



An Economist's Perspective on the Theory of the Firm

Author(s): Oliver Hart

Source: *Columbia Law Review*, Vol. 89, No. 7, Contractual Freedom in Corporate Law (Nov., 1989), pp. 1757-1774

Published by: Columbia Law Review Association, Inc.

Stable URL: <http://www.jstor.org/stable/1122818>

Accessed: 24/08/2008 12:44

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=cra>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

AN ECONOMIST'S PERSPECTIVE ON THE THEORY OF THE FIRM

*Oliver Hart**

An outsider to the field of economics would probably take it for granted that economists have a highly developed theory of the firm. After all, firms are the engines of growth of modern capitalistic economies, and so economists must surely have fairly sophisticated views of how they behave. In fact, little could be further from the truth. Most formal models of the firm are extremely rudimentary, capable only of portraying hypothetical firms that bear little relation to the complex organizations we see in the world. Furthermore, theories that attempt to incorporate real world features of corporations, partnerships and the like often lack precision and rigor, and have therefore failed, by and large, to be accepted by the theoretical mainstream.

This Article attempts to give lawyers a sense of how economists think about firms. It does not pretend to offer a systematic survey of the area; rather, it highlights several ideas of particular importance, and then explores an alternative theoretical perspective from which to view the firm.¹ Part I introduces various established economic theories of the firm. Part II turns to a newer theory of the firm, based not upon human capital structures, but rather upon property rights. Part III synthesizes this property rights-based theory of the firm with more established theories.

I. ESTABLISHED THEORIES

A. *Neoclassical Theory*

Any discussion of theories of the firm must start with the neoclassical approach, the staple diet of modern economists. Developed over the last one hundred years or so, this approach can be found in any

* Professor of Economics, Massachusetts Institute of Technology. B.A. 1969, Cambridge University; M.A. 1972, Warwick University; Ph.D. 1974, Princeton University. Helpful comments from Jeffrey Gordon, Bengt Holmstrom and Jean Tirole are gratefully acknowledged. This Article is based in part on the author's Fisher-Schultze lecture delivered to the Econometric Society in Bologna, Italy in August 1988. Some of the work was done while the author was visiting the Harvard Business School as a Marvin Bower Fellow. He would like to thank that institution for its hospitality and financial support. The author would also like to acknowledge financial assistance from the Guggenheim and Olin Foundations, the Center for Energy and Policy Research at MIT and the National Science Foundation.

1. Several recent surveys provide other perspectives on this material. See, e.g., Holmstrom & Tirole, *The Theory of the Firm*, in *1 Handbook of Industrial Organization* (R. Schmalensee & R. Willig eds., forthcoming 1989); Milgrom & Roberts, *Economic Theories of the Firm: Past, Present and Future*, 21 *Can. J. Econ.* 444 (1988); Williamson, *The Logic of Economic Organization*, 4 *J.L. Econ. & Organization* 65 (1988).

modern-day textbook; in fact, in most textbooks, it is the *only* theory of the firm presented.²

Neoclassical theory views the firm as a set of feasible production plans.³ A manager presides over this production set, buying and selling inputs and outputs in a spot market and choosing the plan that maximizes owners' welfare. Welfare is usually represented by profit, or, if profit is uncertain so that profit-maximization is not well defined, by expected net present value of future profit (possibly discounted for risk) or by market value.

To many lawyers and economists, this is a caricature of the modern firm; it is rigorous but rudimentary. At least three reasons help explain its prolonged survival. First, the theory lends itself to an elegant and general mathematical formalization. Second, it is very useful for analyzing how a firm's production choices respond to exogenous change in the environment, such as an increase in wages or a sales tax.⁴ Finally, the theory is also very useful for analyzing the consequences of strategic interaction between firms under conditions of imperfect competition;⁵ for example, it can help us understand the relationship between the degree of concentration in an industry and that industry's output and price level.

Granted these strengths, neoclassical theory has some very clear weaknesses. It does not explain how production is organized within a firm, how conflicts of interest between the firm's various constituencies—its owners, managers, workers, and consumers—are resolved, or, more generally, how the goal of profit-maximization is achieved. More subtly, neoclassical theory begs the question of what defines a given firm or what determines its boundaries. Since the theory does not address the issue of each firm's size or extent, it does not explain the consequences of two firms choosing to merge, or of one firm splitting itself into two or more smaller firms. Neoclassical theory describes in rudimentary terms how firms function, but contributes little to any meaningful picture of their structure.

B. *Principal-Agent Theory*

Principal-agent theory, an important development of the last fifteen years, addresses some of the weaknesses of the neoclassical approach.⁶ Principal-agent theory recognizes conflicts of interest

2. See, e.g., J. Henderson & R. Quandt, *Microeconomic Theory: A Mathematical Approach* 64–134 (1980); H. Varian, *Microeconomic Analysis* 6–78 (1984).

3. For example, one feasible plan might be to use 10 person-hours and one acre of land to produce one hundred pounds of wheat, while another feasible plan might be to use 12 person-hours and one and one-half acres to produce fifty pounds of corn.

4. See H. Varian, *supra* note 2, at 47; Bishop, *The Effects of Specific and Ad Valorem Taxes*, 82 *Q.J. Econ.* 198 (1968).

