) was ‘persuaded that transactions in the middle range are much more common’. Williamson (1999: 1091) saw ‘hybrids’ as ‘long term contractual relations into which security features have been crafted’. Admittedly, much modern contracting is of this ‘long term’ type. However, such ‘contractual relations’ are not themselves firms, nor within firms. Furthermore, the long term, relational and ‘security’ features of the contractual relations mean that they are not markets either. Market relations are generally more impersonal and are typically short term. But Williamson too is a victim of the false dichotomy. He assumed that if long-term contractual relations are neither firms nor markets, then they must be ‘hybrids’ of the market and the firm. One must ask, however: what agencies make the (long-term) contracts? In fact, these ‘contractual relations’ are all between, not within, firms. If we accept and understand the third (Richardsonian) possibility, of non-market contracts of a ‘relational’ type, then the whole picture is one of relational contracts between firms. There is now no ‘hybrid’ to be seen. By adopting the obfuscatory concept of a firm-market hybrid, Williamson muddled his former claim that the firm is essentially different from the market.

To his credit, Claude Ménard (1995) defined terms such as ‘market’ and ‘organization’ more precisely than most others. In particular, he defined the market as an institution. He rightly pointed to the possibility that markets themselves involve elements of organization and regulation. He thus argued: ‘In all of these situations, the market activities are significantly permeated with organizational factors’ (p. 176). So far so good. He then went off the rails. If some or all markets are organizations then this does not imply that all organizations can be markets. He tried to demonstrate that ‘*organizations can be internally structured as quasi-markets*’. Ménard considered franchising ‘when very strict standards are imposed on independent participants’. He noted that:

Classification [into markets or organizations] becomes particularly difficult when firms are interconnected by a dense web of transactions, with strong commitments to each other and complementarities of their assets, but without formal agreements and, moreover, with property rights on these firms clearly maintained as distinct. (1995: 176)

On this basis he acknowledged intermediate ‘forms’ between markets and hierarchies. These ‘hybrids’ involved ‘specific combinations of markets incentives and modalities of coordination involving some form of hierarchical relationship’ (p. 175).

The problems began when Ménard did not clearly distinguish the notions of ‘organization’, ‘firm’ and ‘hierarchy’. Indeed, he never provided an adequate

definition of the latter two terms, and he seemed to use the terms 'firm' and 'organization' as if they were interchangeable. He referred approvingly to the works of Cheung (1983), Williamson (1985) and others but did not rectify their vagueness over the definition of the firm. To compound the confusion, Ménard generally used the term 'hybrid form' instead of 'hybrid firm' but did not explain the difference between the two terms, nor did he establish a separate and adequate definition of the firm, as an integrated and legally-grounded entity.

Consider Ménard's example of strictly monitored and regulated franchising. With a broad definition of 'organization', this could be described as an 'organized' relationship, but between *two or more* firms or 'legal persons'. Although the relationship has an 'organized' character it does not mean that it is a single firm. The case when 'firms are interconnected by a dense web of transactions, with strong commitments to each other' is a case of relational contracting between multiple firms. Again, the 'organized' character of the relationship does not imply that they everything is organized within a single firm. The fact that 'property rights on these firms [are] clearly maintained as distinct' does not create any taxonomic difficulty. It simply underlines the fact that multiple firms may exist within a single organizational network. Essentially, Ménard's primary error was to allow confusion between the broadly defined organization and the firm. His secondary error was to understress the fact that relational contracting involves the exchange of commodities between *different* firms but *not* on an open market. It is thus a third option, after a market and a firm.

Also conflating the firm and the market, Ken-ichi Imai and Hiroyuki Itami (1984) discussed the supposed 'interpenetration of organization and market' in Japan. However, they defined both market and organization without any reference to property rights or contracts, referring instead to factors such as the degree of durability of the relationship and the use or otherwise of price as a major information signal. By this flawed methodology it is not difficult to find elements of so-called 'organization' in the highly structured and regulated 'markets' of Japan, and to find elements of an alleged 'market' inside many firms. These conclusions follow, however, from the inadequate definitions of 'market' and 'organization' in the first place. In contrast, superior definitions of these terms would lead to the conclusion that markets – in Japan and elsewhere – are often organized to a greater or lesser degree, but that any market is a quite different type of organization from the property-owning and contracting legal entity of the firm.

