Firms as social actors

RICHARD ADELSTEIN*

Department of Economics, Wesleyan University, Middletown, Connecticut, USA

Abstract: This essay asks what firms are, whether they are ‘real’ social actors, and whether their actions can be traced without remainder to the actions of living people or whether there is some irreducible aspect of their existence or operation that must be attributed to the organization itself. It describes firms as ongoing, multilateral relational contracts from whose operation – that is, from performance over time by specific individuals in the roles and relationships defined by the contract – emerge the firm’s idiosyncratic routines and capabilities. It emphasizes the role of entrepreneurs in the creation of firms and the close dependence of organizational capabilities on human performance, and argues that this account is consistent with a reasonable individualism that allows for social outcomes to be determined by the actions and interaction of individuals. It then proposes that firms are nonetheless institutional facts and thus ontologically subjective but epistemically objective components of reality, and concludes with directions for future work.

Scholars and jurists have long puzzled over how to think about organizations, associations of individuals characterized by one or a very few governing objectives and, generally, some sort of central plan designed to achieve them. This is particularly true of business firms, organizations dedicated to the production and sale of goods for profit. Social scientists have not yet resolved the interrelated ontological and epistemological questions of just what these ubiquitous social objects are, whether and in what sense they are ‘real’ social actors, and whether their actions and effects can ultimately be traced without remainder to the actions of living human beings in the here and now or, contrarily, whether there is some irreducible aspect of their existence or operation that can only be attributed to the organization itself. Courts and legal theorists, meanwhile, have struggled with the normative dimension of the problem, the granting of unitary legal personality to business firms and its implications for the rights and powers of individuals and collectives. My purpose here is to address this complex of questions, not comprehensively, but with primary emphasis on the ontological and epistemological issues raised by firms. I will try to say precisely what firms are, how they originate, and how their actions are linked to the intentionality and

*Email: radelstein@wesleyan.edu

I am grateful to David Campbell, David Gindis, Bill Lazonick, Lee McIntyre, Tom Miceli, Heath Pearson, Philip Pettit and two anonymous referees of this journal for helpful criticism. All remaining errors are my own.
behavior of their human participants, and argue that firms are a species of what Searle (1995, 2005) calls institutional facts and thus ontologically subjective but epistemically objective components of reality. Because any adequate treatment of the normative problem of collective rights and powers must be based on some resolution of the ontological and epistemological questions, I leave the former for another time, except insofar as issues of legal personality pertain to the latter. My aim is simply to propose a historically plausible, analytically productive perspective on firms that offers a useful conceptual basis for considering their rights and powers.

1. Thinking about firms

Firms as entities

Before the Civil War, when almost every American firm was either a sole proprietorship or small partnership, these questions were essentially moot. Business relations were governed by contract law; individual owners and partners bore unlimited liability for the obligations of their enterprises; and the forms within which business was conducted were understood as convenient legal artifacts, transparent, insubstantial masks behind which easily identifiable human beings who could be held personally responsible for the conduct of their firms did business. This began to change in mid-century, when the right to incorporate with limited liability was generally extended to any lawful business, and with the rapid growth of large-scale, mass-production enterprises after 1870, the emerging industrial giants took on an undeniable social presence, especially when they adopted the corporate form. Incorporation offered not only limited liability to owners but perpetuity to the corporation itself, legal life that extended beyond the participation of any individual in the enterprise, however important, and a legal personality that enabled the firm to own property, make contracts, and appear in court in its own name, independently of the living people who at any moment might own, operate or work for it. In 1886, the United States Supreme Court declared that corporations had the same constitutional rights to due process and equal protection of the laws that living people did, though as Horwitz (1985: 176–178) shows, this was intended to protect the rights of individual shareholders acting in concert through the contractual artifice of a corporation rather than to recognize the latter as an independent, rights-bearing entity distinct from its human constituents.

The Court’s insistence on seeing corporations as ‘artificial entities’ just as it was bestowing constitutional personhood on them reflected a traditional American jurisprudence of rights vesting solely in living people and a wariness of the power large corporations could wield if granted similar rights. But it was a rearguard action against an advancing conceptual alternative that saw corporations not as a legal artifices or fictions but as real, active social objects that were not created by
the law and thus could only be regulated by it. Gindis (2009: 31–36) traces three stages in the evolution of this real entity theory. The earliest realists, influenced by the German theorist Otto von Gierke, embraced a literal organicism that fully reified corporations and treated them as natural, purposeful social beings separate in principle from their participants and endowed with economic interests identical to those of living people. In this view, a corporation is an ‘autonomous, self-sufficient, self-renewing body’ (Davis, 1897: 294), ‘no fiction, no symbol [but] a living organism and a real person, with body and members and a will of its own’ (Maitland, in Gierke, 1987: xxvi). By 1920, a second generation of realists had largely abandoned this explicit organicism for the milder notion of social collectivity proposed by the American Ernst Freund. They acknowledged that the ‘body corporate’ was not an actual organism but maintained that it was a real social entity, identifiable as such by ‘its unity, its distinctiveness, and its identity in succession’ (Freund, in Gindis, 2009: 33). They believed, moreover, that the social reality of any collective body that the law comes to deem a corporation generally precedes the fact of this legal recognition, and that ‘the law should comply with the fact of the group’s socio-economic existence and attribute legal capacity to an already existing or a potential socio-economic capacity’ (Gindis 2009: 33).