5. See J. Tirole, *The Theory of Industrial Organization* 205–301 (1988).

6. See, e.g., Holmstrom, *Moral Hazard and Observability*, 10 *Bell J. Econ.* 74

between different economic actors, formalizing these conflicts through the inclusion of observability problems and asymmetries of information. The theory still views the firm as a production set, but now a professional manager makes production choices, such as investment or effort allocations, that the firm's owners do not observe. Because the manager deals with the day-to-day operations of the firm, she also is presumed to have information about the firm's profitability that the owners lack. In addition, the manager has other goals in mind beyond the owners' welfare, such as on-the-job perks, an easy life, empire building, and so on. Under these conditions, principal-agent theory argues that it will be impossible for the owners to implement their own profit-maximizing plan directly, through a contract with the manager—in general, the owners will not even be able to tell *ex post* whether the manager has chosen the right plan. Instead, the owners will try to align the manager's objectives with their own by putting the manager on an incentive scheme. Even under an optimal incentive scheme, however, the manager will put some weight on her own objectives at the expense of those of the owners, and conflicting interests remain. Hence, we have the beginnings of a managerial theory of the firm.⁷

Principal-agent theory enriches neoclassical theory significantly, but still fails to answer the vital questions of what defines a firm and where the boundaries of its structure are located. To see why, consider the example of Fisher Body, which for many years has supplied car bodies to General Motors.⁸ Principal-agent theory can explain why it might make sense for GM and Fisher to write a profit-sharing agreement, whereby part of Fisher Body's reward is based on GM's profit from car sales: this encourages Fisher to supply high-quality inputs. The theory does not tell us, however, whether it matters if this profit-sharing agreement is accomplished through the merger of Fisher and GM into a single firm, with GM having authority over Fisher management; or whether GM and Fisher should remain as separate firms; or whether GM and Fisher should merge, with Fisher management having authority over GM management.⁹ In other words, principal-agent theory tells

(1979); Shavell, Risk Sharing and Incentives in the Principal and Agent Relationship, 10 *Bell J. Econ.* 55 (1979). For a recent survey, see Hart & Holmstrom, *The Theory of Contracts*, in *Advances in Economic Theory: Fifth World Congress* 71, 75–106 (T. Bewley ed. 1987).

7. It is also possible to extend the principal-agent view of the firm to analyze conflicts of interests between managers and workers, and those between managers and consumers. See Calvo & Wellisz, *Supervision, Loss of Control and the Optimum Size of the Firm*, 86 *J. Pol. Econ.* 943 (1978).

8. I will discuss the GM-Fisher Body relationship at several points in the text. In doing so, I draw on material from Klein, Crawford & Alchian, *Vertical Integration, Appropriable Rents, and the Competitive Contracting Process*, 21 *J.L. & Econ.* 297, 308–10 (1978); Klein, *Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited*, 4 *J.L. Econ. & Organization* 199 (1988).

9. As a matter of history, GM and Fisher started off as separate firms linked by a

us about optimal incentive schemes, but not (at least directly) about organizational form. Hence, in the absence of a parallel between the two, which turns out to be difficult to draw, principal-agent theory provides no predictions about the nature and extent of the firm.¹⁰

C. *Transaction Cost Economics*

While the neoclassical paradigm, modified by principal-agent theory, progressed along the above lines, a very different approach to the theory of the firm developed under the heading of transaction cost economics. Introduced in Coase's famous 1937 article,¹¹ transaction cost economics traces the existence of firms to the thinking, planning and contracting costs that accompany any transaction, costs usually ignored by the neoclassical paradigm. The idea is that in some situations these costs will be lower if a transaction is carried out within a firm rather than in the market. According to Coase, the main cost of transacting in the market is the cost of learning about and haggling over the terms of trade; this cost can be particularly large if the transaction is a long-term one in which learning and haggling must be performed repeatedly.¹² Transaction costs can be reduced by giving one party authority over the terms of trade, at least within limits. But, according to Coase, this authority is precisely what defines a firm: within a firm, transactions occur as a result of instructions or orders issued by a boss, and the price mechanism is suppressed.¹³

Such an arrangement, however, brings costs of its own. Concentrating authority in one person's hands is likely to increase the cost of errors and lead to greater administrative rigidity. In Coase's view, the boundaries of the firm occur at the point where the marginal cost savings from transacting within the firm equal these additional error and

long-term contract, but after a dispute GM bought Fisher in 1926. Klein, Crawford & Alchian, *supra* note 8, at 310.

10. Drawing a parallel might be possible if, say, profit- or cost-sharing arrangements were only found within a single firm. This is not the case, however. For example, consider cost-plus contracts between the United States Government and private defense contractors. See generally F. Scherer, *The Weapons Acquisition Process: Economic Incentives* 131-309 (1964) for a discussion of defense contracts.

11. Coase, *The Nature of the Firm*, 4 *Economica* 386 (1937).

12. One can distinguish between learning and haggling costs incurred at the beginning of the relationship when the parties reach an initial agreement and those incurred as the relationship proceeds and the parties revise their agreement. For present purposes, the latter costs are more important.

