Division of Research
School of Business Administration

Revised April 1987

THE INSTITUTIONAL BASIS FOR THE FIRM

Working Paper #473R

Scott E. Masten*
University of Michigan

*I would like to thank Keith Crocker, James Krier, Frank Easterbrook, Michael Rosenzweig, George Siedel, Ted Snyder, Valerie Suslow and Martin Zimmerman for helpful comments on an earlier draft of this essay. All the usual disclaimers apply. Research support from the University of Michigan School of Business Administration is also gratefully acknowledged.

THE INSTITUTIONAL BASIS FOR THE FIRM

Scott E. Masten
Graduate School of Business Administration
University of Michigan
Ann Arbor, MI 48109

Fifty years after the publication of Ronald Coase's seminal deliberations on the subject, economists have yet to reach consensus on the nature of the firm. To be sure, the literature offers numerous reasons why transactors might vertically integrate--among them, technological interdependencies, factor distortions, information asymmetries, transaction costs, and risk-sharing and control considerations[1]--each of which has contributed in some way to our understanding of the problems associated with vertical exchange and the motivation to internalize production. But these contributions have also left some central questions unanswered. Theories of vertical integration have, for example, tended to emphasize the failings of market exchange without adequately explaining how "internalization" overcomes those failings. Upon integration, information previously impacted suddenly becomes available, opportunism desists, and input misallocations disappear. But the mechanisms through which vertical integration accomplishes these

---

[1]See, respectively, Chandler, 1966; Vernon and Graham, 1971, and Warren-Boulton, 1974; Arrow, 1975, and Crocker, 1983; Coase, 1937, Williamson, 1975, 1979 and 1985, and Klein, Crawford and Alchian, 1978; Carleton, 1979; and Simon, 1951. Under the category of control, I would also add the recent paper by Grossman and Hart.

feats remain obscure. Moreover, this faith in the salutary properties of internal organization belies the persistence of exchange outside the firm: If a firm can do everything a market can and more, what explains the failure to internalize all transactions?[2]

The lack of authoritative answers to such questions reflects an even more basic controversy over what, if anything, distinguishes organization within a firm from external or market exchange. While many economists continue to regard the firm as a distinct institution, usually ascribing to it some superior control, information or adaptive properties, others reject the notion that any unique governance advantages accrue to integration, noting that neither human nature nor technology or information are altered by the purely nominal act of "internalization." For the latter, the word firm is merely descriptive, a collective noun denoting a particular cluster of otherwise ordinary contractual relationships.

Our view of the firm and its role in economic organization turns ultimately on the resolution of this controversy. Accordingly, the objective

---

[2]As Grossman and Hart elaborate (p. 692-3):

> [I]f vertical integration always reduces transaction costs, any buyer A and seller B that have a contractual relationship should be able to make themselves better off as follows: (i) A buys B and makes the previous owner of B the manager of a new subsidiary; (ii) A sets a transfer price between the subsidiary and itself equal to the contract price that existed when the firms were separate enterprises; and (iii) A gives the manager of B a compensation package equal to the profit of the subsidiary.

Given that administrative intervention could then be adopted "selectively--which is to say only upon showing expected net gains" (Williamson, 1985, p. 133), "how can integration ever be strictly worse than nonintegration; that is, what limits the size of the firm" (Grossman and Hart, p. 693). Also see in general, Williamson, 1985, pp. 131-37.

of this paper is *not* to propose a theory of vertical integration or to explain when one transaction should be internalized and another left on the market. As difficult and important as those problems are, they have already received vastly more attention than the logically antecedent issue of what constitutes a firm. Rather, I wish to explore the question of whether it even makes sense to talk about the firm as a distinct organizational form. Does internal organization exhibit special properties giving the notion of the firm constructive meaning, or is the designation *firm* really just descriptive of a set of commonly observed but otherwise unexceptional contractual relationships?

1. The Theory of the Firm and Its Critics (The Controversy).

The ambiguity regarding the nature of the firm in the economics literature can be traced back at least to Ronald Coase's original contribution. The language Coase adopts often suggests a conception of the firm as a distinct alternative to market exchange (1952, p.333-4):

> Outside the firm, price movements direct production, which is co-ordinated through a series of exchange transactions on the market. Within a firm, these market transactions are eliminated and in place of the complicated market structure with exchange transactions is substituted the entrepreneur-co-ordinator, who directs production. It is clear that these are *alternative methods* of co-ordinating production (emphasis added).

But this apparently discrete view of the firm is subsequently qualified: "Of course, it is not possible to draw a hard and fast line which determines

whether there is a firm or not. There may be more or less direction" (p. 337 fn. 21). Rather, the firm is characterized by the existence of a central contracting agent and a contract "whereby the factor, for certain remuneration (which may be fixed or fluctuating), agrees to obey the directions of the entrepreneur *within certain limits*" (pp. 336-7, emphasis in original). Hence, for Coase, the distinction between the firm and the market appears to be more a matter of degree than kind, the existence of the firm depending on the *amount* of discretion accorded the manager in the contract.

Herbert Simon's early treatment of the employer-employee relationship offers a similar view. Like Coase, Simon defines the employment contract in terms of the authority permitted the employer to control the allocation of effort in response to changing conditions (1951, p. 294):

> We will say that W [a worker] enters into an employment contract with B [a boss] when the former agrees to accept the authority of the latter and the latter agrees to pay the former a stated wage (w). This contract differs fundamentally from a sales contract--the kind of contract that is assumed in ordinary formulations of price theory. In the sales contract each party promises a specific consideration in return for the consideration promised by the other. The buyer (like B) promises to pay a stated sum of money; but the seller (unlike W) promises in return a specified quantity of a completely specified commodity.

Thus, the essential feature of Simon's employment relationship is the discretion left to the employer to direct some dimension of the employee's behavior. Under this definition, however, the employment relationship is analytically indistinguishable from any contract in which one party is empowered to alter some aspect of performance unilaterally. An example would be a fixed-price, variable-quantity contract in which the buyer has the "authority" to determine the volume of trade under the agreement and can thus "direct" the production level of the seller. Such arrangements are not at

all uncommon in long-term contracts. But although they conform in a technical sense to Simon's definition of an employment transaction, the relationship between the buyer and seller in such contracts would not generally be considered that of employer and employee. At best, the distinction between employee and supplier is again a matter of degree. Only the details and not the type of contract entered separate an employment from a commercial transaction, and neither the label *firm* nor *employee* has any force beyond the provisions explicitly adopted in the contract itself.