Contracts and capabilities

The more or less overt collectivism of the entity theorists produced hostile reactions in the 1930s from both legal individualists and legal realists, and Gindis describes third-stage real entity theory in the legal academy as in retreat, preserving the ontological commitments of the second generation but more concerned with practical than philosophical matters, before becoming moribund after 1950. The individualist response to real entity theory has been more vigorous among economists, characteristically committed to one or another shade of ontological reductionism under the banner of ‘methodological individualism’. Their strategy has been to focus on the transaction rather than the decisionmaker, to see firms as a way of organizing (or ‘governing’) economic activity by contractual arrangement rather than market exchange, and to explore the conditions that drive activity from market to contract or vice versa.

Two distinct contractual theories of the firm have emerged since 1970, both claiming descent from Coase (1937) but very different in their depictions of the contracts involved. One, the nexus of contracts model introduced by Alchian and Demsetz (1972) and Jensen and Meckling (1976), illustrates what Hoover (2009: 388) calls the ‘reductionist impulse’ of neoclassical thinking in its decisive rejection of social collectives as causal agents. It treats firms as coalitions of individuals linked through a set of contracts, either with one another directly or with a fictitious central party called ‘the firm’, which is itself nothing more than an imagined common signatory of these agreements, ‘no more than a name’ for what are in essence a potentially large number of distinct contracts among
workers, managers, and contributors of capital (Easterbrook and Fischel, 1985: 89). When markets are competitive and working well, as the nexus theories assume, contractors have options they want to preserve, so contracts tend to be discrete, with clearly specified terms, easy entry and exit, and short duration (Macneil 1978: 856–865). A nexus of discrete contracts is thus not so much the haven from market exchange that Coase described as a seamless extension of market exchange to the domain of production over time, effacing the line between market and firm. Indeed, to Jensen and Meckling (1976: 311), it makes ‘little or no sense to distinguish those things that are “inside” the firm . . . from those things that are “outside” of it’, and from here it is a short step to denying the firm as such any substantial agency or reality at all. ‘[T]he firm itself is not an individual. It is a legal fiction which serves as a focus for a complex process in which the conflicting objectives of individuals . . . are brought into equilibrium within a framework of contractual relations’ (emphasis in original). Firms, that is, do not exist prior to law but are created by it, to be nothing more than placeholders, a convenient linguistic shorthand for the conceptual site of a set of discrete contracts that govern easily terminated relations among people in some sphere of productive activity.

A second, relational strand of contractual theory derives from Macneil (1978) and Williamson (1979) and considers environments in which markets work poorly. It adopts Coase’s depiction of market exchange as an imperfect, costly means of allocating resources that is, as a result, often inferior in production contexts to long-term contractual relationships that replace allocation by price in markets with allocation by plan in firms. In this view, firms are conceptually distinct from markets, an alternative way to organize production that obviates the transaction costs of market exchange by making production governed by a long-term relational contract cheaper than production governed by more discrete exchange. The ability of successful firms to economize on transaction costs means that the contracts that constitute them have real economic value, which suggests that they are ‘real’ social objects in a sense they are not in the nexus theories. But most transaction-cost theorists, clinging to as much of the neoclassical reductionism as this insight will permit, go no further than this; once the contract loses its functional quality, and production in markets becomes cheaper than production in the firm, its value to the contractors disappears with it. The working relationships created and sustained by the individual contractors for as long as their association was profitable are accorded no value to participants beyond their ability to reduce transaction costs and, like any other tool in the workplace, presumably scrapped once they no longer work. This honored the commitment to methodological individualism and kept the beast of collective agency at bay, but, as critics noted, it created difficulties in accounting for either the telos of profit maximization presumed in the models (Dow, 1987: 25–33) or the possibility that organizational experience in adapting to transactional frictions over time might alter the calculus of market or firm (Langlois, 1992).
This last point has been pursued in yet another theory of firms, the capabilities approach, which breaks free from the reductionist impulse and takes firms seriously as collective actors. Rooted in Penrose (1959), Richardson (1972), and Nelson and Winter (1982), capabilities theory describes firms themselves in cognitive terms, as repositories of collective, tacit knowledge and capacities, embodied in firm-specific behavioral routines that are created within the firm and evolve over time in competitive environments. This knowledge is not held by any individual in the firm, nor is it the aggregate of what these individuals know, nor is it meaningful to say that the firm as such holds it as a kind of thinking collective. Rather, the firm’s knowledge and capabilities are manifest in the operation of the firm itself, and change as new participants and technologies come and go and the firm gains experience in the market. From this perspective, we see what the firm ‘knows’ or ‘intends’ and what it is capable of doing by observing what it does and how it does it, just as we might infer these things from watching individuals act. The firm itself is manifested at every moment in the performance of its distinctive behavioral routines by its participants of the moment, a process in which the personal identities of the individuals involved and the nuanced effects of their interactions combine to produce collective actions, the outcomes of the routines, that could not have been predicted from full knowledge of the participants in isolation. Routines are emergent social phenomena, and the firms that display them, while not living organisms, are nonetheless autonomous, even lifelike social beings existing apart from their temporal participants, capable of intentionality and social action in their own right, able to learn, change and reproduce over a life of indeterminate length.

Entities redux

Seeking to revivify real entity theory in a fourth incarnation, Gindis (2007, 2009) focuses primarily on the idea’s fate and prospects in the legal academy. Apart from the role of foil he assigns to the nexus theories, he is silent on the relational theories and mentions capabilities theory only in passing, despite its clear resemblance to second- and third-generation real entity theory. But the concept of emergence is crucial to his account. Drawing on an eclectic range of legal and philosophical sources, Gindis argues for the rough-and-ready Freundian view that a firm’s ‘power to cause, power to stay unified, power to avoid dissolution’ is sufficient to demonstrate its reality, and that its organizational unity or coherence, the cohesiveness of the bonds it creates among its participants, and its persistence through time evince an emergent ‘whole which can no longer be viewed as reducible to any of its constituent parts or members’ (Gindis 2007: 278, 282).