13. A related idea can be found in Simon, *A Formal Theory of the Employment Relationship*, 19 *Econometrica* 293 (1951) (arguing that it is efficient for employee to accept employer's authority if employee is approximately indifferent about tasks he performs, but employer has a strict preference). It is also worth noting that the superior adaptive properties of the employment relation were emphasized by Chester Barnard at around the same time that Coase was writing. See C. Barnard, *The Functions of the Executive* 139-60 (1938) (discussing incentives necessary to induce individuals to contribute to organizations).

rigidity costs.¹⁴

Coase's ideas, although recognized as highly original, took a long time to catch on.¹⁵ There are probably two reasons for this. First, they remain to this day very hard to formalize. Second, there is a conceptual weakness, pointed out by Alchian and Demsetz,¹⁶ in the theory's dichotomy between the role of authority within the firm and the role of consensual trade within the market. Consider, for example, Coase's notion that an employer has authority over an employee—an employer can tell an employee what to do.¹⁷ Alchian and Demsetz questioned this, asking what ensures that the employee obeys the employer's instructions. To put it another way, what happens to the employee if he disobeys these instructions? Will he be sued for breach of contract? Unlikely. Probably the worst that can happen is the employee will be fired. But firing is typically the sanction that one independent contractor will impose on another whose performance he does not like. To paraphrase Alchian and Demsetz's criticism, it is not clear that an employer can tell an employee what to do, any more than a consumer can tell her grocer what to do (what vegetables to sell at what prices); in either case, a refusal will likely lead to a termination of the relationship, a firing. In the case of the grocer, this means that the consumer shops at another grocer.¹⁸ Thus, according to Alchian and Demsetz's argument, Coase's view that firms are characterized by authority relations does not really stand up.¹⁹

Finding Coase's characterization of the firm wanting, Alchian and Demsetz developed their own theory, based on joint production and monitoring. Transactions involving joint or team production require careful monitoring so that each actor's contribution can be assessed. According to Alchian and Demsetz, the best way to provide the monitor with appropriate incentives is to give him the following bundle of rights, which effectively define ownership of the capitalist firm: 1) to be a residual claimant; 2) to observe input behavior; 3) to be the central

14. Coase, *supra* note 11, at 395.

15. In Coase's words, they were "much cited and little used" (until the 1970s). Coase, *The Nature of the Firm: Influence*, 4 *J.L. Econ. & Organization* 33, 33 (1988).

16. Alchian & Demsetz, *Production, Information Costs, and Economic Organization*, 62 *Am. Econ. Rev.* 777 (1972).

17. Coase, *supra* note 11, at 404.

18. Alchian & Demsetz, *supra* note 16, at 777–78, 783–84. But cf. Masten, *A Legal Basis for the Firm*, 4 *J.L. Econ. & Organization* 181, 186–87 (1988) (law makes a distinction between an employer-employee relationship and one between independent contractors, in that an employee owes her employer duties of loyalty and obedience that do not exist between independent contractors).

19. It bears noting that the second part of Coase's thesis, maintaining that firms suppress the price mechanism, is also flawed. The use of prices to allocate resources within a multidivisional firm—the phenomenon of transfer pricing, probably more common now than it was when Coase wrote—seems a fairly immediate counterexample. For a recent discussion of the use of transfer pricing, see Eccles & White, *Price and Authority in Inter-Profit Center Transactions*, 94 *Am. J. Soc.* S17 (Supp. 1988).

party common to all contracts with inputs; 4) to alter membership of the team; and, 5) to sell rights 1–4.²⁰ We will return to some of these ideas below, but at this stage it suffices to note that the theory suffers from the same criticism levelled at Coase: it is unclear why the problems of joint production and monitoring must be solved through the firm and cannot be solved through the market. In fact, one does not need to look far to see examples of market solutions to these problems, such as auditing between independent contractors.

At the same time that doubts were being expressed about the specifics of Coase's theory, Coase's major idea—that firms arise to economize on transaction costs—was increasingly accepted. The exact nature of these transaction costs, however, remained unclear. What lay beyond the learning and haggling costs that, according to Coase, are a major component of market transactions? Professor Oliver Williamson has offered the deepest and most far-reaching analysis of these costs.²¹ Williamson recognized that transaction costs may assume particular importance in situations where economic actors make relationship-specific investments—investments to some extent specific to a particular set of individuals or assets.²² Examples of such investments include locating an electricity generating plant adjacent to a coal mine that is going to supply it; a firm's expanding capacity to satisfy a particular customer's demands; training a worker to operate a particular set of machines or to work with a particular group of individuals; or a worker's relocating to a town where he has a new job.²³

In situations like these, there may be plenty of competition before the investments are made—there may be many coal mines next to which an electricity generating plant could locate or many towns to which a worker could move. But once the parties sink their investments, they are to some extent locked into each other. As a result, external markets will not provide a guide to the parties' opportunity costs once the relationship is underway. This lack of information takes on great significance, since, in view of the size and degree of the specific investment, one would expect relationships like these to be long lasting.²⁴

In an ideal world, the lack of ex post market signals would pose no problem, since the parties could always write a long-term contract in advance of the investment, spelling out each agent's obligations and

20. See Alchian & Demsetz, *supra* note 16, at 783.

21. See generally O. Williamson, *The Economic Institutions of Capitalism* (1985) [hereinafter *Economic Institutions*]; O. Williamson, *Markets and Hierarchies: Analysis and Antitrust Implications* (1975). For another significant analysis of these costs, see generally Klein, Crawford & Alchian, *supra* note 8.

22. See *Economic Institutions*, *supra* note 21, at 30.

23. *Id.* at 95–96.

