

AN ECONOMIST'S PERSPECTIVE
ON THE THEORY OF THE FIRM

Oliver Hart

Discussion Paper No. 51

3/89

Program in Law and Economics
Harvard Law School
Cambridge, MA 02138

The Program in Law and Economics is supported by
a grant from the John M. Olin Foundation.

AN ECONOMIST'S PERSPECTIVE
ON THE THEORY OF THE FIRM

Oliver Hart*

*Massachusetts Institute of Technology

Helpful comments from Jeffrey Gordon, Bengt Holmstrom and Jean Tirole are gratefully acknowledged. Part of this work was done while 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 support from the Guggenheim and John M. Olin Foundations, the Center for Energy and Policy Research at MIT and the National Science Foundation.

Abstract for:

**"An Economist's Perspective on the Theory
of the Firm"**

by Oliver Hart

This article is an attempt to give lawyers a sense of how economists view the firm. Included is a discussion of the neoclassical theory, principal-agent theory, transaction cost economics, the firm as a nexus of contracts, and the recent literature on the firm as a set of property rights.

An Economist's Perspective on the Theory of the Firm

Oliver Hart

My purpose in writing this article is to give lawyers a sense of how economists think about firms. The article does not pretend to be a systematic survey of this area; rather, I shall feel free to highlight those ideas which I think are particularly important. For other perspectives on the same material, I would refer the reader to three recent surveys: Holmstrom and Tirole (1989), Milgrom and Roberts (1988), and Williamson (1989).

An outsider to the field 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. On the one hand, most formal models of the firm are extremely rudimentary, with the firm in question bearing little relation to the complex organizations we see in the world. On the other hand, those theories which attempt to incorporate real-world features of corporations, partnerships and the like often lack precision and rigor, and have therefore by and large failed to be accepted in the theoretical mainstream.

1. The Neoclassical Theory

An example of the first category--an approach which is rigorous but rudimentary--is the standard neoclassical theory of the firm. This theory, which has been developed

over the last hundred years or so, is the staple diet of modern economists and is what one will find in any modern-day textbook (see, e.g., Varian (1978)). In fact, in most textbooks, it will be the only theory of the firm presented.

According to this theory, the firm is a collection or set of feasible production plans. This production set is presided over by a manager who, buying and selling inputs and outputs in a spot market, chooses that plan which maximizes owners' welfare.¹ Welfare is usually taken to be represented by profit, or, if the firm's environment is uncertain, so that profit-maximization is not well-defined, by expected net present value of profit (possibly discounted for risk) or by market value.

To most lawyers--and to many economists, too--this is a caricature of the modern firm. Why has it survived for so long? Let me offer three reasons (in no particular order of importance): (1) The theory lends itself to an elegant and general mathematical formalization. (2) It is very useful for analyzing how a firm's production choice responds to exogenous changes in the environment, e.g., an increase in wages or a sales tax (see Varian (1978) or Bishop (1968)). (3) The theory is also very useful for analyzing the consequences of strategic interaction between firms under conditions of imperfect competition (see, e.g., Tirole (1988)); for example, it can help us to understand the relationship between the degree of concentration in an industry and that industry's output and price level.

¹ 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 a half acres to produce fifty pounds of corn.

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 a firm is. In particular, since each firm's size (or extent) is taken as given, the theory does not tell us what would happen if two firms chose to merge to become a single firm (would the assumption of profit-maximization become less reasonable?), or if one firm split itself up into two smaller firms. To put it in another way, the theory tells us nothing about the structure of firms.

2. Principal-Agent Theory

Some of the weaknesses of the neoclassical approach are addressed in an important development of the theory which has occurred in the last fifteen years: principal-agent theory.²

Principal-agent theory introduces conflicts of interest between different economic actors through the inclusion of asymmetries of information or observability problems. In this theory, the firm is still viewed as a production set, but now the production decisions are taken by a professional manager who chooses actions, e.g., effort, which are unobservable to the firm's owners, or has information about the firm's profitability that the owners do not have. In addition, the manager is assumed to have other goals than

² See, e.g., Holmstrom (1979) and Shavell (1979); or, for a recent survey, Hart and Holmstrom (1987).

just the owners' welfare, e.g., on the job perks, an easy life, empire building, etc. Under these conditions, it will be impossible for the owners to implement the profit maximizing plan directly--say, through a contract with the manager: in general, they won't 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 scheme, however, the manager will put some weight on his (or her) own objectives at the expense of the owners', and a conflict of interest will remain. Hence we have the beginnings of a managerial theory of the firm.³