I believe that Gindis and the capabilities theorists are right: modern firms, corporations or otherwise, are social objects whose operational characteristics do not arise from any aggregation of the qualities of their participants but emerge from the interactions among them, and they are not just real social actors that
do cause events, but political actors as well, insofar as they subject some people
to the limited authority of others. But given the scientific and philosophical
commitments at stake in competing claims that firms are unitary, active social
objects irreducible to their human components or merely conceptual markers
of a complex process involving only human actors, it is surprising that neither
he nor they say precisely what these active social objects are. Without a clear
account of exactly what it is in the material and social universe to which the
names firm and the firm’s acts are attached, it is hard to pin down the claim
that firms are real or to get very far into how they work. If reductionists are
often too keen to deny the existence of active collectives and so deny themselves
access to knowledge of real phenomena, perhaps an ‘anti-reductionist impulse’
leads those who claim reality for collective actors to avoid the complex details
of how the collective actually performs ‘on the ground’, in the mistaken belief
that to trace all the actions of the firm to the actions of living people in the here
and now is to deny the social reality of the firm itself.

Gindis’s fourth stage entity theory differs from all its predecessors in an
important respect. For them, the firm is a social reality before it is recognized by
the law; it is not, in its essence, primarily a legal phenomenon. The law clearly
provides the general background of social order that the growth of firms requires,
and often seeks to regulate the operation of firms in various ways or direct firms
into certain legally defined patterns so as to serve a variety of ends. But it can
only respond to the behavior of firms; it does not create them or make their
behavior possible. How the state chooses to regulate the activity of firms, and to
what sorts of enterprises it will or will not grant legal privileges, obviously has
consequences for the success and continuity of those firms (cf. Blair, 2003). But
Gindis, following Searle (2005: 15–17), goes much further, concluding that legal
recognition as a distinct entity is a condition precedent of the firm as a real entity.
‘The law’s creation of legal entity status is thus a fundamental institutional fact
about the firm’ (Gindis, 2009: 39), a proposition nexus theorists would endorse.
I argue below that the second-generation theorists were right on this question:
legal status enables independently originated firms to be regulated, but, by itself,
it does not constitute them.

In Section 2, I draw on the relational contract and capabilities approaches to
examine the sense in which firms act ‘on the ground’ as a means to understanding
exactly what they are. I argue that firms are constantly changing, multilateral
relational contracts from whose operation – that is, from performance over time
by specific men and women in the roles and relationships defined by the contract –
emerge the firm’s idiosyncratic routines and capabilities. I stress the role of
entrepreneurs in the creation of firms and the close dependence of organizational
capabilities on human performance, and argue that my account of routines
emerging from contracts and the reality of firms as collective social actors is
consistent with a reasonable individualism that allows for social outcomes to be
determined by the actions and interaction of individuals. This being said, I use
Searle’s rigorous, compelling account of social reality to argue in Section 3 that firms are real institutional facts, though not in the way Searle describes them, epistemically objective, ontologically subjective social actors with strong links to ontologically objective reality. A brief conclusion sketches directions for further inquiry.

2. What firms are, and how they act

Relational contracts

Firms are not nothing, ontologically empty boxes within which individual contracts are said to be located. They are distinct social objects constituted by what Macneil (1978) calls relational contracts. Macneil depicts the exchange environment as a continuum in two dimensions: one measuring the ability of bargainers to presentiate, to foresee and bring the relevant future to bear on negotiations in the present; the other identifying transactions as discrete, like an instantaneous exchange between strangers in their substantive isolation from other transactions, or relational, embedded in what Commons (1968: 143–153) called a ‘going concern’ that all sides have an interest in preserving beyond the single transaction at hand. Relational transactions take place within a larger, continuing network of personal bonds that create a context for the transaction and affect the terms of future negotiations, just as they are affected by other transactions in the past. Macneil argues that the degree of uncertainty that surrounds a transaction and the intensity and duration of the parties’ pre-existing relationship influence the way they choose to organize their relations with one another. Where the degree of discreteness is high and the parties can presentiate effectively, contracts are easily made and rescinded and precise terms can usually be fully elaborated. But as transactions become more relational, or it becomes harder to characterize adequate performance when the contract is made because the circumstances under which it will have to take place cannot be anticipated, or both, the parties are likely to account for the possibility of future conflict by leaving such details unspecified in a relational contract and agreeing instead to procedures for taking decisions and resolving disputes as they arise without terminating the relationship itself.

Relational contracts thus focus on the going concern rather than the instant transaction, replacing the easy ‘in-and-out’ of discrete contracts with periodic renegotiation of terms by the parties intended to preserve the relationship; and because both past and future outcomes matter to the parties, they are subject to gradual, conservative change intended to institutionalize successful adaptations and reduce conflict. They are truly social contracts in a way that discrete contracts, even those with many parties, are not. The very idea that firms are nothing, fictions without substance, suggests that the contracts of which such firms are the ‘nexus’ govern very discrete transactions, in
which resources are highly mobile, and short-term price fluctuations are more important to contractors than durable personal or collective relationships. Were it otherwise, and long-term contractual relationships more important than short-term advantage, the parties would see the contracts themselves as having real value to them, something to be preserved, and thus not nothing. But every ongoing firm has this quality to some degree. When participants say they are loyal to a firm, or that they are not, they mean that the roles and relationships through which they interact at work are or are not satisfying to them, things to be preserved or discarded. It is the roles and relationships expressed in the relational contract, the substance of the contract itself, as it is and has been manifested in the actual operation of the firm during each participant’s experience, that is ‘the firm’ in their (and others’) minds, the something to which they do or do not feel attachment.