While the conception of the firm offered by both Coase and Simon is essentially descriptive, more recent treatments of the vertical integration decision have clearly taken a more constructive view. Transaction-cost economics, in particular, has emphasized the distinctive features of internal organization and adopted as a central precept that the analysis of organizational form "mainly involves a comparative institutional assessment of discrete institutional alternatives" (Williamson, 1985, p. 42). Internal organization, in the transaction-cost framework, somehow alters the underlying environment or the rules under which the transactors operate. Among the advantages believed to arise under integration are better access to information and greater managerial control and flexibility. In *Markets and Hierarchies*, for instance, Oliver Williamson credits the firm with superior auditing and conflict resolution properties: Internal auditors are believed to have superior access to the information necessary for decision-making; and internal dispute resolution mechanisms are perceived as more responsive to changing circumstances than court adjudication, encumbered as it is by strict procedural requirements (1975, pp. 29-30). Changes in the information structure of exchange have also formed the basis for more formal models of

the vertical integration decision, such as in Kenneth Arrow's (1975) analysis that treats integration as "essentially a way of acquiring predictive information" (p. 176; also see Crocker, 1983).

Although the views expressed by these authors conform to common perceptions of the role and advantages of the firm, the source of those advantages remain obscure. Where does the authority of management to direct production or settle disputes come from, and why are employees less able to hide or distort information than independent contractors? Armen Alchian and Harold Demsetz offered a major criticism of authority-based theories of the firm in their 1972 article (p.777):

> It is common to see the firm characterized by the power to settle disputes by fiat, by authority, or by disciplinary action superior to that available in the conventional market. This is delusion. The firm does not own all its inputs. It has no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting between any two people. I can "punish" you only by withholding future business or by seeking redress in the courts for any failure to honor our exchange agreement. That is exactly all that any employer can do. He can fire or sue, just as I can fire my grocer by stopping purchases from him or sue him for delivering faulty products.

More recently, David Evans and Sanford Grossman have adopted and extended this line of argument. Targeting the transaction-cost literature, they criticize what they consider an overly sanguine view of internal organization (p. 115):

> A failure to contract leads to chaos. A failure by employees of different divisions of an integrated company to coordinate activities also leads to chaos. Advocates of the transaction-cost theory of integration appear to believe that contractual problems evaporate upon merger. Indeed, transactions within integrated businesses are afflicted by the same contractual problems as transactions between unintegrated businesses.

The possibility that integration relieves information asymmetries is explicitly rejected: "Common ownership creates neither new information nor expertise. Nor does it create new auditing opportunities. Nonintegrated companies permit arbitrators to audit their operations when there is a dispute between them" (pp. 119-20). And the factors that undermine market exchange, especially opportunistic behavior, are ubiquitous: "It is strange that the transactions cost literature assumes that the mere act of integration transforms selfish humans into selfless ones whose only goal is their company's welfare. Common ownership, unfortunately, eradicates neither indolence nor dishonesty" (p.121).[3]

In sum, these authors deny the existence of administrative solutions to contractual failures, asserting that the same transactional frictions confront employers as independent contractors. The firm, at least in governance respects, is thus no more than a coalition or "nexus" of contractual relationships, and the choice faced by transactors is only among the details to include in the contract. As Steven Cheung puts it, "It is not quite correct to say that a 'firm' supercedes 'the market.' Rather, one type of contract supercedes another type" (1983, p. 10). "Thus it is futile," he continues, "to press the issue of what is or is not a firm. . . .The important questions are why contracts take the forms observed and what are the economic implications of different contractual and pricing arrangements" (p. 18).[4]

---

[3]Also see Grossman and Hart, p. 695, and Cheung, p. 6. The last of these criticisms is inaccurate. The transaction-cost literature has consistently and explicitly recognized subgoal pursuit by employees as a problem within firms. See, for example, Williamson, 1975, pp. 117-131.

[4]Cf. Jensen and Meckling, 1976, pp. 310-1, and Alchian and Demsetz, p. 778.

2. A Legal Basis for the Firm.

In large measure, the debate over the nature and role of the firm just described has taken place in a vacuum, on one side, the institutionalists appealing to an intuitive sense that integration exhibits special properties and, on the other, the nominalists insisting that the problems associated with market exchange are immutable except only to the extent that a single, universal set of contracting principles can be brought to bear on the behavior of the transactors. Neither side has given careful consideration to the premises that underlie the dispute, however. Is there a sense, for example, in which all economic relationships are contractual, or do institutional designations sometimes carry with them certain rights, responsibilities or authority? And if institutions do exhibit special properties, what is the source of those distinctions?

Part of the dissonance surrounding this issue stems from the many uses of the terms *contract* and *institution*. Both economists and lawyers often use the term contract in a broad sense encompassing both 'agreement' and 'transaction' in meaning. Accordingly, a promise to deliver steel in 60 days and a purchase from a grocer are often categorized equivalently as contractual. Under this expansive usage, the relationships comprising a firm would also accurately be termed contractual.

There is, however, another more restrictive connotation to the term contract, and that is as a formal, legal commitment to which each party gives express approval and to which a particular body of law applies (cf., e.g., Klein, 1980, p. 358; and Clark, pp. 60-61). Not all transactions and agreements are contractual in this way. For example, under

this definition, it is meaningful to distinguish between a simple exchange and a bona fide contract. In this sense, "cancelling an order" differs from "breaching a contract," to use Stewart Macaulay's dichotomy (1963, p. 61). And the contractual nature of the firm becomes less apparent. Employees with special skills or those represented by a union may have an explicit contract with management, but the typical employment relationship is "at will," that is, "a simple exchange of a day's work for a day's pay" (see, e.g., Note, 1983, pp. 449-50, fn. 5).