24. *Id.* at 61. For empirical evidence on the importance of relationship-specific investments and lock-in effects, see Joskow, *Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence*, 4 *J.L. Econ. & Organization* 95 (1988).

the terms of the trade in every conceivable situation. In practice, however, thinking, negotiation and enforcement costs will usually make such a contract prohibitively expensive. As a result, parties must negotiate many of the terms of the relationship as they go along. Williamson argues that this leads to two sorts of costs. First, there will be costs associated with the ex post negotiation itself—the parties may engage in collectively wasteful activities to try to increase their own share of the ex post surplus; also, asymmetries of information may make some gains from trade difficult to realize.²⁵ Second, and perhaps more fundamental, since a party's bargaining power and resulting share of the ex post surplus may bear little relation to his ex ante investment, parties will have the wrong investment incentives at the ex ante stage.²⁶ In particular, a far-sighted agent will choose her investment inefficiently from the point of view of her contracting partners, given that she realizes that these partners could expropriate part of her investment at the ex post stage.²⁷

In Williamson's view, bringing a transaction from the market into the firm—the phenomenon of integration—mitigates this opportunistic behavior and improves investment incentives. Agent *A* is less likely to hold up agent *B* if *A* is an employee of *B* than if *A* is an independent contractor. However, Williamson does not spell out in precise terms the mechanism by which this reduction in opportunism occurs. Moreover, certain costs presumably accompany integration. Otherwise, all transactions would be carried out in firms, and the market would not be used at all. Williamson, however, leaves the precise nature of these costs unclear.²⁸

D. *The Firm as a Nexus of Contracts*

All the theories discussed so far suffer from the same weakness:

25. Economic Institutions, *supra* note 21, at 21.

26. *Id.* at 88–89.

27. *Id.* at 30–32.

28. Williamson argues that a major benefit of integration comes from the fact that the party with authority can resolve disputes by fiat (as opposed to litigation), while a major cost comes from the fact that the party with authority cannot commit himself to intervene selectively in the affairs of other parties. See *id.* at 76, 133–35. Williamson, however, is not very clear about what mechanisms are at work here. For example, a boss may try to resolve a dispute, but what guarantee is there that the parties will follow his edicts? To paraphrase Alchian and Demsetz, what disciplinary power does a boss have that an independent contractor does not? A similar issue arises with regard to selective intervention. In what activities will the boss intervene, and how will this intervention be enforced? What power to intervene does a boss have that an independent contractor does not have? See *supra* notes 16–18 and accompanying text.

The greater powers of a boss relative to an independent contractor can be understood if one takes a property rights-based view of the firm—in particular, if one recognizes that a firm consists of nonhuman assets as well as human assets and that a boss typically has control over these nonhuman assets. See *infra* notes 33–46 and accompanying text.

while they throw light on the nature of contractual failure, none explains in a convincing or rigorous manner how bringing a transaction into the firm mitigates this failure.

One reaction to this weakness is to argue that it is not really a weakness at all. According to this point of view, the firm is simply a nexus of contracts,²⁹ and there is therefore little point in trying to distinguish between transactions within a firm and those between firms. Rather, both categories of transactions are part of a continuum of types of contractual relations, with different firms or organizations representing different points on this continuum.³⁰ In other words, each type of business organization represents nothing more than a particular "standard form" contract. One such "standard form" contract is a public corporation, characterized by limited liability, indefinite life, and freely transferable shares and votes. In principle it would be possible to create a contract with these characteristics each time it is needed, but, given that these characteristics are likely to be useful in many different contexts, it is much more convenient to be able to appeal to a "standard form." Closely held corporations or partnerships are other examples of useful "standard forms."

Viewing the firm as a nexus of contracts is helpful in drawing attention to the fact that contractual relations with employees, suppliers, customers, creditors and others are an essential aspect of the firm. Also, it

serves to make it clear that the personalization of the firm implied by asking questions such as "what should be the objective function of the firm" . . . is seriously misleading. *The firm is not an individual* The "behavior" of the firm is like the behavior of a market, i.e., the outcome of a complex equilibrium process.³¹

At the same time, the nexus of contracts approach does less to resolve the questions of what a firm is than to shift the terms of the debate. In particular, it leaves open the question of why particular "standard forms" are chosen. Perhaps more fundamentally, it begs the question of what limits the set of activities covered by a "standard form." For example, corporations are characterized by limited liability, free transferability of shares, and indefinite life. But what limits the size of a corporation—what are the economic consequences of two corpora-

29. The nexus of contract theory is often associated with Jensen and Meckling. See Jensen & Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 *J. Fin. Econ.* 305, 310 (1976).

30. Note that lawyers' and economists' ideas of what constitutes a contract may differ. Economists tend to view contracts as relationships characterized by reciprocal expectations and behavior; lawyers consider the enforceable legal duties implicit in such relationships and look for formalization through the standard indicia of contract formation, such as offer and acceptance. See Gordon, *The Mandatory Structure of Corporate Law*, 89 *Colum. L. Rev.* 1549, 1549-50 (1989).

31. Jensen & Meckling, *supra* note 29, at 311.

tions merging or of one corporation splitting itself into two? Given that mergers and breakups occur all the time, and at considerable transaction cost, it seems unlikely that such changes are cosmetic. Presumably they have some real effects on incentives and opportunistic behavior, but these effects remain unexplained.

II. A PROPERTY RIGHTS APPROACH TO THE FIRM

One way to resolve the question of how integration changes incentives is spelled out in recent literature that views the firm as a set of property rights.³² This approach is very much in the spirit of the transaction cost literature of Coase and Williamson, but differs by focusing attention on the role of physical, that is, nonhuman, assets in a contractual relationship.

Consider an economic relationship of the type analyzed by Williamson, where relationship-specific investments are important and transaction costs make it impossible to write a comprehensive long-term contract to govern the terms of the relationship. Consider also the nonhuman assets that, in the postinvestment stage, make up this relationship. Given that the initial contract has gaps, missing provisions, or ambiguities, situations will typically occur in which some aspects of the use of these assets are not specified. For example, a contract between GM and Fisher might leave open certain aspects of maintenance policy for Fisher machines, or might not specify the speed of the production line or the number of shifts per day.