As in a successful marriage, the archetypical relational contract, daily negotiations in successful firms are undertaken with the intention and expectation of maintaining a valuable larger relationship. The past matters, loyalty is evoked, and the relationship itself is constantly adapted by the parties of the moment to changing conditions within the firm, such as turnover of personnel or the introduction of new technology, and in the competitive environment outside it. At any moment, the activity of the firm manifests the terms of the agreement among its participants at that moment, and adaptive change in its organization and operation is made possible by, and reflected in, concurrent change in the terms of the contract and the identities of the contractors. As firms, relational contracts may encompass large numbers of people intent on preserving the relationship among them, and take on a broader social character. Often, Macneil observes, they create ‘a minisociety with a vast array of norms beyond those centered on’ immediate transactions that can survive the eventual replacement of every founding participant. Given its adaptive history, such a contract ‘may or may not include an “original agreement”; and if it does, may or may not result in great deference being given it’ (1978: 901, 890).

A firm is not a nexus of many discrete contracts, but a single, constantly changing relational contract among individuals whose goods and services must be combined over time to create the firm’s product. Its subject is the precise terms under which this combination is to take place at the present moment, and agreement to it creates an obligation in every constituent to perform in more or less well-specified ways in concert with the others for as long as they choose to remain bound by the contract. The contract that constitutes the firm governs production within it at the same time. That is, it codifies the consensual relationships and obligations of the firm’s participants to one another in the particular process of production they agree to undertake together, though the explicitness of its terms and the prevalence of unspoken working rules (‘organizational culture’) that are part of it vary from firm to firm. As a social contract, a ‘meeting of the minds’ of the individual contractors, the firm as
such resides precisely in the agreement and obligations of the parties expressed in the contract. It is an idea, a mental construction, not a tangible object or a living being. Observing all the factories and offices of a modern corporation, an anthropologist from another planet, able to see what humans do but unable to read their thoughts, could see almost everything the firm does, but could not see the firm itself.

The state (‘law’) need play no role in the creation or persistence of relational contracts, apart perhaps from providing a background of stable social order. This is because the point of the contract is the relation; the parties make and comply with the contract because they want to have and preserve the relation that contract constitutes and governs. The state has its reasons for wanting to regulate the relations of people who live together as spouses, and to encourage people to seek legal recognition of their agreement to do so, it confers benefits on couples who submit to the requirements for legal marriage or penalizes those who do not. But if two people actually contract to live as spouses, in their own words or signs, with no witnesses or ceremony, and then live together accordingly, they will have created the same social object as a legally married couple, less the legal standing itself, and may live lives identical to those they would have lived had they been legally married, less whatever legal entanglements marriage might have produced. The essence of the marriage is the relational contract between the spouses, not the law’s certification of it, and the contract can be enforced by consent and policed by trust even if the state is not involved at all. Once the relational contract exists in the minds of the parties, the relationship is created and the parties see themselves as spouses, a perception others might share, even if the state imposes its imprimatur and regulations only later or not at all. Indeed, ‘common law marriages’ may be recognized as legal given a showing that a couple has lived together as spouses, that is, performed together in a relational contract, even with no prior exchange of promises. People can create relational contracts to live as spouses with or without benefit of law; a legal marriage without a relational contract truly is a fiction.

This is true of firms as well. People everywhere participate in successful relational contracts to produce and sell goods without engaging the state at all. The ‘informal’ economies of rich and poor countries alike are populated by uncounted numbers of small, often transient but nonetheless functional firms providing a variety of products in plain sight but under the state’s regulatory radar, and, of course, where the law makes trade in certain goods illegal, only illegal firms trade in those goods. All that is required for any firm to exist and operate is a relational contract among the participants, a meeting of minds as to the governance of the enterprise, and something to give each of them confidence that the others will perform as agreed. Personal reputation or market discipline might be that something, or trust and affection, as in a marriage. Or the parties might agree to an extralegal enforcement system where they have no recourse to the legal one, as in organized crime. It is certainly easier to sustain a large or
complex relational contract if it is recognized and supportively regulated by the
law. But while the law can do much to help firms succeed or fail, firms do not
need legal recognition to exist, and the law does not constitute or create them.
That is the business of the relational contract, and the contractors.

Asset specificity
Williamson (1979: 235–239) acknowledges Macneil’s contribution to his own
thinking, and the Coasean production contracts Williamson posits as adaptive
responses to transaction costs in markets are clearly relational in character.
He cites three critical dimensions of production transactions that combine
with the opportunistic behavior and bounded rationality of ‘contractual man’
(Williamson, 1985: 43–67) to drive transactions from market to contract:
imperfect knowledge, the frequency of transactions, and most importantly, asset
specificity, the extent to which particular human and material resources become
significantly more valuable in some uses than others. Where costly assets are
specific and substantial investments of time and money leave particular buyers
and sellers little choice but to deal with one another, the time and effort needed to
negotiate terms of trade and the opportunities they present for strategic behavior
and exploitation are potent inducements to contract for both sides. If these
bilateral monopolists must deal at arm’s length with one another on a regular
basis, the pressure to replace the cost and acrimony of recurrent bargaining
with the ‘once-and-for-all’ negotiation of a long-term relational contract will
be stronger still. In this way, Williamson argues, the typical business enterprise
extends itself backward and forward along the production chain by offering
the free agents who provide its inputs and purchase its outputs the chance
to renounce their independence and join the contract themselves. Where the
relationship between the firm and a still independent trading partner is a recurrent
bilateral monopoly, the contract is likely to expand to bring that partner into
the firm. But when the goods or services traded across the firm’s boundary become
sufficiently generic to support effective competition on one side or the other, the
pressure to expand the contract subsides and the firm stops growing.