This brings us to the meaning of *institution*. Here again the term has broad and narrow applications. The broad sense is essentially descriptive; it includes all modes or conventions for transacting as well as the organizations embodying them and is aptly applied to any pattern of behavior or collection of relationships that occurs with enough frequency to merit a label. Markets and credit in their generic senses are examples of economic institutions under this definition.

In its narrower sense, *institution* denotes a more established arrangement, a relationship or organization whose existence or boundaries are defined and administered by an exogenous authority. Membership or participation in such an institution typically confers particular rights or responsibilities and establishes the rules and procedures that govern the conduct of the transactors. Congress is governed by the constitution, an agency of government by its enabling legislation, and a corporation by its charter and the laws of incorporation. The relationships among members of these bodies are all contractual in the broad sense, but the organizations themselves represent distinct institutions in that term's narrow sense by virtue of the peculiar rules and procedures that regulate admissable behavior within them.

By the same reasoning, the restrictions placed on acceptable provisions, enforcement mechanisms and procedures by the broader system of contract law make a *contract* in the narrow sense also an *institution* in the narrow sense. The definition and regulation of allowable behavior by a dominant authority or set of institutions is what gives institutional designations of this type constructive as opposed to merely descriptive meaning.

The question of the nature of the firm can now be reformulated in these terms. To say the firm is contractual in the broad sense is tautological. On the other hand, the fact that membership in the firm rarely requires an explicit contract invalidates the definition of the firm as a nexus of contractual relationships in the narrow sense. The real issue is not whether the relationships are contractual, however, but whether the firm is an institution in its constructive as opposed to merely descriptive meaning. Does the law, for example, distinguish between commercial and employment transactions in a meaningful way? In particular, are any rights, authority or responsibilities given to employers or employees (or any limitations placed on them) that are not available in analogous form to commercial transactors? Whether internal organization represents a distinct alternative to market exchange depends on whether such differences exist and how they influence the incentive structure within and between firms.

Answering these question involves examining in some detail the rules and standards applied to market and integrated transactions as defined by the broader legal and political institutions that regulate economic activity. The nature of the firm is thus not subject to a priori reasoning but is ultimately a question of fact: Are there mechanisms or sanctions available in employment transactions that are not similarly available to independent

contractors? In this paper, I examine the status of the employment relationship in the legal system and compare it to corresponding doctrines in commercial contract law to see whether rules of law establish an institutional basis for the traditional view of the firm in economics.[5]

Duties and Obligations: Ironically, economists have either downplayed or rejected outright the role of the law in defining the firm, divorcing the economic concept from the "legal fiction" (see, Jensen and Meckling, 1976, pp. 310-311; and Rubin, 1978, p. 225).[6] Even a cursory examination of the case law governing the relationship between employers and employees, however, reveals a set of obligations and responsibilities that are indeed unique to

---

[5]In what follows, I have attempted to cite the most common rules of law as reported in original cases, legal treatises and case books. While some exceptions and cross-jurisdictional differences in either the laws or how they are applied exist, the general principles cited are fairly standard in American common law.

In an earlier draft I had also relied on a number of legal encyclopediae for statements of the general rules governing employment transactions. I was advised by legal counsel (see the acknowledgements), however, that such sources were considered less than authoritative by legal scholars. Although for expository purposes I have sometimes found it convenient to retain the language of these sources, in every case I have supported the quotations with supplementary citations from more authoritative sources.

[6]Jensen and Meckling, to whom the phrase "nexus of contractual relationships" is attributable, offer seemingly contradictory statements on the importance of law to organization. On one hand they assert, "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" (1976, p. 311). In a footnote following shortly thereafter, however, they observe, "This view of the firm [as a nexus for contracting relationships] points up the important role which the legal system and the law play in social organizations, especially, the organization of economic activity.... [Various] government activities affect both the kinds of contracts executed and the extent to which contracting is relied upon. This in turn determines the usefulness, productivity, profitability and viability of various forms of organization" (p. 311, fn. 14).

employment transactions and which often coincide precisely with the traditional emphasis in economics on the information and authority advantages of internal organization. Upon entering an employment relationship, for example, every employee accepts an implied duty to "yield obedience to all reasonable rules, orders, and instructions of the employer."[7] The importance of authority in employer-employee relationships is given further weight by the criteria courts use to adjudicate disputes over the nature of a particular transaction. The overriding consideration expressed by the courts in such cases is the control exercised by an employer and, especially, whether the latter is concerned with the manner in which work is performed and not solely with its outcome: In evaluating a transaction, "The first, and seminal, inquiry is whether the alleged employer... has the *right* to control...the details of the alleged employee's work;"[8] whereas, "an 'independent' contractor is generally defined as one who in rendering services exercises an independent employment or occupation and represents his employer only as to the results of his work and not as to the means whereby it is to be done...."[9] Hence, the traditional emphasis in economics on the authority of management to direct the efforts of employees is at least nominally supported by the law governing employment transactions.

Comparison of commercial and employment law also provides support for the informational advantage commonly attributed to internal organization in the

---

[7]53 Am Jur 2nd section 97; Restatement of Agency (2nd), sections, 2, 220, 385.

[8]Pitts v. Shell Oil Co, 463 F.2nd 331 (1972), emphasis in original; Restatement (Second) of Agency, section 2.

[9]56 Corpus Juris Secundum (hereafter CJS) 45; Restatement of Agency (2nd), sections 2, 14.

traditional view of the firm. In commercial transactions, laws regarding the transfer of information are fairly liberal. As a rule, "one party to a business transaction is not liable to the other for harm caused by his failure to disclose to the other facts of which he knows the other is ignorant and which he further knows the other, if he knew them, would regard as material in determining his course of action in the transaction in question."[10] The most prominent exception to that rule concerns the existence of a fiduciary relationship between the transactors such as that of principal and agent or employer and employee.[11] Where such a relationship exists, the agent is legally obligated to reveal relevant information to the principal.[12] This exception to the general rule of nondisclosure has also become codified in the law of master and servant as the employee's duties to disclose and account, under which a subordinate is obliged "to communicate to

---

[10]Restatement (2nd) of Torts, Sec. 551; also see Restatement of Contracts (2nd), Sec. 303.