A similar defect is found among the economists of what might be termed (after their origin) the 'Warwick School', namely Keith Cowling, Christos Pitelis and Roger Sugden. They are so keen to emphasize an extra-legal notion of power and control that the legal and contracting realities are largely ignored. Accordingly, their work contains no clear definition of the market, largely ignores relational exchange, and makes no substantial distinction between employment and self-employment, or between employment and slavery. Observing that power is exercised in both firms and markets, the peculiar and inadequately defined notion of 'market hierarchies' (Pitelis, 1991) emerged to symbolize yet another conflation of the firm with the market. Here, however, instead of seeing everything through market lenses, the hierarchical power relations within organizations become eyeglasses to view the system as a whole. The conceptual focus becomes one of coordination and control. Again, the crucial error here is that legal and historical specificities – particularly concerning markets and contracts – are downplayed. As a result, instead of the prominent 'new institutionalist' mistake of treating institutions as if they were all

markets, the inverted error becomes one of treating all markets as if they were non-market organizations, and all contracts as if they were all simply matters of power and administrative control.

Accordingly, Cowling and Sugden (1998: 62) ‘question at the outset the focus on market versus non-market activity’ to ‘view the firm as a nexus of strategic decision-making’ (p. 67). The firm, thus defined, includes any subcontracted or subsidiary units. For Cowling and Sugden (p. 64) ‘the power to make strategic decisions is the power to plan the overall direction of production in the firm.’ They argue that

the crucial distinguishing feature of transactions within a modern corporation is that they are subject to strategic decision-making from one centre. In essence, the foundation for this case is that, by definition, strategic decisions are the most important determinants of a firm’s broad direction. (p. 66)

As a result, ‘the strategic decision-making approach sees firms as encompassing market and non-market transactions’ (p. 68). Skipping this unwarranted inference, to deal with a relatively minor point first, Cowling and Sugden, refer to the ‘modern’ corporation. This might suggest that they are observing a new reality, and making appropriate amendments to ‘the theory of the firm’ to deal with novel phenomena. But it seems perfectly clear that ‘strategic decision-making from one centre’ is not exclusively ‘modern’ but has been a feature of enterprises since medieval times.

Essentially, the Cowling-Sugden argument boils down to this: (1) matters of power and control are supremely important, hence (2) we must redefine the firm as a centre of strategic decision-making, and (3) a focus on the distinction between market and non-market activity is not important. Proposition (1) is plausible but proposition (2) does not logically follow from it. The words ‘firm’ or ‘corporation’, which already have a proper meaning that has not been successfully challenged, cannot legitimately be used to apply to something else, such as a ‘strategic centre’. In addition, proposition (3) does not follow from either (1) or (2). The fact that power and control are paramount does not imply that other matters can be ignored.

In particular, if matters of power and control are supremely important then this might require, rather than exclude, consideration of the legal and contractual realities. A corporate ‘centre of strategic decision-making’ exercises its power and control partly through legal instruments. Furthermore, the varied nature of those legal instruments has important consequences for the scope and type of power or control. The differences between self-employment, employment and slavery – largely ignored by the Warwick school – are substantial differences of power and control over individuals, and yet cannot be understood with recognition of legal realities. A firm’s power over an employee is different from that of a self-employed contractor, because in the latter case the firm has no right to interfere in the nature and pattern of work. In addition, a firm’s power over its internal divisions can be significantly different from the power of a firm over a subcontractor, and so on. These distinctions are backed up by the very real powers of the state and the courts. An emphasis on power and control is no excuse for ignoring the legal realities. On the contrary, it is the occasion to introduce them, alongside other more subtle mechanisms of power.