Take the position that the right to choose these missing aspects of usage resides with the *owner* of the asset. That is, ownership of an asset goes together with the possession of residual rights of control over that asset; the owner has the right to use the asset in any way not inconsistent with a prior contract, custom, or any law. Thus, the owner of Fisher assets would have the right to choose maintenance policy and production line speed to the extent that the initial contract was silent about these.³³

32. See generally Grossman & Hart, *The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration*, 94 *J. Pol. Econ.* 691 (1986); Holmstrom & Tirole, *supra* note 1; O. Hart & J. Moore, *Property Rights and the Nature of the Firm* (Massachusetts Institute of Technology, Dep't of Economics Working Paper No. 495, 1988). This literature owes much to the earlier property rights literature on the efficiency of private property in an externality-free world. See, e.g., Demsetz, *Toward a Theory of Property Rights*, *Am. Econ. Rev.*, May 1967, at 347.

33. This view of ownership seems consistent with the standard one adopted by lawyers:

But what are the rights of ownership? They are substantially the same as those incident to possession. Within the limits prescribed by policy, the owner is allowed to exercise his natural powers over the subject-matter uninterfered with, and is more or less protected in excluding other people from such interference.

The owner is allowed to exclude all, and is accountable to no one.

O. Holmes, *The Common Law* 193 (1963 ed.).

Finally, identify a firm with all the nonhuman assets that belong to it, assets that the firm's owners possess by virtue of being owners of the firm. Included in this category are machines, inventories, buildings or locations, cash, client lists, patents, copyrights, and the rights and obligations embodied in outstanding contracts to the extent that these are also transferred with ownership. Human assets, however, are not included. Since human assets cannot be bought or sold, management and workers presumably own their human capital both before and after a merger.

We now have the basic ingredients of a theory of the firm. In a world of transaction costs and incomplete contracts, *ex post* residual rights of control will be important because, through their influence on asset usage, they will affect *ex post* bargaining power and the division of *ex post* surplus in a relationship. This division in turn will affect the incentives of actors to invest in that relationship. Hence, when contracts are incomplete, the boundaries of firms matter in that these boundaries determine who owns and controls which assets.³⁴ In particular, a merger of two firms does not yield unambiguous benefits: to the extent that the (owner-)manager of the acquired firm loses control rights, his incentive to invest in the relationship will decrease. In addition, the shift in control may lower the investment incentives of workers in the acquired firm. In some cases these reductions in investment will be sufficiently great that nonintegration is preferable to integration.³⁵

Note that, according to this theory, when assessing the effects of integration, one must know not only the characteristics of the merging firms, but also who will own the merged company. If firms *A* and *B* integrate and *A* becomes the owner of the merged company, then *A* will presumably control the residual rights in the new firm. *A* can then use those rights to hold up the managers and workers of firm *B*. Should the

34. This consolidation of ownership and control points to an important lacuna in the property rights approach. The approach makes no distinction between ownership and control, assuming that both rest with the same entity. In most of the formal models that have been developed, such an arrangement turns out to be optimal since agents are assumed to be risk-neutral and to have sufficient wealth to buy any asset. If managers were risk-averse and had limited wealth, however, this conclusion would no longer be valid. Moreover, from a descriptive point of view, the assumption that owners manage is seriously inadequate; while it may apply to small firms such as partnerships or closed corporations, it certainly does not apply to large, publicly held corporations. For how the ownership/control dichotomy might affect the property rights approach, see *infra* notes 58–59 and accompanying text.

35. It is important to emphasize that the property rights approach distinguishes between ownership in the sense of possession of residual control rights over assets and ownership in the sense of entitlement to a firm's (verifiable) profit stream. In practice, these rights will often go together, but they do not have to. The property rights approach takes the point of view that the possession of control rights is crucial for the integration decision. That is, if firm *A* wants to acquire part of firm *B*'s (verifiable) profit stream, it can always do this by contract. It is only if firm *A* wants to acquire control over firm *B*'s assets that it needs to integrate.

situation be reversed, a different set of control relations would result in *B* exercising control over *A*, and *A*'s workers and managers would be liable to holdups by *B*.

It will be helpful to illustrate these ideas in the context of the Fisher Body-General Motors relationship.³⁶ Suppose these companies have an initial contract that requires Fisher to supply GM with a certain number of car bodies each week. Imagine that demand for GM cars now rises and GM wants Fisher to increase the quantity it supplies. Suppose also that the initial contract is silent about this possibility, perhaps because of a difficulty in predicting Fisher's costs of increasing supply. If Fisher is a separate company, GM presumably must secure Fisher's permission to increase supply. That is, the status quo point in any contract renegotiation is where Fisher does *not* provide the extra bodies. In particular, GM does not have the right to go into Fisher's factory and set the production line to supply the extra bodies; Fisher, as owner, has this residual right of control. The situation is very different if Fisher is a subdivision or subsidiary of GM, so that GM owns Fisher's factory. In this case, if Fisher management refuses to supply the extra bodies, GM always has the option to fire management and hire someone else to supervise the factory and supply extra bodies (they could even run Fisher themselves on a temporary basis). The status quo point in the contract renegotiation is therefore quite different.