[11]"Agency, properly speaking, relates to commercial or business transactions, while service has reference to actions upon or concerning things. Service deals with matter of manual or mechanical execution. An agent is the more direct representative of the master, and clothed with higher powers and broader discretion than a servant. ... Agency, in its legal sense, always imports commercial dealings between two parties by and through the medium of another. An agent negotiates or treats with third parties in commercial matters for another....

"A servant is a person subject to the command of his master as to the manner in which he shall do his work, and the master is the one who not only prescribes the work but directs, or may direct, the manner of doing the work" (American Savings Life Ins. Co. v. Riplinger, 60 S.W. (2nd) 115, 249 Ky. 1933, quoted in Dykstra and Dykstra, Business Law, 1969, p. 354). Note also that a servant or employee is always an agent but that not all agents are servants.

[12]See Restatement of Agency, section 381, Restatement of Torts (ibid) and Kronman, 1978.

[his employer] all facts which he ought to know."[13] Again, the law distinguishes between employment and commercial transactions in a way that apparently supports the superior access to information traditionally assumed to accrue to integration in the economics literature.

The obligations of an employee to an employer, moreover, involve much more than simple compliance to managerial directives and disclosure of information. In addition to obedience, an employer has the right to expect loyalty, respect and faithfulness from his employees and that each will "conduct himself with such decency and propriety of deportment as not to work injury to the business of his employer."[14] These duties extend so far as to require the employee to maintain "friendly relations" with the employer.[15]

Criticizing the concept of the firm used by agency theorists, Robert Clark discusses some of the specific implications the requirement of loyalty has for employees.[16] For example, in comparison to the obligations of an independent contractor, the duty of loyalty places larger demands on a manager employed to run a business on behalf of others:

---

[13]56 CJS 67, see also Clark, pp. 71-72, and Michigan Crown Fender v. Welch, 1920.

[14]53 Am Jur 2nd; Restatement of Agency (2nd), sections 380, 387.

[15]Restatement of Agency (2nd), section 380.

[16]Clark's position on the concept of the firm adopted by agency theorists is consonant in many respects with the one advanced here. The present paper goes further, I believe, in its analysis of the controversy over the nature of the firm, in examining the sanctions supporting the duties accruing to employees and the role of termination, and in interpreting the economic implications of the legal distinctions between employment and commercial relationships, especially as they apply to the advantages commonly associated with internal organization.

> The independent contractor usually has relatively fixed obligations under his contract. If the contract does not call for a particular performance, he does not have to do it [whereas w]ith corporate managers, the open-endedness of legally imposed duties is more substantial. . . .Case law on managers' fiduciary duty of care can fairly be read to say that the manager has an affirmative, open-ended duty to maximize the beneficiaries wealth (1985, p. 73).

The duty of loyalty also restricts the ability of a manager to benefit at the expense of owners. "Essentially," Clark continues, "the fiduciary cannot take any compensation from the beneficiaries or any other advantages from his official position (even when doing so does not seem to deprive the beneficiaries of any value they would otherwise get) except to the extent provided in an above-board actual contract..." (p. 73). The rules that apply to the relationship between a owner and manager also apply to master and servant or employer and employee relationships.[17]

Finally, although I have to this point emphasized the employee's duties, employers also assume certain responsibilities upon entering an employment transaction. Along with authority to direct an employee's behavior comes liability under the doctrine of *respondeat superior* for any harms the employee causes to third parties in the course of his employment. "This doctrine," however, "does not apply to the relationship between employer and independent contractor because the employer cannot exercise control over the manner of the contractor's work" (Dykstra and Dykstra, 1969, p. 356). Thus, responsibility for outcomes is assigned to the party ostensibly in control.

---

[17] See Restatement of Agency (2nd), section 25.

Overall, the duties accruing to employment transactions under the law are consistent, at least superficially, with the authority and information properties often associated with internal organization. In fact, the intent of employment law seems to be to make the employee as much as possible an extension of the employer. The duties discussed above, moreover, accrue independently of whether or not they are explicitly contained in a contract.[18] The applicable case law depends on whether the relationship is between independent contractors or employer and employee. In other words, the set of rules that govern the relationship depend on the institution chosen.

Sanctions and procedures: Duties of obedience and disclosure, of course, have force only if mechanisms exist to implement them. Statements such as the following, quoted in Batt's *The Law of Master and Servant*, have tended to undermine such distinctions (1967, p.2, from Mercy Docks and Harbour Board v. Coggins and Griffith (Liverpool) Ltd., [1947] A.C. I):

> "I take no orders from anybody," said the cranedriver: "A sturdy answer," said Lord Simmons, "which meant that he was a skilled man and knew is job and would carry it out in his own way. Yet ultimately he would decline to carry it out in [his employers'] way at his own peril, for in their hands lay the only sanction, the power of dismissal."

If the power of management to enforce its directives rests solely on the threat of dismissal, as this passage suggests, then--a nominal duty to obey notwithstanding--the authority of management is no different from that of an independent transactor engaged in ordinary market exchange. Absent an

---

[18] Restatement of Agency (2nd), sections 220, 225.

express contract to the contrary, the ability to terminate negotiations and any further dealings is also the "only sanction" normally available to commercial transactors dissatisfied with the terms of trade currently tendered. Obviously, the claim of special authority in employment transactions cannot be based on the threat of termination alone. Otherwise, as Alchian and Demsetz maintained, there can be no material distinction between employment transactions and any other simple exchange in a market setting.

The large volume of case law devoted to defining and differentiating between employment and commercial transactions makes it difficult, however, to dismiss the notion that institutional form matters. If courts cared only about the *details* of a relationship as stipulated in a contract when evaluating performance, then the *type* of relationship the parties had entered would be irrelevant. That transactors are willing to spend resources litigating the designation of a transaction is a strong indication that institutional labels do create legal distinctions having practical implications for the transacting parties. This suggests that it may be appropriate to look beyond the duties nominally accruing to employment transactors to a more fundamental set of questions: Is termination really the only sanction or are there other penalties available to employers and employees, and if so, how do these differ from those available to commercial transactors? Is the ability to write and enforce contracts the same in commercial and employment settings or do the criteria for performance and the penalties for breach differ across institutions? In other words, do the circumstances under which an employer can *fire* or *sue* and the remedies and penalties prescribed by the law differ depending on the mode of organization?