The Cowling-Sugden ‘redefinition’ of the firm is extremely inexact. They give no clear criterion to determine a ‘centre of strategic decision-making’ or the boundaries of its control. Consider the case of a large corporation that has a number of smaller subcontractors and suppliers – such as Benetton, or Marks and Spencer. If we mistakenly defined a firm in terms of a broad notion of control, then the large corporation, plus all the subcontracted suppliers, would together be regarded as a single firm. However, this would simply – and with potential confusion – shift the definition of ‘the firm’ from one type of phenomenon to another. Clearly, we require two terms. One – the firm – would describe a productive organization constituted as singular legal entity. The other – such as a ‘supplier network’ – would describe the entire clustered complex of subordinate subcontractors that are dependent on the contracts of a dominant organization. It is simply confusing to shift the word ‘firm’ from the former to the latter.

If it were defined clearly, it would be possible to consider the ‘centre of strategic decision-making’ as an additional key analytical category in industrial economics. But this would not be a theory of the firm or corporation. Neither would the distinction between the firm and the market be undermined or deemed irrelevant.

A firm, a conglomerate, a ‘centre of strategic decision-making’ and a ‘supplier network’ are different. A firm is a single ‘legal person’. A conglomerate is a set of firms, wholly or partly owned by a holding firm and acting as a single entity. A supplier network is a set of firms, each of which may be dependent on the (written or unwritten) contracts of a single firm, without necessarily being wholly or partly owned by the dominant firm.

Although the purported ‘hybrid’ cases disappear when we examine them closely, there is one important case where the boundaries are often difficult to draw. This concerns the distinction between an employment contract and a contract for services. A firm may change the status of its workers from employed to self-employed, hiring them to do the same work much as before. The key legal and substantive difference here is that, under an employment contract, the employer has the right to control and interfere with the manner and process of work. There is no such right over a self-employed worker: there is simply the right to obtain the contracted good or service at the agreed quantity and quality. In practice, however, the line between employment and self-employment is difficult to determine. Nevertheless, although the line is difficult to draw, the distinction is still real (Masten, 1991). Difficulties of demarcation between types do not imply that differences of type are non-existent.

It should also be pointed out, however, that according to the definition of the firm proposed in section 8 below, employment relationships are not a necessary feature of a firm. Partnerships without employees, worker cooperatives, or enterprises with slaves, can also be firms. As argued elsewhere, the particular type of firm that is dominated by an employment relationship can be described as the ‘capitalist firm’ (Hodgson, 1999: 220–46).

7. Some Voices in the Wilderness

We have already noted Jensen and Meckling’s (1976) recognition of the legal basis of the firm. Subsequent analyses of the legal realities are rare, and two deserve a brief mention here. In an excellent but neglected paper on ‘a legal basis for the firm’ Scott Masten (1991) shows that the distinction between employment and

self-employment has a number of important ramifications. On this basis, he goes further to argue that there is a close correspondence between legal forms and internal relations of power and control:

Overall, the duties accruing to the employment contract under the law are consistent, at least superficially, with the authority and information properties often associated with internal organization. . . . In a formal sense, the distinction between the ownership and governance roles of the firm is a spurious one, however. Ownership itself is a condition sustained by legal rules and remedies. (Masten, 1991: 201–8)

Masten rightly argues that the governance of the firm cannot be understood without recognition and analysis of legal rules.⁹

In another neglected paper, Louis Putterman (1988: 244) considered three different conceptions of the firm: ‘the firm as saleable commodity, the firm as production coalition and the firm as polity or association.’ He argued that ‘the place of the concept of *the ownership of the firm* is poorly developed in most of this literature, with many writers either ignoring the concept entirely or arguing that it is of no importance.’ However, ‘the concept of ownership of firms is crucial to an understanding of internal governance issues’ (p. 243).

Putterman pointed out that while most economists ignore the concept of the ownership of the firm they implicitly assume that firms must be ownable and saleable, especially if they are to be operated efficiently. He wrote: ‘The firm in contract law and the firm that is traded in the marketplace is not the team as such but rather an entity that enters into contractual relations with team members, in the role of employer’ (p. 245). Accordingly, in Putterman’s view, the firm is a commodity.