To put it very simply, if Fisher is a separate firm, Fisher management can threaten to make both Fisher assets and their own labor unavailable for the uncontracted-for supply increase. In contrast, if Fisher belongs to GM, Fisher management can only threaten to make their own labor unavailable. The latter threat will generally be much weaker than the former.³⁷

Although the status quo point in the contract renegotiation may depend on whether GM and Fisher are one firm rather than two, it does not follow that the outcomes after renegotiation will differ. In fact, if the benefits to GM of the extra car bodies exceed the costs to Fisher of supplying them, we might expect the parties to agree that the bodies should be supplied, regardless of the status quo point. However, the divisions of surplus in the two cases will be very different. If GM and Fisher are separate, GM may have to pay Fisher a large sum to persuade it to supply the extra bodies. In contrast, if GM owns Fisher's plant, it may be able to enforce the extra supply at much lower cost since, as we have seen in this case, Fisher management has much reduced bargaining and threat power.

Anticipating the way surplus is divided, GM will typically be much more prepared to invest in machinery that is specifically geared to

36. See *supra* note 8.

37. If current Fisher management is indispensable for the operation of Fisher assets, there is, of course, no difference between the two threats. It is rare, however, that current management is completely irreplaceable.

Fisher bodies if it owns Fisher than if Fisher is independent, since the threat of expropriation is reduced.³⁸ The incentives for Fisher, however, may be quite the opposite. Fisher management will generally be much more willing to come up with cost-saving or quality-enhancing innovations if Fisher is an independent firm than if it is part of GM, because Fisher management is more likely to see a return on its activities. If Fisher is independent, it can extract some of GM's surplus by threatening to deny GM access to the assets embodying these innovations. In contrast, if GM owns the assets, Fisher management faces total expropriation of the value of the innovation to the extent that the innovation is asset-specific rather than management-specific, and GM can threaten to hire a new management team to incorporate the innovation.³⁹

So far, we have discussed the effects of control changes on the incentives of top management. But workers' incentives will also be affected. Consider, for example, the incentive of someone who works with Fisher assets to improve the quality of Fisher's output by better learning some aspect of the production process. Suppose further that GM has a specific interest in this improvement in car body quality, and that none of Fisher's other customers cares about it. There are many ways in which the worker might be rewarded for this, but one important reward is likely to come from the fact that the worker's value to the Fisher-GM venture will rise in the future and, due to his additional skills, the worker will be able to extract some of these benefits through a higher wage or promotion. Note, however, that the worker's ability to do this is greater if GM controls the assets than if Fisher does. In the former case, the worker will bargain directly with GM, the party that

38. It should be emphasized that there is no inconsistency in assuming that an initial contract is incomplete and at the same time that the parties anticipate how the ex post surplus will be divided up as a result of this incompleteness. For example, suppose there are many individually unlikely states with similar characteristics to an uncontracted-for increase in demand. It may be prohibitively expensive for the parties to contract for each of these states, and yet they may be well aware of the average degree to which their investments will be expropriated as a result of not contracting for these states.

39. Under some conditions expropriation problems can be avoided regardless of organizational form. One possibility is for the parties to write an ex ante profit-sharing agreement. However, a profit-sharing agreement may be insufficient to encourage ex ante investments to the extent that some returns from an asset's use are unverifiable. Examples of unverifiable returns are effort costs, nonmonetary rewards such as perks, and monetary returns that can be diverted so that they do not show up in the firm's accounts.

Another way the parties might overcome expropriation problems is to share investment expenditures. For example, if Fisher and GM are independent, Fisher could compensate GM for its later hold-up power by contributing towards GM's initial Fisher-specific investment. Note, however, that this strategy will work only to the extent that either GM contractually agrees to make the investment or Fisher can make part of the investment on GM's behalf. Otherwise, GM could use an up-front payment from Fisher to make a *non*-relationship-specific investment.

benefits from the worker's increased skill.⁴⁰ In the latter case, the worker will bargain with Fisher, who only receives a fraction of these benefits, since it must in turn bargain with GM to parlay these benefits into dollars. In consequence, the worker will typically capture a lower share of the surplus, and his incentive to make the improvement in the first place will fall.

In other words, given that the worker may be held up no matter who owns the Fisher assets—assuming that he, himself, does not—his incentives are greater if the number of possible hold-ups is smaller rather than larger. With Fisher management in control of the assets, there are two potential hold-ups: Fisher can deny the worker access to the assets, and GM can decline to pay more for the improved product.⁴¹ As a result, we might expect the worker to get, say, a third of his increased marginal product (supposing equal division with Fisher and GM). With GM management in control of the Fisher assets, there is only one potential hold-up, since the power to deny the worker his increased marginal product is concentrated in one agent's hands. As a result, the worker in this case might be able to capture half of his increased marginal product (supposing equal division with GM).⁴²

The above reasoning applies to the case in which the improvement is specific to GM. Exactly the opposite conclusion would be reached, however, if the improvement were specific to Fisher, such as the worker learning how to reduce Fisher management's costs of making car bodies, regardless of Fisher's final customer (a cost reduction, furthermore, which could not be enjoyed by any substitute for Fisher management). In that event, the number of hold-ups is reduced by giving control of Fisher assets to Fisher management rather than GM. The reason is that with Fisher management in control, the worker bargains with the party who benefits directly from his increased productivity, whereas with GM management in control, he must bargain with an indirect recipient; GM must in turn bargain with Fisher management to benefit from the reduction in costs.

40. This is not quite correct since the worker will actually bargain with GM management rather than with GM shareholders, who are arguably the ultimate beneficiaries. However, it is approximately correct to the extent that, perhaps because GM management is on an incentive scheme, GM management benefits from an increase in GM's profit or market value. For the remainder of the discussion, we will, at a cost both in precision and realism, ignore the distinction between management and shareholders, and also treat management as a monolithic group. But see *supra* note 34; *infra* note 46 (explaining how this analysis can be generalized to include more complicated forms of group ownership); *infra* notes 58–59 and accompanying text.