Although a fairly well-established body of common law governs both employment and commercial relations and many of the basic rules of behavior apply to both (see below), a closer look at the case law reveals there are indeed a number of differences in the mechanisms and penalties available to commercial and employment transactors and in their application by the courts. For instance, despite the passage from Batt's *The Law of Master and Servant* quoted above, the threat of dismissal is not the only sanction available to employers. An employee who fails to uphold his duties may actually be held liable for damages if that failure were to cause injury to his employer's business. "In an absence of a waiver of the breach, the employer may recover damages from his employee ... for involving his employer in loss through his own negligence or wrongful act..., or *generally for any failure to perform the duties devolving on him under the employment contract*"[19]

Thus, whereas the personal conduct and loyalty of participants in commercial exchanges are strictly a matter of business judgement, the law obligates an employee, with few exceptions, to act in his employer's interest. Disloyalty may, of course, be tempered by reputation considerations in both types of transactions, but formal legal sanctions for such behavior are available only to employers. An independent supplier who recruited personnel from a commercial customer might jeopardize future dealings but could not generally be held legally accountable for the customer's loss. Were the raider an employee, however, the enticement of fellow workers to join a competing concern is likely to constitute a breach

---

[19] 56 C.J.S. 500, emphasis added; Restatement of Agency (2nd) sections 399, 400, 401.

of the duty of loyalty and entitle the former employer to recover damages.[20] This aspect of the law is also what creates common law remedies for insider trading as well as a number of other practices prohibited to employees but open to outside contractors.[21] The important point is that the law entitles an employer to recover damages from a disloyal or uncooperative subordinate and thereby differentiates the incentives of employees from those of independent contractors in a discrete fashion, altering the payoff to uncooperative employee behavior in a way that arguably supports the authority commonly attributed to employers in the theory of the firm.

Similar sanctions support the employee's disclosure duty. Specifically, an employee is legally accountable for any pecuniary losses sustained by his employer as a result of a failure to disclose relevant facts and is liable for damages and the return of any ill-gotten gains from a failure to uphold that obligation.[22] An independent subcontractor bears no such responsibility and is free, among other things, to exploit profit opportunities that arise in the course of the contract's performance.[23] The fact that an employee is less likely to profit successfully from nondisclosure should reduce his incentive to distort or conceal information from his employer.

---

[20]See Frederick Chusid & Co. v. Marshall Leeman & Co.

[21]See, Clark, pp. 74-5.

[22]See Michigan Crown Fender Co. v. Welch. Moreover, "It is not necessary that the employer suffer actual loss before he is entitled to recover ill gained profits from an employee or agent" Byer v. International Paper Co., 314 F. 2nd 831 (1963).

[23]See, for example, Clark, p. 73-75.

The objection made by Evans and Grossman that commercial contractors may employ arbitrators to audit a firm and resolve disputes does not undermine this argument. Ultimately, arbitrators must rely on the provision of information by the disputants. An uncooperative actor (especially one that assessed a small probability of success but wished to buy time) could impede arbitration and conceivably force the issue into court. The differential treatment courts bestow on integrated and independent transactors under the laws regarding disclosure will influence both the prospect and timing of successful access to information. The employee's duty to disclose makes it much more likely that an employer would receive a summary judgement in his favor than would an independent contractor in an otherwise similar situation. The prevailing doctrine would place a larger burden on a plaintiff to demonstrate a right to access the internal records of a nonintegrated defendant. Obviously, the delays and other costs such litigation could impose are likely to create a greater barrier to accessing information from an independent contractor than from an integrated division. In this light, the radical change from asymmetric to costless information that traditional theories associate with integration exaggerates but does not contradict the reduction in institutional barriers to information that prevailing legal doctrine affords internal organization.

The procedures for resolving disputes in commercial and employment transactions also differ--and do so in ways that enhance the flexibility of internal organization. As Williamson points out, managerial directives possess a presumptive validity that is reflected in the rules governing conflicts between employers and employees, even in collective bargaining settings:

> Even where the collective agreement lists certain offenses or the parties negotiate plant rules, management may normally supplement the listed offenses or negotiated rules. Rules prescribed by management are subject to arbitrator review, but they carry a presumptive validity and will be upheld so long as they are reasonably related to achieving efficient operation and maintaining order and are not manifestly unfair or do not unnecessarily burden employees' rights.
>
> Management also is entitled to have its orders obeyed and may discipline employees for refusing to obey even improper orders. Arbitrators almost uniformly hold that an employee must obey first and then seek recourse through the grievance procedure, except where obeying would expose him to substantial risks of health and safety [Summers, cited in Williamson, 1985, p.249].

Disagreements in commercial dealings, by comparison, require "mutual consent before adaptations can be effected" (Williamson, 1985, p.249). This difference may be important. In the time necessary to negotiate a mutually advantageous adjustment, the opportunity to act may well have passed. The potential liability of an employee for failing to obey or delaying performance of a reasonable directive, on the other hand, is likely to encourage immediate compliance, thereby promoting responsive adaptation to changing circumstances.

The extension of liability for the torts of the employee to the employer under the doctrine of *respondeat superior* also affects incentives in employment transactions. First, it motivates employers to monitor employees' activities more closely. And second, to the extent that an employee may seek legal compensation from an employer for personal liability to third parties for actions undertaken at the direction of the employer, it may reduce the employee's reluctance to follow orders by lessening his need to assess the consequences of those actions. Interestingly, while the monitoring function *per se* does not define the firm, as Alchian and Demsetz and others have

suggested it does, the incentives created by the supporting institutional structure is likely to encourage more intensive monitoring within firms than between them--even holding claims on residual income constant.