However, while Putterman’s reinstatement of the legal realities is highly commendable, he gives them a relatively excessive and virtually exclusive emphasis. His argument that the firm is a commodity is valid. But it does not exclude the additional fact that the firm is an organization, even if it is treated in law as a singular ‘legal person’.¹⁰

8. Concluding Remarks

Clearly, in modern economies, there are many different, interlocking structures associated with the sphere of production. However, no good argument has been found to abandon the original, legally-based definition of the firm. In addition to a legally grounded definition of the firm, what is required is a plurality of other concepts and definitions to dissect the complex reality.

From this vantage point, the claims that the boundaries of the firm have broken down, and that there is no essential distinction between firms and markets, have been shown here to be unproven or false. Internal markets within firms prove to be chimerical. The alleged cases of firm-market ‘hybrids’ turn out to be interlocking relations or networks between multiple and distinct legal firms or legal persons, rather than a single, encompassing firm.

Recognition of the general absence of markets and commodity exchange inside all firms is important for several reasons. It dispenses with confused terms such as ‘internal market’, ‘continuum’ and ‘hybrid’. It also helps to underline the relevance of the boundaries of the firm and the vital interface between non-market and market

modes of coordination. Any analysis of the formation and role of these boundaries has vital implications for corporate and public policy. Crucially, a neglect of legal realities impairs any attempt by the social scientist to advise on appropriate legal structures for economic performance. If economists or sociologists proclaim that legal realities are unimportant, then they will be inadequately equipped to advise us on the effects of legal changes such as corporate takeovers, divestitures or privatizations. Furthermore, a possible implication of blurring the distinction between the firm and the market is to allow the firm to be treated as if it were a market, and perhaps left to its own devices.

Above all, without attention to legal relations, social scientists are ill equipped to intervene in the long debate concerning the limitation of abuses of corporate power. They will be less able to evaluate the conditions involved in any legal incorporation of a firm by the state, and the nature of the *quid pro quo* for society in return for the legal privilege of limited liability. Without attention to these features, the social scientist may become dangerously indifferent to policies that extend or diminish the scope of corporate power or the real market.

Like all myths, perceptions of veracity have been reinforced more by mere repetition than by solid theoretical argument or by empirical authentication. The mythical conflation of the firm and the market allows many economists to ignore the reality of non-market organization in the private sector and bring everything there under the conceptual umbrella of market analysis. By many, corporate control and authority is treated purely as a matter of free contract. They can thus ignore the reality of control and authority within the private capitalist corporation but remain critical of public sector bureaucracy and state planning. Such misconceptions are aided by the lack of clear and adequate definitions of 'firm' or 'market' in social science.

The lack of clarity and precision over the definition of the firm also impairs progress in the resolution of the current debate between transaction cost and capabilities approaches to the theory of the firm. While one side has claimed victory (Williamson, 1999), it has done so by dissolving the distinction between the firm and the market and by making the 'market' transaction ubiquitous. In contrast, some defenders of the capabilities approach rightly insist that the firm is separate from the market and this has important implications for learning within the firm (Rumelt *et al.*, 1991).

In summary, the case for retaining a legally-based definition of the firm is as follows. At least in countries where the rule of law prevails, laws are part of socio-economic reality and have real effects. In general, legal relations are not mere formalities, but are backed with the powers and sanctions of the state legal system. Laws set bounds on our behaviour, including the kind of contracts which may be concluded. The courts and police, even if they are not used frequently, can be used to enforce the fulfilment of contracts or to enforce the collection of damages for non-performance. These are clearly matters of power and control. The fact that legal relations do not constitute the whole story is beyond the point. The firm is a legally grounded entity but it is not defined here entirely in legal terms. The case has been made here for a definition of the firm in which legal structures are an indispensable feature among others.

The following challenge is posed to those that wish to dispense with a legally-oriented definition of the firm: please provide us with a good reason for doing so. If legal criteria are regarded as dispensable, then the onus is on the researcher to identify the powers and forces that are sufficient to counter the sanctions and

powers of state authority. As yet, no such reasons have been adequately articulated in the literature.