For Williamson, asset specificity is a problem to which firms are the solution.
The parties seek the security of a long-term relationship as a way to avoid the
costs of recurrent bargaining over specific assets; as noted, the relationship itself
has little value to them beyond the transaction costs it obviates. Markets are
the nonpareil of efficient organization, and firms are second-best alternatives in
exchange environments that cannot support markets. The sources of transaction
cost, bounded rationality, opportunism, asset specificity and the rest, are static
and conditioned by technologies and ways of doing things as they exist at
a particular moment in time. In this light, as Lazonick (1991: 213–227; cf.
Langlois, 1992: 102–105) points out, firms are adaptive rather than innovative.
They take the exchange environment as given and try to optimize within it
rather than envisioning new, constraint-relaxing responses to the environment
and introducing the new technological and managerial strategies these require. In contrast, the nineteenth-century firms chronicled by Chandler (1977, 1984) were Schumpeterian innovators, radical disturbers of the established order. In Chandler’s framework, entrepreneurs envision an innovation, like deploying a new mass-production technology, formulate a strategy to profit by investing in the technology, and then invest in the new organizational structures and managerial talent needed to realize the productive potential of the technology. Entrepreneurs thus innovate twice, investing first in the technology and then in the organization, and it is the two together that enable the firm to prosper. The second investment includes more than just purchasing the skill to coordinate the movement of people and materials through a complex production process. It also means gaining the willing cooperation of owners, managers, and workers with the governing objectives of the enterprise through tolerable working conditions and a mutually agreeable division of the surplus created by the technology.

Human asset specificity, the greater value of one’s labor in one firm than in any other, takes on a very different character for Lazonick than for Williamson. As the founding participants of a new firm work together to solve the myriad of context-specific, practical problems involved in actually building and operating, say, a nineteenth-century railroad or steel mill, working relationships are negotiated, formally and informally, at the same time that vast amounts of practical information, about furnaces and locomotives but also about productively integrating men and machines in satisfactory ways, are acquired and put to use through those relationships. These working relationships, the substance of the relational contract, are thus both the means by which new knowledge is created and, in the actual operation of the production process they make possible, the principal repository of the knowledge itself. In time, these initial relationships become the firm’s idiosyncratic routines and procedures, gradually changing working rules that assign roles and delineate relationships among human actors that, when put into practice, constitute, as Lazonick sees them, the powers of the collective as such to see, think, decide, and act. The existing knowledge distilled in these routines, and the new problem-solving knowledge they produce in turn, are not held by, or accessible to, any one participant, nor are they the aggregate of what all the participants, past or present, knew and know. They are ‘possessed’, that is manifested and put to use, by the firm itself in its daily operation, and the collective sensory and cognitive powers possessed by firms often far exceed the similar powers of any living person (cf. Nelson and Winter, 1982: 59–65, 99–107).

Far from a problem, for Lazonick human asset specificity achieved by consent is highly valuable, for two reasons. Because it represents greater experience and more nuanced communication and cooperation within the routines, it ‘unbounds’ human rationality by supplementing it with a more expansive collective rationality. And because sustained cooperation in a mutually agreeable relational contract fosters cordiality, teamwork and trust within firms, it relaxes

**Routines and capabilities**

Lazonick’s work is an early venture into capabilities theory, for which internal routines and procedures, ‘habitual pattern[s] of behavior embodying knowledge that is often tacit and skill-like’ (Langlois and Robertson, 1995: 1), are the key to understanding the origin and development of firms. A firm’s capabilities, the skills, experience, and knowledge it demonstrates in its daily operation, ‘consist largely of the ability to perform and sustain a set of routines’ (Nelson and Winter, 1982: 142) and, along with the routines themselves, change over time as new participants bring new knowledge and skills to the contract and the firm ‘learns by doing’ in a competitive environment. Firms as such are understood as de facto ‘truces’ or ‘treaties’ (ibid.: 107–112) that mediate divergent individual interests and induce participants to do their part in the routines despite their differing interests. Repeated performance of the routines builds trust and reinforces the truce, which itself becomes an aspect of the firm that simultaneously enables it to operate and inhibits change in its established routines, the latter in time reducing the firm’s ability to respond flexibly to new profit opportunities (Langlois and Robertson, 1995: 101–108).

In this perspective, firms arise to create and coordinate the capabilities needed to exploit new opportunities, and expand or contract as necessary to align capabilities with opportunities in changing environments. Their boundaries are determined primarily by the relative costs of developing needed capabilities internally, which turn largely on technology, existing capabilities, and the flexibility of routines, or, if possible, contracting to buy them from other firms, which generates Williamsonian, ‘static’ transaction costs. Because the latter diminish as firms gain experience with them, they are eventually eclipsed by the ‘dynamic’ transaction costs of a firm’s not having the capabilities it needs when it needs them, so that, in the long run, the firm’s scope of operations depends almost entirely on the relationship between its own capabilities and those of other firms. In time, as knowledge diffuses and new technologies are assimilated, all firms become ‘more capable,’ so there is generally flux in the organization of industries as capabilities shift among old and new firms (Langlois and Robertson, 1995: 20–45). Indeed, one of the virtues of the capabilities approach is that, unlike Williamson’s, it suggests not just why firms expand, but also why they contract (cf. Langlois, 2003).