Thus, there appear to be differences both in the obligations assumed to accrue to transactors in employment and commercial settings and in the sanctions and procedures that support those obligations. In other words, although both employment and commercial transactors have the right to sue, what they can sue for and their expectations of success differ in relation to the transaction's designation and the common law that applies. Again, the legal distinctions examined so far seem to support the authority and information properties often ascribed to internal organization in the traditional view of the firm.

Termination and contracting: Another property often associated with the firm is the greater flexibility to terminate a relationship the doctrine of employment-at-will affords an employer (see, e.g., Rubin, 1978). Employment-at-will empowers either party to an employment transaction to terminate the relationship "for good cause, bad cause, or no cause at all" (see, e.g., Note, *Stanford Law Review*, p.335). The doctrine is more a source of symmetry than contrast between commercial and employment exchanges, however. Commercial transactors engaged in a simple exchange across a market interface also have the right, in the absence of a contract, to discontinue a trading relationship unilaterally. Thus, rather than distinguishing the firm, the doctrine of employment-at-will seems to place employment and commercial dealings (again, absent a formal contract) on an equal footing with regard to termination.

The law also treats employment and commercial contracts similarly in many respects. First, employers and employees, like commercial transactors, can generally contract out of the at-will setting by mutual consent (see Epstein). Stipulated damage provisions, for instance, are generally upheld under the same circumstances in both contracts *of service* and contracts *for services* (see Note, *J. Law Reform*, p. 457). The rules governing employment and commercial contracts of unspecified duration are also similar. While employment contracts of unspecified term provide no guarantee of continued performance (Note, *Stanford Law Review*, p. 345), indefinite term commercial contracts require only "reasonable notification" for unilateral termination (see UCC section 2-309).

Even statutory prohibitions against specific performance in employment contracts offer little distinction in practice. Even though an employee cannot ordinarily be compelled to render services to an employer, specific performance is infrequently applied in commercial settings as well.[24] Rather, the penalties normally imposed for nonperformance of commercial and employment agreements are similar: lost profits versus lost wages "less what can be earned by reasonable effort in similar employment" (Corbin, section 958).

Yet despite these similarities, differences do exist in the treatments courts afford employment and commercial contracts. In particular, the duties discussed above provide an important distinction regarding what constitutes

---

[24]Specific performance may be required if a service involves a special or unique ability. In these cases, however, the contract is considered as being *for services* rather than *of service*; that is, as a commercial rather than employment transaction.

breach of contract in employment and commercial settings. Although only "substantial performance" of the contract is required in either contracts for services or of service, the duties and obligations automatically accruing to every employment transaction provide employers wider latitude to suspend employment contracts: An employment contract may be discharged on the basis of employee indolence, dishonesty, disloyalty or disrespect, among other offenses. Even where employment is protected under a collective bargaining agreement, fighting, insubordination, the use of profanity or abusive language to supervisors, theft, dishonesty, gambling, and the possession or use of drugs or alcohol have all been found to constitute "just cause" for termination of an employment contract (see Steiber and Murray, 1983, p. 323). But although indolence or disrespect may be bad business in commercial transactions, such behavior would not constitute an actionable cause for discharge of a commercial contract.

The burden of proof in contract disputes also supports the employer's authority. Although an employee may contest a claim of unsatisfactory performance, it is the employee's burden to show that his behavior was in fact satisfactory: "The general rule ... is that the employee may question the honesty and good faith of his employer's dissatisfaction, and a feigned dissatisfaction is not sufficient justification to avoid continuation of a contract of employment..., [but] the burden of proof [rests] upon the employee to show that the claimed dissatisfaction is not in good faith."[25] Given the inclusiveness of employees' duties, proving bad faith termination is likely to be difficult in practice.

---

[25]Coker v. Wesco Materials Corp., 308 S.W. 2nd 884.

Such differences in the criteria for discharge of contractual obligations can be interpreted as further contributing to employers' control over the method of production relative to commercial contractors. In comparison to an employee, an independent subcontractor is relatively immune to contract cancellation on the basis of behavior that is largely extraneous to the central purpose of the agreement. The difficulty of proving bad faith termination in employment transactions, moreover, further increases the incentive of employees (relative to independent contractors) to perform satisfactorily and avoid conflicts, especially in cases where the unsatisfactory performance is either marginal in nature or difficult to substantiate in court. In these respects, then, the law serves to make dismissal a more credible threat when made by an employer than by an independent contractor. As a result, an employee under threat of dismissal is more likely to accept managerial redirection--holding reputation considerations constant--than a commercial supplier with common law protections against breach of contract.

Finally, even where an express contract does not exist, there may actually be differences in the ability of parties to terminate the relationship. Although an agreement normally carries the force of law only if the parties have expressly bound themselves to performance in writing or by other express means, modern contract law will sometimes infer the existence of a contractual obligation if, for instance, one party has relied on the performance of the other. In particular, if one party is induced to undertake investments in support of a transaction, the value of which would not be recoverable if the other party failed to perform, the courts may treat the transaction as though a formal agreement had in fact been accepted

despite insufficient evidence that an explicit bargain had been struck. A distributor who incurred advertising expenditures in reliance on the delivery of brand name merchandise, for example, might be able to recover damages if the manufacturer failed to deliver as promised, even if a formal contract was never stipulated.

Related doctrines under which courts have inferred contractual obligations despite the lack of formal written agreements include promissory estoppel and implied or quasi-contract (see, e.g., Note, *Stanford Law Review*, 1974; and Note, *J. of Law Reform*, 1983). Such remedies have generally been denied in an employment setting, however: "Most courts have been reluctant to find any contractual obligation of just cause discharge in an at-will setting, despite employee reliance on express or implied promises of job security" (Note, *J. of Law Reform*, 1983, p. 455). Instead, the inference of a contractual obligation has been held to a stricter standard requiring "independent consideration": "To avoid arbitrary discharge, employees must provide their employers with consideration, such as monetary contribution, property transfer, or other financial benefit, arising independently of their jobs. Only after providing such 'independent' consideration have employees been able to enforce employer promises of job security. ... In any other commercial setting these employer promises would create binding contractual obligations; in the employment setting, however, they have not been viewed as legally binding." (ibid., pp. 449-50). As a result of this distinction, the threat of termination may be more powerful in employment than commercial settings. A commercial transactor who has relied on the performance of another would have less fear of termination given the potential of recovering the value of his reliance in a court of law. An employee, on the other hand,

who does not have such protection would likely be much more careful to avoid giving an employer reason to reconsider the relationship. Hence, on the margin, the employee again has a greater incentive to accede to the employer's demands than would a commercial contractor.