The market has been defined above. Let us attempt two more definitions. We may define an organization as a special type of institution involving (1) criteria to establish its boundaries and to distinguish its members from its non-members, (2) a principle of sovereignty concerning who is in charge and (3) a chain of command delineating responsibilities within the organization. These are basic conditions of organizational unity and integrity, which could exist independently of the law.

A firm is a special type of organization that is legally recognizable and devoted to production. *A firm is defined as an integrated and durable organization involving two or more people, acting openly or tacitly as a 'legal person', capable of owning assets, set up for the purpose of producing goods or services, with the capacity to sell or hire these goods or services to customers.* As a 'legal person,' the firm may carry legal entitlements and liabilities in its own right. These entitlements include the right of legal ownership of the goods as property up to the point that they are exchanged with the customer, and the legal right to obtain contracted remuneration for the produced services. Corporate legal liabilities may also be incurred in the production and provision of those goods or services. Accordingly, subject to taxation, the owners of the firm have the legal right to the residual income of the firm. A sense in which a firm is integrated is that it is regarded as a 'legal person' owning its products and entering into contracts. The sense in which a firm is durable is that it constitutes more than a transient contract or agreement between its core members and it incorporates structures and routines of some expected and potential longevity.

Note that this definition relies not simply on legal, but also on other relations or concepts. Nevertheless, it is argued here that the firm is a distinct legal entity: it is technically a 'legal person'. It owns its products and sells or hires them to others. It enters into written or unwritten contracts with its workforce and its customers. Accordingly, its external relations are dominated by commodity exchanges or markets. Internally, however, neither markets nor commodity exchange rules the firm. The firm is an organization that manages the resources in its possession, essentially by administrative control rather than by internal contract. Accordingly, one of the key features of a firm is that it is an organized enclave, apart from markets and exchange. Unless we understand that institutional reality we do not understand the firm.

A genuine query emerges: why do so many economists evade the obvious, everyday, legally grounded, definition of the firm? There are two likely and related reasons. One is the tendency to associate the so-called 'economic' attributes of a social arrangement with specific subjective perceptions of the agents involved, rather than with the character of the social structures or institutions. The other is the related desire to make economics general and ahistorical, rather than to associate it with historically specific institutions (Hodgson, 2001).

Some writers repel attempts to define the firm as an identifiable legal and non-market entity. They suggest, for instance, that the boundaries between the firm and the market are being eroded in modern capitalism. Some claim that such developments undermine the idea of the firm as a non-market entity; others claim that ownership-based and legal definitions of the boundaries of firms are increasingly irrelevant for understanding the organization of economic activities (Zucker, 1991; Helper *et al.*, 2000). Nevertheless, even if such arguments are true, then we would still have the problem of defining and identifying the past historical

phenomenon of the firm. The extinction of a species does not exempt us from the task of taxonomy.

In truth, however, the firm as a legally-grounded entity is not extinct. In modern capitalism there may have been an important growth of 'market-like devices' within firms, and some 'infusion of hierarchy' into exchange networks (Zenger and Hesterly, 1997: 211). But there has been no intrusion of markets proper into firms, as defined here. And, incidentally, a hierarchically-organized network of autonomous producers is not properly described as a market. Distinctions between firms, markets and other arrangements remain real, and the legal status of the firm has not as yet been undermined.

It may be possible to classify the 'hybrid' cases identified by the authors cited here as 'hybrid organizations' in the sense that they are enduring cases of complex relational contracting between multiple firms. Leaving aside the merits or demerits of such a classification, the argument here is that it would still be no basis to abandon the legally-grounded definition of the firm. If the possibility of a 'hybrid organization' is admitted then it must be clearly distinguished from a firm. The term 'hybrid firm' remains misleading, and true 'internal markets' are absent within the firm as defined here. Given the plague of confusion propagated by the term 'hybrid' it may be advisable, in any case, to use a term such as 'network' instead of 'hybrid organization' (Powell, 1990).