Joined to the complementary contractual theory of Macneil and Williamson, the capabilities approach is a powerful analytical tool for understanding the historical development of firms. Like the contractualists, it maintains that the firm’s working rules are the firm, and like Hayek (1945) and his Austrian successors, it emphasizes the production and use of knowledge in social orders
and expands their cognitive perspective from the individual to the firm itself as a sensing, deliberating, and acting agent. It advances the evolutionary program of both traditions by making clearer exactly what is evolving and how the selection mechanism works. And it plausibly suggests why existing firms grow, shrink, or stand pat in imperfect markets and how they evolve the internal procedures and routines that enable them to make the decisions and take the actions that this requires. But it does not explain the firm’s origins, how human decisionmakers might come to take the risk of creating the firm in the first place. For at the moment the firm is first created, before anyone has had a chance to see how it actually functions, how it performs in competition, or to what kinds of situations it might have to respond, it is very hard for potential contractors to see clearly enough what the properties of the various alternatives are to make the organizational choices that both Williamson and the capabilities theorists demand of them. Given this inevitable uncertainty, the decision of the founding participants to surrender their autonomy and commit themselves to a relational contract must turn on factors other than their confident knowledge of the consequences for them in the future of choosing between market and contract now.

**Entrepreneurship**

The final piece to the puzzle lies in the Austrian theory of entrepreneurship, particularly the idea of *cognitive leadership* proposed by Witt (1999; cf. Yu, 1999: 29–31). Following Kirzner (1973), Austrians link the creation of firms to the entrepreneurial *discovery* of new profit opportunities and the entrepreneur’s need for the close cooperation of other people to exploit them. Discovery is not simply stumbling upon objective information that already exists ‘out there’, waiting to be found. It is the product of active imagination and personal judgment, of seeing how purposes that are not being served by the existing deployment of resources or ways of organizing production might be served, and profits for the entrepreneur and her associates earned as a result, by doing things differently. Judgment about the existence of such opportunities and the best way to exploit them may not be forthcoming at all, and where it is, it often proves mistaken, a situation competition exposes soon enough. But the willingness of entrepreneurs and their associates to assume the risk of acting on their vision of how things might be done and suffer the consequences of error is an indispensable element of the market process. Whether the new enterprise succeeds or fails, every entrepreneur contributes through the results of her experiment to the creation of new knowledge about how (or how not) to advance human interests by reorganizing the way production is carried out, knowledge that would not otherwise be available to anyone and whose dispersion through the market enables production to be adapted quickly and effectively to changing conditions.

Statements of the capabilities model like Lazonick’s come close to reifying the firm as a supraindividual being that acquires knowledge, makes decisions,
and acts in the market through the operation of its routines. But the Austrians’
identification of the entrepreneur’s personal leadership as the vital force that gives
birth to the relational contract enables a richer contract-capabilities model to be
constructed on a firmly individualist foundation. Austrians portray the firm as a
consensual, planned production order that comes about when an entrepreneur
succeeds in persuading others to cast their lot with her new way of conceiving
the production of some good, so that abandoning free agency and associating
in the firm to execute the entrepreneur’s plan serves their interests as well as
the entrepreneur’s (Adelstein, 2005). The novelty of the entrepreneur’s proposal
and the unknowability of its fate in the market means that those who take
this risk do so more as an act of faith than a rational choice, a belief in the
person of the entrepreneur and the ‘rightness’, for them, of joining their interests
with the interests of the entrepreneur in the contract she envisions. So it is the
entrepreneur’s act of cognitive leadership, offering a persuasive, inspiring vision
of a new way to produce, that provides a provisional sketch of how the new firm
will work, induces agreement to the relational contract that constitutes it, and
sets the evolution of its collectively defined capabilities in motion. Once a group
of people consent to join an entrepreneur in a contract to organize production
according to her plan, what had once been the entrepreneur’s necessarily vague,
speculative vision of how a hypothetical firm might work becomes a real firm
that must mobilize people to act and compete under its name. The founding
contractors cannot anticipate all, or even most, of the circumstances to which
the agreement will have to apply, though the character of the relationships they
establish at the outset will strongly condition the future development of the
contract. But from the first day of operations, daily interactions within the firm
begin to create the distinctive, constantly changing routines, procedures and
organizational culture that are embodied at every moment in the behavior of the
changing roster of contractors.

This is the sense in which firms have an existence of their own, independent
of their transient participants and indefinitely sustainable as new parties to the
contract replace the old over time. When we say that a firm ‘acts’ in a certain
way, we mean that specific individuals do specific things in the context of the
contract that constitutes their firm that the contractors (and others) agree to
describe as actions of the firm. And though firms exist only in the minds of
people, in operation they possess a quality I call physicality. At every moment
of their existence, firms are physically manifested in the performance of their
participants in the roles and relationships of the contract at that
moment. This physicality is not complete, for human intentionality continues to
condition the actual performance of the firm: were the roles and relationships
of an existing contract to be exactly preserved but all the current participants
replaced by other people, the behavior of the resulting social actor, however
one might describe it relative to the original firm, would surely differ from the
original, much as the actual performance of a symphony varies from orchestra
to orchestra. But that the firm’s routines (or the symphony) are in fact being performed is an outwardly physical phenomenon, visible to any observer.