3. Conclusion

Ex ante, contracting is a flexible institution. Transactors can, at least in principle, design each relationship to suit their particular needs. In all but the simplest exchanges, however, the process of exploring and stipulating details of a transaction can become expensive very quickly. In addition, many basic terms and conditions are likely to be common across transactions. To minimize the costly duplication of identical provisions in individual contracts, the law provides a set of standard doctrines and remedies to deal with recurring contractual events. Thus, both court determined penalties for contract breach and common law application of *force majeure* criteria can be interpreted as substitutes for the redundant stipulation of common provisions by individual transactors. At the same time, courts recognize the diversity of transactions and give parties wide latitude to augment or modify the terms of the agreement by mutual consent. The existence of a standard set of doctrines to govern contractual exchange and the ability to "'contract out of or away from' the governance structures of the state by devising private orderings" (Williamson, 1983, p. 520) combine to provide a degree of both economy and flexibility in constructing contractual relationships.

There is, however, no reason to believe that the distribution of transactions is unimodal, so that the logic that justifies the existence of an original set of common law doctrines may also warrant the establishment of alternative sets of norms and conventions (i.e., institutions) to govern disparate clusters of transactions. Granted, the flexibility afforded transactors in designing contracts admits the possibility of altering rules and remedies to effect any desired structure. Employer-employee relations could be replicated, for instance, through detailed stipulation of the duties and sanctions defined in the law of master and servant in a contract, in which case the contract rather than the case law would become the reference point in the event of a dispute. But accomplishing this would, for all intents and purposes, require reviewing and repeating the entire case law in each contract, obviously forfeiting a substantial economy.[26] Reliance on common law doctrines, in contrast, permits transactors to choose that combination of legal "defaults" or "presets" that most closely approximates the ideal arrangement simply by identifying the class of transactions that the parties intended, to which they may again make incremental adjustments by mutual consent.[27]

The differences in legal defaults, sanctions and procedures governing commercial and employment transactions provide a constructive, as opposed to merely descriptive, connotation to the notion of the firm. The failure to

---

[26]Clark suggests that exact replication may not, in fact, be possible: "[Basic fiduciary duties] cannot be bargained around (unless, perhaps, one is willing to depart from the manager or stockholder roles or to modify their parameters drastically)" (p. 64). See also footnote 24.

[27]The question of how such rules evolve in the first place, and whether specific rules (such the doctrine of employment-at-will) promote efficient organization are obviously worthy of attention but beyond the scope of this paper.

document adequately such differences made it difficult for traditional theorists to identify a basis for the special managerial authority or access to information usually attributed to integration and led a number of prominent authors to deny the existence of a governance role for internal organization. At most, the firm took on constructive meaning only relative to "those assets (e.g. machines, inventories) that it owns" (Grossman and Hart, 1986, p. 692).

At one level, 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. But a change in legal status obviously does not physically transform an asset. What is altered is the relationship of economic actors to those assets, their rights and responsibilities as defined and supported by the legal system. As Harold Demsetz put it, "the problem of defining ownership is precisely that of creating properly scaled legal barriers to entry" (Demsetz, 1982, p. 52), that is, of establishing penalties that promote or discourage specific behavior. In this respect, there is no substantive difference between the power of a manager to direct an employee and an owner's ability to restrain the use or removal of an asset. Just as control over individuals is influenced by the rules and penalties prescribed in the law, so is control over physical capital. In either case, an agent's incentives to comply depend on the sanctions the principal can bring to bear. Thus, although particulars may differ--sanctions for disobedience versus those for theft,

for instance--ownership, like managerial authority, is ultimately a governance issue.[28]

In this paper, I have addressed the question of what distinguishes organization within the firm from external or market exchange by reviewing the legal literature on employment transactions and comparing it with corresponding doctrines in commercial contract law. Although by no means comprehensive, the investigation reveals that the law does in fact recognize substantial differences in the obligations, sanctions and procedures governing the two types of exchange, and that these distinctions are likely to alter the incentives of actors across institutional modes in a meaningful way.[29] Moreover, they do so in a manner that appears to support the conventional view of the firm in economics. On one side, the results lend support to the authority, flexibility and information advantages commonly attributed to internal organization. The effects of the law are such that an employee interested in preserving an employment relationship appears to have stronger incentives to comply with the demands of his employer than would an independent contractor similarly situated. Again, the distinctions and

---

[28]While Grossman and Hart are generally critical of transaction-cost theories of the firm for reasons similar to those outlined in this paper, their use of legal opinion to define ownership in terms "residual rights of control over assets" is compatible with the approach adopted here. They also suggest that their theory may be extended to residual rights of control over actions as a basis for analyzing "the relative advantages of contractor-contractee and employer-employee relationships" (p. 717). Among other things, the present paper details the source and nature of those rights in the duties and sanctions defined by the legal system.

[29]Changes in legal rules over time and cross-jurisdictional differences in the way courts handle employment disputes should alter the incentives of transactors to integrate exchange and the form that integration takes, raising the possibility of empirical tests.

responsibilities uncovered seem to have the intent of making the employee, to as great an extent as possible, an extension of his employer.

At the same time that the laws promote obedience and disclosure by an employee, however, they are also likely to discourage employee initiative and investments in information acquisition and, generally, to require greater monitoring by the employer. An employee's liability for nondisclosure, for instance, not only decreases his incentive to withhold information from his employer but also his incentive to accumulate information in the first place, increasing the need for employee oversight.[30] The employer's liability for the torts of his employees is also likely to encourage greater supervision of subordinates. The strain that this increased attention places on the finite capacities of managers to effectively administer production and exchange represents the principle disadvantage of internalizing successive transactions and ultimately limits the size of the firm.