Above all, a muddled reality is no excuse for muddled definitions. Likewise, a mutable reality is no justification for elastic ideas. Accordingly, even if the boundaries between the firm and the market are breaking down in reality, the conceptual distinction between these two terms is still necessary to make sense of such a statement. In order to describe or understand such a tangled reality we need clear concepts and careful definitions to guide us. Without them we are conceptually blind. Regrettably such definitions are relatively scarce and it is an urgent task to provide them.

Notes

1. This paper expands considerably on some arguments in Hodgson (2001, ch. 17).
2. For a useful discussion of the growth of joint ventures, strategic alliances and networks see Hage and Alter (1997), who provide a typology of these inter-firm phenomena.
3. See Foss (2002).
4. For a critique of Williamson's concept of hierarchy see Hamilton and Feenstra (1995).
5. See Khalil (1997) for an interesting discussion of the individuality of the firm. However, according to the argument in the present paper, Khalil's notion of 'consented goal' is an important aspect of an organization but not sufficient to define the firm.
6. According to the definition here, all markets involve some degree of competition. The question of whether the outcome is efficient or inefficient is another issue, depends on a further definition of efficiency, and is not concluded by this definition of the market alone.
7. Some ambiguities arose when Coase (1937: 388) pointed out that in a department store there may be 'competitive bidding for space'. Furthermore: 'In the Lancashire cotton industry, a weaver can rent power and shop-room and can obtain looms and yarn on credit.' These passages might be interpreted as suggesting the existence of markets inside the firm. However, from the perspective of the present article, a department store franchisee is outside the boundaries of the franchising department store firm, and the Lancashire weavers to which Coase referred were self-employed producers, even if they worked inside a mill owned by another firm.
8. In contrast, Bradach and Eccles (1989) propose a 'quasifirm'. What is interesting, however, is their defence of clear, 'ideal types' in the analysis of the complex structures of modern industry. Bradach and Eccles (1989: 116) argue that 'the ideal types of market and hierarchy serve as a useful starting point for studying the organization of industry.' Here is muted plea for clear concepts to act as

markers to help understand real-world complexity. However, they then spoil their case by stating that markets and hierarchies are not mutually exclusive.

9. See also Foss and Foss (2001) for a critique of overly narrow and a-legal concept of ownership in the new property rights economics. Khalil (1997: 523–4) makes some similar points in critiques of Hart and Moore (1990) and Marx.
10. Although David Ellerman (1992) is ultimately concerned to make a normative case for the worker-owned firm, his book pays attention to legal relations as well as matters of rights and power.