41. We assume that no payment was specified for the improved product in the initial contract.

42. For a formal treatment of the division of surplus, see O. Hart & J. Moore, *supra* note 32, at 11. The numbers one-half and one-third should not be taken too seriously. The important point is that, in the context described, the worker is likely to get a larger share of his increased marginal product when GM controls the assets than when Fisher does.

Up to this point we have assumed that GM management will control GM assets. This, however, need not be the case; in some situations it might make more sense for Fisher management to control these assets—for Fisher to buy up GM. One thing we can be sure of is that if GM and Fisher assets are sufficiently complementary, and initial contracts sufficiently incomplete, then the two sets of assets should be under common control. With extreme complementarity, no agent—whether manager or worker—can benefit from any increase in his marginal productivity unless he has access to both sets of assets (by the definition of extreme complementarity, each asset, by itself, is useless). Giving control of these assets to two different management teams is therefore bound to be detrimental to actors' incentives, since it increases the number of parties with hold-up power.⁴³ This result confirms the notion that when lock-in effects⁴⁴ are extreme, integration will dominate nonintegration.⁴⁵

These ideas can be used to construct a theory of the firm's boundaries. First, as we have seen, highly complementary assets should be owned in common, which may provide a minimum size for the firm. Second, as the firm grows beyond a certain point, the manager at the center will become less and less important with regard to operations at the periphery in the sense that increases in marginal product at the periphery are unlikely to be specific either to this manager or to the assets at the center. At this stage, a new firm should be created since giving the central manager control of the periphery will increase hold-up problems without any compensating gains. It should also be clear from this line of argument that, in the absence of significant lock-in effects, nonintegration is always better than integration—it is optimal to do things through the market, for integration only increases the number of potential hold-ups without any compensating gains.⁴⁶

Finally, it is worth noting that the property rights approach can explain how the purchase of physical assets leads to control over human assets. To see this, consider again the GM-Fisher hypothetical. We showed that someone working with Fisher assets is more likely to

43. See *id.* at 11, 19.

44. For examples of lock-in effects, see *supra* notes 22–24 and accompanying text.

45. Klein, Crawford & Alchian, *supra* note 8, at 300. However, Klein, Crawford and Alchian fail to provide a formal justification for this notion.

46. In the above we have concentrated on ownership by an individual or by a homogeneous and monolithic group (“management”). However, the analysis can be generalized to include more complicated forms of group ownership, such as partnerships, or worker-, manager-, or consumer-cooperatives. It turns out that these will be efficient when increases in agents' marginal products are specific to a group of individuals of variable composition, rather than to a fixed group. For example, if the increase in an agent's marginal product can be realized only if the agent has access to a majority of the members of a management team, as well as to a particular asset, then it will be optimal to give each of the managers an equal ownership share in the asset and equal voting rights, and adopt majority rule. See O. Hart & J. Moore, *supra* note 32, at 19.

improve Fisher's output in a way that is specifically of value to GM if GM owns these assets than if Fisher does. This result can be expressed more informally as follows: a worker will put more weight on an actor's objectives if that actor is the worker's boss, that is, if that actor controls the assets the worker works with, than otherwise. The conclusion is quite Coasian in spirit, but the logic underlying it is very different. Coase reaches this conclusion by assuming that a boss can tell a worker what to do; in contrast, the property rights approach reaches it by showing that it is in a worker's self-interest to behave in this way, since it puts him in a stronger bargaining position with his boss later on.

To put it slightly differently, the reason an employee is likely to be more responsive to what his employer wants than a grocer is to what his customer wants is that the employer has much more leverage over his employee than the customer has over his grocer. In particular, the employer can deprive the employee of the assets he works with and hire another employee to work with these assets, while the customer can only deprive the grocer of his custom and as long as the customer is small, it is presumably not very difficult for the grocer to find another customer.

III. PROPERTY RIGHTS AND THE ESTABLISHED THEORIES OF THE FIRM

The property rights approach has features in common with each of the approaches described previously.⁴⁷ It is based on maximizing behavior (like the neoclassical approach); it emphasizes incentive issues (like the principal-agent approach); it emphasizes contracting costs (like the transaction cost approach); it treats the firm as a "standard form" contract (like the nexus of contracts approach);⁴⁸ and, it relies on the idea that a firm's owner has the right to alter membership of the firm: the owner has the right to decide who uses the firm's assets and who doesn't.⁴⁹ Its advantage over these other approaches, however, is its ability to explain both the costs and the benefits of integration; in particular, it shows how incentives change when one firm buys up another one.

Some react skeptically to the notion that a firm can be characterized completely by the nonhuman assets under its control.⁵⁰ That is, there is a feeling that one should be able to make sense of a firm as a mode of organization, even if there are no definable assets on the scene. In his analysis of GM's decision to acquire Fisher Body in 1926, Professor Klein argues that getting control over Fisher's organizational

47. See *supra* notes 2-31 and accompanying text.

48. In the language of the property rights approach, "firm" is shorthand for a collection of assets; "ownership" is shorthand for the possession of residual rights of control over these assets.