Principal-agent theory enriches neoclassical theory significantly, but still fails to answer the vital question of what a firm is (or what determines the boundaries of a firm). To see why, consider the example of Fisher Body, which for many years has supplied car bodies to General Motors. 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 (since this encourages Fisher to supply high-quality input). 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

³ It is also possible to analyze extensions of the principal-agent view of the firm in which conflicts of interest between managers and workers and managers and consumers are explored. See, e.g., Calvo and Wellisz (1978).

should merge with Fisher management having authority over GM management.⁴ The point is that principal-agent theory tells us about optimal incentive schemes, but not 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, by itself, has no predictions about the nature and extent of the firm.⁵

3. Transaction Cost Economics

While the neoclassical paradigm was progressing along the above lines, a very different approach to the theory of the firm was being developed under the heading of transaction cost economics. Starting with Coase's famous 1937 article, transaction cost economics traces the existence of firms to the thinking, planning, and contracting costs which accompany any transaction, but which are usually ignored in 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 through the market. According to Coase, the main cost of transacting in the market is the cost of learning and haggling over the terms of the trade; this cost can be particularly large if the transaction is a long-term one in

⁴ As a matter of history, GM and Fisher started off as separate firms linked by a long-term contract, but after a dispute GM bought Fisher up in 1926. For interesting discussions of the relationship, see Klein, Crawford and Alchian (1978) and Klein (1988).

⁵ Drawing a parallel might be possible if, say, profit or cost-sharing arrangements were only ever found within a single firm. Then, one might conclude that GM and Fisher would have to merge. This is not the case, however. For example, consider cost-plus contracts between the U.S. Government and private defense contractors.

which learning and haggling must be performed repeatedly. This cost can be reduced by giving one party (or group) authority over the terms of trade, at least within limits. But, according to Coase, this is precisely what defines a firm: within a firm, the price mechanism is suppressed and transactions occur as a result of instructions or orders issued by a boss.⁶

Such an arrangement brings costs of its own, however: concentrating authority in one person's hands is likely to increase the cost of errors and also to lead to an additional bureaucratic burden. In Coase's view, the boundaries of the firm occur at the point where the cost savings from transacting within the firm are, at the margin, just offset by these additional error and bureaucratic costs.

Coase's ideas, although recognized as highly original from early on, took a long time to catch on.⁷ There are probably two reasons for this. First, they remain to this day very hard to formalize. Second, they contain a conceptual weakness, which was pointed out by Alchian and Demsetz in their 1972 paper. Alchian and Demsetz questioned Coase'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, i.e., an employer can tell an employee what to do. For Alchian and Demsetz the crucial question is: what ensures that the employee obeys the employer's instructions? To put it another way, what happens to the

⁶ A related idea can be found in Simon (1951).

⁷ In Coase's words, they were "much quoted but little used" (until the 1970's).

employee if he disobeys these instructions? Will he be sued for breach of contract? Unlikely. Probably, the worst that can happen is that 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, it is not clear that an employer can tell an employee what to do, any more than I can tell my grocer what to do (what vegetables to sell me at what prices); in either case, a refusal is likely to lead to a termination of the relationship, i.e., a firing. (In the case of the grocer, this means that I shop at another grocer.)⁸

So Coase's view that firms are characterized by authority relations doesn't really stand up. It is also worth noting that the second part of Coase's thesis, that the price mechanism is suppressed within firms, is also not convincing. Transfer pricing within a multidivisional firm--probably a more common phenomenon now than it was in Coase's day--seems to be a fairly immediate counterexample.

Finding Coase's characterization of the firm wanting, Alchian and Demsetz developed their own theory, based on joint production and monitoring. Alchian and Demsetz argued that transactions involving joint or team production require careful monitoring so that each agent'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:

⁸ Masten (1988) has recently pointed out that an employee owes an employer a duty of loyalty that one independent contractor does not owe another (see also Coase (1937)). This may give the employer additional leverage over the employee in some circumstances. The empirical significance of this difference is unclear, however.

1) to be a residual claimant; 2) to observe input behavior; 3) to be the central party common to all contracts to inputs; 4) to alter memberships of the team; and, 5) to sell rights 1-4 (see Alchian and Demsetz (1972), p. 783). We will return to some of Alchian and Demsetz's ideas below, but at this stage it is enough to note that their theory suffers from the same criticism that they made of Coase's : it is unclear why the problems of joint production and monitoring cannot also be solved through the market (i.e., via a contract), and in fact one does not need to look far to see examples of joint ventures or 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 becoming increasingly accepted. But what exactly was the nature of these transaction costs? What lay behind the learning and haggling costs which, according to Coase, are a major component of market transactions? The deepest and most far-reaching analysis of these costs is to be found in the work of Williamson (1975, 1985) (see also Goldberg (1976) and Klein, Crawford and Alchian (1978)). Williamson recognized that transaction costs are likely to be particularly important in situations where economic agents make relationship-specific investments, i.e., investments to some extent specific to a particular set of individuals or assets. Examples of such investments are the location of an electricity generating plant next door to a coal mine which is going to supply it; a firm's expansion of capacity to satisfy the demands of a particular customer; the training which a worker undertakes to operate a particular set of machines or to work with a particular group of individuals; or a worker's relocation to a town where he has a new job.