References

- Abbott, C.C., *Rise of the Business Corporation*. Ann Arbor: Edwards Brothers, 1936.
- Anderson, G.M. and Tollinson, R.D. “Adam Smith’s Analysis of Joint-Stock Companies”, *Journal of Political Economy*, December 1982, 90(6), pp. 1237–56.
- Berle, A.A. and Means, G.C., “Corporations and the Public Investor”, *American Economic Review*, March 1930, 20(1), pp. 54–71.
- Bradach, J.L. and Eccles, R.G., “Price, Authority and Trust: From Ideal Types to Plural Forms”, *Annual Review of Sociology*, 1989, 15, pp. 97–118.
- Cheung, S.N.S., “The Contractual Nature of the Firm”, *Journal of Law and Economics*, April 1983, 26(2), pp. 1–21.
- Coase, R.H., “The Nature of the Firm”, *Economica*, November 1937, 4, pp. 386–405.
- Coase, R.H., “The Nature of the Firm: Origin, Meaning, Influence”, *Journal of Law, Economics, and Organization*, Spring 1988, 4(1), pp. 3–47. Reprinted in Williamson, Oliver E. and Winter, Sidney G., eds, *The Nature of the Firm: Origins, Evolution, and Development*. Oxford and New York: Oxford University Press, 1991.
- Commons, J.R., *Legal Foundations of Capitalism*. New York: Macmillan, 1924.
- Cowling, K. and Sugden, R., “The Essence of the Modern Corporation: Markets, Strategic Decision-Making and the Theory of the Firm”, *The Manchester School*, January 1998, 66(1), pp. 59–86.
- Doeringer, P.B. and Piore, M.J., *Internal Labor Markets and Manpower Analysis*. Lexington, MA: Heath, 1971.
- Dore, R., “Goodwill and the Spirit of Market Capitalism”, *British Journal of Sociology*, 1983, 34(4), pp. 459–82.
- Du Bois, A.B., *The English Business Company After the Bubble Act, 1720–1800*. New York: Commonwealth Fund, 1938.
- Eccles, R.G., “The Quasifirm in the Construction Industry”, *Journal of Economic Behavior and Organization*, 1981, 2, pp. 335–57.
- Ellerman, D.P., *Property and Contract in Economics: The Case for Economic Democracy*. Oxford: Basil Blackwell, 1992.
- Florence, P.S., “New Measures of the Growth of Firms”, *Economic Journal*, June 1957, 67(2), pp. 244–48.
- Foss, N., “Coase vs Hayek: Economic Organization in the Knowledge Economy”, *International Journal of the Economics of Business*, 2002, 9(1), pp. 9–35
- Foss, K. and Foss, N., “Assets, Attributes and Ownership”, *International Journal of the Economics of Business*, 2001, 8(1), pp. 19–37.
- Goldberg, V.P., “Relational Exchange: Economics and Complex Contracts”, *American Behavioral Scientist*, 1980, 23(3), pp. 337–52.
- Greif, A., “The Study of Organizations and Evolving Organizational Forms Through History: Reflections from the Late Medieval Family Firm”, *Industrial and Corporate Change*, 1996, 5(2), pp. 473–501.
- Hage, J. and Alter, C., “A Typology of Interorganizational Relationships and Networks”, in Hollingsworth, J. Rogers and Boyer, Robert, ed., *Contemporary Capitalism: The Embeddedness of Institutions*. Cambridge and New York: Cambridge University Press, 1997, pp. 94–125.
- Hamilton, G.G. and Feenstra, R.C., “Varieties of Hierarchies and Markets: An Introduction”, *Industrial and Corporate Change*, 1995, 4(1), pp. 51–87. Reprinted in Dosi, Giovanni, Teece, David J. and Chytry, Josef, eds, *Technology, Organization, and Competitiveness: Perspectives on Industrial and Corporate Change*. Oxford and New York: Oxford University Press, 1998.
- Handlin, O. and Handlin, M.F., “Origins of the American Business Corporation”, *Journal of Economic History*, May 1945, 5(1), pp. 1–23.
- Hart, O.D. and Moore, J., “Property Rights and the Nature of the Firm”, *Journal of Political Economy*, 1990, 98(6), pp. 1119–58.