**Emergence and individualism**

An essential aspect of the firm’s constituent routines is that they are emergent phenomena: they arise not from the actions or qualities of the contractors as isolated individuals, but from the interactions between them structured by the relational contract that constitutes the firm. They entail novelty, in that the routines display properties that cannot be inferred or predicted solely from the isolated behavior of the participants, and what Hodgson (2003: 164–165) calls *reconstitutive downward causation*, in that they simultaneously result from the behavior of the firm’s human participants and in part cause that behavior through the effects they have on the perceptions and intentions of the participants. But the contract-capabilities model is fully consistent with a reasonable individualist insistence that all social phenomena be ultimately traceable without remainder to the actions of living people. As Hodgson (2007) points out, social phenomena are social precisely because they involve interactive relations among individuals, so no explanation of such phenomena can be based solely on the behavior of individuals in isolation. A reasonable individualism thus cannot demand that social phenomena be explained entirely in terms of individuals *alone*, only that they be explained entirely in terms of individuals and the interactions among them, a formulation that leaves ample room for emergent routines and the reality of firms as social actors, so long as every action of a firm *supervenes* on the actions of its living participants. This means that the actions of the participants completely determine the actions of the firm, such that any change in the actions of the firm necessarily involves a change in the actions of at least one of the participants (McLaughlin and Bennett, 2008).

The contract-capabilities model clearly meets this requirement. The indispensable role of the entrepreneur in inducing the relational contract in the first instance means that a firm cannot come into existence other than through the interaction of the entrepreneur and the other founding participants; though in appropriate circumstances, firms may play both these roles, as when a consortium of existing firms contracts to create a new one. Once a firm begins operation, its existence at every moment requires the ongoing consent of its human participants to the roles and relationships created by the contract and manifested in the firm’s routines. Actual performances of the routines, the things we call ‘acts of the firm’, are nothing more or less than the actual performances of specific human participants in those roles and relationships at that moment. This supervenience means that firms cannot act, or exist, independently of living men and women. Were all the firm’s participants to disappear in a flash, the firm would disappear with them, even if its buildings and machines remained behind. Firms are created by and reside within the minds of living people, and act in the world solely through the intentions, actions, and interactions of people
who have beliefs and attitudes about firms and the structure firms lend to their relationships with other people.

3. Are firms real?

Relational contracts as social facts

Assume that there is an external reality ‘out there’, of which social objects like firms might be a part, and that it consists in true statements about what Searle (1995: 2) calls brute facts. These are features of the world whose existence is independent of any human state of mind (or intentionality) regarding them: Jupiter orbits the sun, there is snow on Mount Everest, and hydrogen atoms have one electron are all brute facts. As a test of claims to reality, social or otherwise, assume further that every successful claimant must ultimately supervene on brute facts, so that changing anything that really happens means changing some brute facts. Firms, I have said, are ideas, more specifically relational contracts, and thus manifestly not brute facts. Indeed, given the nature of their subject, social scientists almost never confront brute facts, since almost nothing they study is independent of human intentionality. Can ideas be real?

Searle’s ingenious answer begins by dividing the way the world really is into two ontological boxes. Brute facts are ontologically objective, true no matter what anyone thinks, but some features of the world as it is are ontologically subjective, true only because of what people think (this paper is a dollar bill). How we regard or judge such things is an epistemic matter, and here too there are two possibilities. We may regard aspects of reality as epistemically objective, taking them to be real whatever any individual’s attitudes or feelings toward that reality might be (apples cost a dollar), or as epistemically subjective, expressing some evaluative judgment of it that is true, if at all, only because someone thinks it is (apples are too expensive). Epistemic objectivity is tricky, a matter of social ‘consensus’. It signifies a reality that ‘exists independently of my representations, though not independently of all representations’ (Hoover, 2009: 396) – this dollar bill is money, whether I think it is or not, though if enough people come to agree that it isn’t money, then it isn’t, at least for them. So if the statement $X$ is a firm refers to a feature of the real world, two things must be true. First, since the existence of $X$ is clearly ontologically subjective, ontologically subjective things like $X$ must be shown to be real, reducible to brute facts. Second, the real thing $X$ must epistemically objectively be a firm; some critical mass of relevant people must agree to think of $X$ as a firm and act as if it was, however they might feel about $X$.

On the first question, Searle is admirably explicit. He takes ‘the atomic theory of matter and the evolutionary theory of biology’ as given, ‘no longer up for grabs’ (1995: 6), and though he does not use the term, he appeals to what Davidson (1980: 214) calls the supervenience of mental states on the physics
of brains to draw a straight line from the brute facts of physical reality to the ontologically subjective reality of intentional states like \(X\).

We live in a world made up entirely of physical particles in fields of force. Some of these are organized into systems. Some of these systems are living systems and some of these living systems have evolved consciousness. With consciousness comes intentionality, the capacity of the organism to represent objects and states of affairs in the world to itself. (1995: 7)

That there can be no change in intentional states without some change in the brute facts of the brain is not an uncontroversial position, though it is widely held and prominently defended by Kim (1993: 53–78, 131–160). I adopt it here as more persuasive than the alternative; if intentionality does not fully depend ontologically on the physics of brains, it is very hard to say clearly and plausibly on what else it could possibly depend.

Intentionality is a capacity of individual minds, so addressing the second question requires some alignment of individual intentionalities to produce the social consensus that bestows epistemic objectivity on the statement \(X_{safir m}\) and makes the firm real. Searle’s solution is collective intentionality, by which he means not the intentionality of a collective being but only that a group of individuals share particular beliefs, desires, or intentions, and more specifically ‘cases where I am doing something only as part of our doing something’. He illustrates this with football players, whose intentions are to carry out their unique part of a complex play only as part of the team that is executing the play, and who would not do what they were doing were this not the case. The allusion to the ‘we consciousness’ of team players is apt, as the crucial element of collective intentionality is ‘a sense of doing (wanting, believing, etc.) something together, and the individual intentionality that each person has is derived from the collective intentionality that they share’ (1995: 23–25, emphases in original).