49. See Alchian & Demsetz, *supra* note 16, at 783 (manager should have right to alter membership of production team).

50. See, e.g., Klein, *supra* note 8, at 205-08.

assets rather than their physical capital was the crucial motivating factor:

By integrating with Fisher, General Motors acquired the Fisher Body organizational capital. This organization is embedded in the human capital of the employees at Fisher but is in some sense greater than the sum of its parts. The employees come and go but the organization maintains the memory of past trials and the knowledge of how to best do something (that is, how to make automobile bodies).⁵¹

Klein's conclusion is in no way inconsistent with the property rights approach. The control of physical capital can lead to control of human assets in the form of organizational capital.⁵² However, Klein appears to argue that his conclusion would hold true even if physical assets were irrelevant.⁵³ The problem with this point of view is that, in the absence of physical assets, it is unclear how GM can get control over an intangible asset like organizational capital by purchasing Fisher. For example, what is to stop Fisher management from trying to reassert control of the organizational capital after the merger? Klein writes:

A threat that all the individuals will simultaneously shirk or leave if their wages were not increased to reflect the quasi-rents on the organizational capital generally will not be credible. After vertical integration the Fisher brothers will not be able to hold up General Motors by telling all the employees to leave General Motors and show up on Monday morning at a new address.⁵⁴

This conclusion is reasonable when physical capital is important since it would be difficult at best for Fisher employees to find a substitute for this capital, particularly by Monday morning. However, it is not reasonable in the absence of physical assets. In this case, to paraphrase Alchian and Demsetz, the Fisher brothers have no more ability to hold up GM by telling all the employees to leave GM or, more generally, by countermanding GM's instructions, when Fisher is separate than when Fisher belongs to GM. Their ability to do so will be determined by factors such as the motivation, talent, knowledge and charisma of the Fisher brothers; the quality of worker information;⁵⁵ and the degree of worker inertia—factors that do not seem to have anything

51. *Id.* at 208.

52. See *supra* notes 32–46 and accompanying text. Note that the observation that the whole of organizational capital is typically greater than the sum of its parts is equivalent to the observation that the total output of a group of workers typically exceeds the sum of the workers' individual outputs, to the extent that there are complementarities.

53. Klein, *supra* note 8, at 208 n.11.

54. *Id.* at 208.

55. See G. Mailath & A. Postlewaite, *Workers Versus Firms: Bargaining Over a Firm's Value* 14 (University of Pennsylvania, Center for Analytic Research in Economics and the Social Sciences Working Paper No. 88–11, 1988).

to do with ownership structure. To put it another way, GM's response to a hold-up attempt by the Fisher brothers will be the same whether GM owns Fisher or Fisher is independent: to try to persuade Fisher workers to desert the Fisher brothers and join GM.⁵⁶

As noted previously, one of the weaknesses of the property rights approach as described here is that it does not take account of the separation of ownership and control present in large, publicly held corporations.⁵⁷ In principle, it should be possible to extend the existing analysis to such situations. A public corporation can still be usefully considered a collection of assets, with ownership providing control rights over these assets. Now, however, the picture is more complicated. Although owners (shareholders) typically retain some control rights, such as the right to replace the board of directors, in practice they delegate many others to management, at least on a day-to-day basis.⁵⁸ In addition, some of the shareholders' rights shift to creditors during periods of financial distress. Developing a formal model of the firm that contains all these features, and that includes also an explanation of the firm's financial structure, is an important and challenging task for future research. Fortunately, recent work suggests that the task is not an impossible one.⁵⁹

CONCLUSION

This Article began with the observation that the portrayal of the firm in neoclassical economics is a caricature of the modern firm. It then went on to discuss some other approaches that attempt to develop a more realistic picture. The end product to date is still, in many ways,

56. This is not without qualification. It can be argued that if GM acquires Fisher, Fisher workers become liable for damages if they try to organize a new firm since, as employees, they owe GM a duty of loyalty. See Masten, *supra* note 18, at 189. But, in practice, employees *do* leave to form new firms. Moreover, the courts facilitate this process by sometimes hesitating to enforce covenants not to compete even when such covenants are explicit. See, e.g., E. Farnsworth, *Contracts* § 5.3, at 337-38 (1982) (courts enforce non-compete covenants "only if . . . employee acquired confidential information" in course of employment). Thus, it is unclear how important this factor could have been in the GM-Fisher acquisition.

57. See *supra* note 34.

58. See, e.g., Clark, *Agency Costs Versus Fiduciary Duties*, in *Principals and Agents: The Structure of Business* 55 (J. Pratt & R. Zeckhauser eds. 1985); Easterbrook & Fischel, *Voting in Corporate Law*, 26 *J.L. & Econ.* 395 (1983); Fama & Jensen, *Separation of Ownership and Control*, 26 *J.L. & Econ.* 301 (1983).

59. See, e.g., Grossman & Hart, *One Share-One Vote and the Market for Corporate Control*, 20 *J. Fin. Econ.* 175 (1988); Harris & Raviv, *Corporate Governance: Voting Rights and Majority Rules*, 20 *J. Fin. Econ.* 203 (1988); P. Aghion & P. Bolton, *An 'Incomplete Contract' Approach to Bankruptcy and the Financial Structure of the Firm* (Stanford University, Institute for Mathematical Studies in the Social Sciences Technical Report No. 536, 1988); C. Kahn & G. Huberman, *Default, Foreclosure, and Strategic Renegotiation* (paper presented at Conference on Economics of Contract Law, Duke University, March 1988).

a caricature, but perhaps not such an unreasonable one. One promising sign is that the different approaches economists have used to address this issue—neoclassical, principal-agent, transaction cost, nexus of contracts, property rights—appear to be converging. It is to be hoped that in the next few years the best aspects of each of these approaches can be drawn on to develop a more comprehensive and realistic theory of the firm. Such a theory would capture the salient features both of modern corporations and of owner-managed firms, and would illuminate the issues for economists and lawyers alike.