- Hayek, F.A., ed., *Collectivist Economic Planning*. London: George Routledge, 1935.
- Helper, S., MacDuffie, J.P. and Sabel, C., "Pragmatic Collaborations: Advancing Knowledge While Controlling Opportunism", *Industrial and Corporate Change*, 2000, 9, pp. 443–87.
- Hessen, R., *In Defence of the Corporation*. Stanford: Hoover Institute, 1979.
- Hessen, R., "Corporations", in Eatwell, John, Milgate, Murray and Newman, Peter, eds, *The New Palgrave Dictionary of Economics*. London: Macmillan, vol. 1, 1987, pp. 675–7.
- Hodgson, G.M., *Economics and Institutions: A Manifesto for a Modern Institutional Economics*. Cambridge and Philadelphia: Polity Press and University of Pennsylvania Press, 1988.
- Hodgson, G.M., *Evolution and Institutions: On Evolutionary Economics and the Evolution of Economics*. Cheltenham: Edward Elgar, 1999.
- Hodgson, G.M., *How Economics Forgot History: The Problem of Historical Specificity in Social Science*. London and New York: Routledge, 2001.
- Holmström, B., "The Firm as a Subeconomy", *Journal of Law, Economics, and Organization*, 1999, 15, pp. 74–102.
- Hunt, B.C., "The Joint-Stock Company in England, 1800–1825", *Journal of Political Economy*, February 1935a, 43(1), pp. 1–33.
- Hunt, B.C., "The Joint-Stock Company in England, 1830–1844", *Journal of Political Economy*, June 1935b, 43(3), pp. 331–64.
- Imai, K. and Itami, H., "Interpenetration of Organization and Market: Japan's Firm and Market in Comparison with the US", *International Journal of Industrial Organisation*, 1984, 6(4), pp. 285–310.
- Jensen, M.C. and Meckling, W.H., "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, 1976, 3, pp. 305–60.
- Khalil, E.L., "Is the Firm an Individual?", *Cambridge Journal of Economics*, July 1997, 21(4), pp. 519–44.
- Klein, B., "Contracting Costs and Residual Claims: The Separation of Ownership and Control", *Journal of Law and Economics*, 1983, 26, pp. 367–74.
- Langlois, R.N., "Capabilities and Coherence in Firms and Markets", in Montgomery, Cynthia A., ed., *Resource-Based and Evolutionary Theories of the Firm: Towards a Synthesis*. Boston: Kluwer, 1995, pp. 71–100.
- Machlup, F., "Theories of the Firm: Marginalist, Behavioral, Managerial", *American Economic Review*, March 1967, 57(1), pp. 1–33.
- Marsden, D., *The End of Economic Man? Custom and Competition in Labour Markets*. Brighton: Wheatsheaf Books, 1986.
- Marshall, A., *Industry and Trade*. London: Macmillan, 1919.
- Masten, S.E., "A Legal Basis for the Firm", in Williamson, Oliver E. and Winter, Sidney G., eds, *The Nature of the Firm: Origins, Evolution, and Development*. Oxford and New York: Oxford University Press, 1991, pp. 196–212.
- Ménard, C., "Markets as Institutions versus Organizations as Markets? Disentangling Some Fundamental Concepts", *Journal of Economic Behavior and Organization*, 1995, 28(2), pp. 161–182.
- Ménard, C., "On Clusters, Hybrids, and Other Strange Forms: The Case of the French Poultry Industry", *Journal of Institutional and Theoretical Economics*, March 1996, 152(1), pp. 154–83.
- North, D.C., "Markets and Other Allocation Systems in History: The Challenge of Karl Polanyi", *Journal of European Economic History*, 1977, 6, pp. 703–16.
- Pitelis, C., *Market and Non-Market Hierarchies: Theory of Institutional Failure*. Oxford: Basil Blackwell, 1991.
- Powell, W.W., "Neither Market nor Hierarchy: Network Forms of Organization", *Research in Organizational Behavior*, 1990, 12, pp. 295–336.
- Putterman, L., "The Firm as Association versus the Firm as Commodity: Efficiency, Rights, and Ownership", *Economics and Philosophy*, 4(2), October 1988, pp. 243–66.
- Richardson, G.B., "The Organisation of Industry", *Economic Journal*, 1972, 82, pp. 883–96.
- Robertson, D.H., *The Control of Industry*. London: Nisbet, 1923.
- Rumelt, R.P., Schendel, D. and Teece, D.J., "Strategic Management and Economics", *Strategic Management Journal*, 1991, 12, pp. 5–29.
- Sened, I., *The Political Institution of Private Property*. Cambridge: Cambridge University Press, 1997.
- Smith, A., *An Inquiry into the Nature and Causes of the Wealth of Nations*, 2 vols, first published 1776, ed. by Roy H. Campbell and Andrew S. Skinner. London: Methuen, 1976.
- Williamson, O.E., *Markets and Hierarchies: Analysis and Anti-Trust Implications: A Study in the Economics of Internal Organization*. New York: Free Press, 1975.
- Williamson, O.E., *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. London: Macmillan, 1985.

- Williamson, O.E., "Strategy Research: Governance and Competence Perspectives", *Strategic Management Journal*, 1999, 20, pp. 1087–1108.
- Zenger, T.R. and Hesterly, W.S., "The Disaggregation of Corporations: Selective Intervention, High-Powered Incentives, and Molecular Units", *Organization Science*, May-June 1997, 8(3), pp. 209–22.
- Zucker, L.G., "Markets for Bureaucratic Authority and Control: Information Quality in Professions and Services", *Research in the Sociology of Organizations*, 1991, 8, pp. 157–90.