Clearly, the decision to integrate involves tradeoffs and the choice among institutions ultimately depends of the relative merits of the alternatives. Unfortunately, institutions--unlike quantities of inputs, for example--do not constitute a continuum along which the best choice can be derived at the margin; properties such as rights, ownership and authority are inherently binary in nature. Consequently, optimization over institutional modes

---

[30]Steven Wiggins (1985) associates an employee's reduced incentives to both withhold and acquire information with the receipt of a fixed wage. This seems inappropriate for several reasons. First, many commercial contracts have fixed or definite payment schedules. Hence, the use of fixed wages does not adequately distinguish employment transactions. Second, an employee's ability to influence joint surplus by his information acquisition and transmission activities can also serve strategic purposes in negotiations over the distribution of the resulting gains from trade. The ability to exploit such strategies will differ across organizational modes only if there are institutional differences of the type described in this paper.

requires direct comparison of the net benefits of each alternative under the relevant range of environments. As this paper demonstrates, conscious attention to the institutional structure in which transactors operate is essential to the analysis of those alternatives.

References

Alchian, A.A. and H. Demsetz, "Production, Information Costs, and Economic Organization," American Economic Review, 62, December 1973, pp. 777-795.

Arrow, K.J., "Vertical Integration and Communication," Bell Journal of Economics, Spring 1975, pp. 173-183.

Batt, E.R., The Law of Master and Servant (fifth edition by G.J. Webber). London: Pitman and Sons, 1967.

Carleton, D.W., "Vertical Integration in Competitive Markets Under Uncertainty," Journal of Industrial Economics, 27, March 1979, pp. 189-209.

Chandler, A.D., Strategy and Structure. New York: Doubleday and Co., Anchor Books Edition, 1966.

Cheung, S.N.S., "The Contractual Nature of the Firm," Journal of Law and Economics, 26, April 1983, pp. 1-21.

Clark, R.C., "Agency Costs versus Fiduciary Duties," in J. Pratt and R. Zeckhauser, eds., Principals and Agents: The Structure of Business, 1985, pp. 55-79, 217-219.

Coase, R., "The Nature of the Firm," reprinted in G. Stigler and K. Boulding (eds.), AEA Readings in Price Theory, Chicago 1952, pp. 331-51.

Corbin, A., Contracts, 4, 1960.

Crocker, K.J., "Vertical Integration and the Strategic Use of Private Information," Bell Journal of Economics, 14, Spring 1983, pp. 236-248.

Demsetz, H., "Barriers to Entry," American Economic Review, 72, March 1982, pp. 47-57.

Dykstra, G.O. and Dykstra, L.G., Business Law: Text and Cases (3rd (Uniform Commercial Code) edition). New York: Pitman, 1969.

Epstein, R.A., "In Defence of the Contract at Will," The University of Chicago Law Review, 51. 1984, pp. 947-82.

Evans, David, and Sanford Grossman, "Integration," in D. Evans (ed.), Breaking Up Bell. New York: North-Holland Publishing Co., 1983, pp. 95-126.

Geotz, C.J., and Scott, R.E., "The Mitigation Principle: Toward a General Theory of Contractual Obligation," Virginia Law Review, 69, September 1983, pp. 967-1024.

Grossman Sanford J. and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical Integration," Journal of Political Economy, 94, August 1986, pp. 691-719.

Klein, B., Crawford, R.G., and Alchian, A.A., "Vertical Integration, Appropriable Rents and the Competitive Contracting Process," Journal of Law and Economics, 21, October 1978, pp. 297-326.

_____, "Transaction-Cost Determinants of 'Unfair' Contractual Arrangements," American Economic Review, 70, May 1980, pp. 356-62.

Kronman, A.T. "Mistake, Disclosure, Information, and the Law of Contracts," Journal of Legal Studies, 7, January 1978, pp. 1-34.

Macaulay, S., "Non-Contractual Relations in Business," American Sociological Review, 28, February 1963, pp. 55-70.

Jensen, M.C., and Meckling, W.H., "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," Journal of Financial Economics, 3, October 1976, pp. 305-360.

Rubin, P.H., "The Theory of the Firm and the Structure of the Franchise Contract," Journal of Law and Economics, 21, April 1978, pp. 223-233.

Simon, Herbert A., "A Formal Theory of the Employment Relationship," Econometrica, 19, July 1951, pp. 293-305.

Steiber, J., and Murray, M., "Protection Against Unjust Discharge: The Need for a Federal Statute," Journal of Law Reform, 16, Winter 1983, pp. 319-341.

Vernon, J.M. and D.A. Graham, "Profitability of Monopolization by Vertical Integration," Journal of Political Economy, 79, July/August 1971, pp. 924-25.

Warren-Boulton, F.R., "Vertical Control with Variable Proportions," Journal of Political Economy, 82, July/August 1974, pp. 783-802.

Wiggins, S. N., "The Comparative Advantage of Institutions: Spot Markets, Long Term Contracts, and Firms," (mimeo, October 1985).

Williamson, O.E., Markets and Hierarchies. New York: Free Press, 1975.

_____, "Credible Commitments: Using Hostages to Support Exchange," American Economic Review, 83, September 1983, pp. 519-540.

_____, The Economic Institutions of Capitalism. New York: Free Press, 1985.

Note, "Challenging the Employment-At-Will Doctrine Through Modern Contract Theory," Journal of Law Reform, 16, Winter 1983, pp. 449-64.

Note, "Implied Contract Rights to Job Security," Stanford Law Review, 26, 1974, pp. 335-369.

American Jurisprudence

American Law Reports

Corpus Juris Secundum, Master and Servant

Restatement (Second) of Agency

Restatement (Second) of Contracts

Restatement (Second) of Torts

Uniform Commercial Code

Byer vs. International Paper Co., 314 F. 2nd 831 (1963)

Coker vs. Wesco Materials Corp., 308 S.W. 2nd 884 (1963)

Frederick Chusid & Co. vs. Marshall Leeman & Co., 326 F. Supp. 1043 (1971)

Michigan Crown Fender Co. v. Welch, 13 A.L.R. 896 (1920)

Pitts vs. Shell Oil Co., 463 F. 2nd 33 (1972)