But collective intentionality often takes much milder forms, as when you and I decide to take a walk together, and in this sense it is ubiquitous. Once it exists, for Searle, a social fact is created for the individuals who share it: the players’ doing their parts creates the social fact that \(the\ team\ executes\ the\ play\), and our decision to take a walk together creates the social fact that \(we\ are\ walking\ together\) (1995: 26). I would make explicit collective intentionality’s ‘they consciousness’ as well: even people not part of the team can be part of a larger consensus on the social fact itself, that the team exists and \(it\ is\ executing\ the\ play\), that those people across the street are walking together. In this way, the social fact becomes the same fact for ‘everyone’, and thus epistemically objective.

Given the kind of collective intentionality I have described at work in them, relational production contracts, the \(X\) of my story, clearly are real social facts that invoke the collective intentionality of both contractors and others and through this intentionality constitute ‘downward causes’ of the perceptions and acts of the people who share it. But social facts only exist, they do not necessarily act.
One would rarely say that *those people walking together* did something as a social unit. But football teams and firms are social facts that *do* act in their own right. To become a social actor, a social fact must also be an *institutional fact*.

**Firms as institutional facts**

Social facts become institutional facts with the creation of a specific collective intentionality that Searle (1995: 43–51) summarizes as *X counts as Y in C*, where *X* is some object, *Y* is a symbolic status that the collective intentionality assigns to *X* so that *X* can perform a function for the group, and *C* specifies the context in which *X* will be recognized as *Y*. Once this is done, *X* is *Y* becomes an institutional fact, not just for the group for whom the ‘status function’ is performed but with the same substance for outsiders, so the fact becomes part of a collective intentionality that includes ‘everyone’. Thus, we all agree to treat a specific kind of paper as money, so we can use it to facilitate exchange in the context of commerce. As noted, the ‘we all’ is tricky, and context matters – a ruble won’t buy much in Paraguay, though it is an institutional fact for everyone that rubles are money in Russia. But the idea is clear enough, and the present account can easily be reduced to it: ‘we all agree that when people perform in accordance with a certain relational contract (*X*), we will think of them and treat them as a unitary body called a firm (*Y*), so they can perform more effectively in producing goods for sale (*C*). However this collective intentionality is created, and among whom, once it exists, the relational contract in performance becomes a particular kind of social fact, one that ‘everyone’ will henceforth treat as an active social unit.

This differs from Searle’s recent account of corporations (2005: 15–16) in two particulars. First, as noted, Searle believes that for firms to be institutional facts, they must be legally recognized. More importantly, he is troubled by the idea that firms themselves are ideas, and thus invisible. Earlier, he had insisted that for institutional facts to meet the test of reality, every *X* had to be a ‘physical realization, some brute fact... on which we can impose [the] status function. Thus there are no institutional facts without brute facts’ (1995: 56), though brute facts will often ‘not be manifested as physical objects but as sounds coming out of peoples’ mouths or as marks on paper – or even thoughts in their heads’. It is, moreover, misleading to think of institutional facts as social objects, ‘independently existing entities on analogy with the objects studied by the natural sciences’. In this case, ‘process is prior to product. Social objects are always, in some sense... constituted by social acts; and in a sense, the object is just the continuous possibility of the activity’ (ibid.: 35–36, emphasis in original).

I think this account is correct. The coupling of collective intentionality and the primacy of performance, ‘we consciousness’ and physicality, is what creates social actors like football teams and firms. As firms, professional football teams
are institutional facts that exist and act all week, their actions consisting in the visible performance of a broad range of routines by many people, most of them not players. On the weekend, the players on the field become a second institutional fact, in which their executing visible routines within a physically marked space according to specific rules becomes the team playing the game. Beyond the bruteness of the fact of collective intentionality, there is a second, physical manifestation of the institutional facts we call the acts of the firm or team in the visible performance of characteristic routines by real people in real time. But Searle now views firms as ‘free-standing Y terms’, such that ‘[t]here is indeed a corporation as Y, but there is no person or physical object X that counts as Y’ (2005: 15–16). Y is not the firm itself, but a set of status roles to be performed by individuals (we all agree that X is treasurer). This view, like my own, reduces the acts of the firm to the acts of individuals in their status roles but, like the nexus theorists, it does this by removing the firm as such from the collective intentionality that creates it and thus from social reality itself. I believe this is mistaken. A characterization of a firm’s act that described only the actions of the individual participants performing their parts in a complex routine would omit both the collective intentionality and the emergent, coordinated performance that together prompt us to agree that a firm is acting and thus make that institutional fact real. It is true that the performance of routines cannot be completely reduced to the brute facts of outward physical activity; human intentionality conditions the actual performance of any routine. But the brute fact of collective intentionality also creates a shared perception that the routines are being performed by a firm, and with it the institutional fact of the firm itself as a distinct social actor.

4. Conclusion

I have argued that firms are relational contracts from whose performance emerge their distinctive routines and capabilities, that the sense of ‘the firm’s acts’ that this implies is grounded in the behavior of living people, and that firms are real social actors. Though the questions with which we began are far from answered, three paths for further work seem open. As Foss (2003: 196) notes, capabilities theorists have not pursued the logical links between the intentionality and behavior of individuals in routines and the operation and evolution of the routines, and a sustained microanalytics of capabilities is needed. Progress in this might enable the example of firms to enter the philosophical debate over emergence, reduction, and supervenience (cf. McIntyre, 2007; Hoover, 2009: 390–394). And this in turn might inform inquiry into how routines and procedures actually work in real firms, how they structure working relationships and the effects these have on participants and others, and what political or moral qualities firms might possess as a result.
References


Firms as social actors 349