In situations like these, while there may be plenty of competition before the

investments are made, e.g., there may be many coal mines an electricity generating plant can locate nearby, or many towns to which a worker can move, once the investments have been sunk, the parties will be locked into each other (at least to some extent). As a result, external markets will not provide a guide as to the parties' opportunity costs once the relationship is underway. Moreover, this lack of information can be significant since, given the size and degree of the specific investment, one would expect relationships like these to be long lasting.

In an ideal world, the lack of ex-post market signals would not be a problem, since the parties could always write a long-term contract in advance of the investment spelling out each agent's obligations and the terms of the trade in every conceivable state of the world. In practice, however, thinking, negotiation, and enforcement costs will usually make such a contract prohibitively expensive. As a result, many of the terms of the relationship will have to be negotiated by the parties as they go along. Williamson argues that this leads to two sorts of costs. First, there will be ex-post costs associated with the negotiation itself; e.g., the parties may engage in collectively wasteful activities in order to try to increase their own share of the ex-post surplus; also, some gains from trade may fail to be realized because of asymmetries of information. Second, and perhaps more fundamentally, 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, i.e., a far sighted agent will choose his investment inefficiently from the point of view of the group as a whole, given that he realizes that he is liable to be held up by others (i.e., part of his investment will be expropriated) at the ex-post stage.

In Williamson's view, bringing a transaction from the market into the firm (i.e., integration) will mitigate this opportunistic behavior and improve investment incentives. The idea is that 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, integration must presumably be accompanied by some costs, since otherwise there would be a tendency for all transactions to be carried out in firms, i.e., the market would cease to be used at all. The precise nature of these costs is also left unclear, however.⁹

4. The Firm as a Nexus of Contracts

All the theories discussed so far suffer from the same weakness: while they can throw light on the nature of contractual failure, none explains in a convincing or rigorous manner how this failure is mitigated by bringing a transaction into the firm.

One reaction to this failure is to argue that it is not really a failure at all.

⁹ Williamson (1985) argues that a major benefit of integration comes from the fact that the party with authority can resolve disputes by fiat (as opposed to going through 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. Williamson is not very clear about what mechanisms are at work here, however. 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?

As will become clear below, I believe that these questions can be answered, but only by expanding the picture of a firm to include nonhuman assets under its control, as well as human assets.

According to this point of view, which is often associated with Jensen and Meckling (1976), the firm is simply a nexus of contracts, 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 simply representing different points on this continuum --i.e., particular standard form contracts. An example of such a standard form contract would be a public corporation, which is characterized by limited liability, indefinite life, and free transferability of shares (and votes). In principle, it would be possible to create a set of contracts with these characteristics each time it is needed, but, given that the characteristics are likely to be found useful in many different contexts, it is much more convenient to be able to appeal to a standard form. Closely held corporations or partnerships would be other examples of useful standard forms.

Jensen and Meckling's view of the firm as a nexus of contracts is helpful in drawing attention to the fact that contractual relations with employees, suppliers, customers, creditors, etc., 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." (Jensen and Meckling (1976) p. 311).

At the same time, the Jensen and Meckling 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. And, perhaps more fundamentally, it begs the question of what limits the set of activities covered by a

standard form. For example, let us accept that corporations are characterized by limited liability, free transferability of shares and indefinite life. One still wants to know what limits the size of a corporation, i.e., what are the economic consequences of two corporations' merging or one corporation's splitting itself into two. Given that we see mergers and breakups occurring all the time, and at considerable transaction cost, it seems unlikely that such changes are cosmetic. Presumably there are some real effects on incentives and opportunistic behavior, but what are they?

5. The Firm as a Set of Property Rights

One way to resolve the question of how integration changes incentives is spelled out in a recent literature that views the firm as a set of property rights (see Grossman and Hart (1986), Hart (1988), Hart and Moore (1988), and Holmstrom and Tirole (1989)).¹⁰ The approach is very much in the spirit of the transaction cost literature of Coase and Williamson, but differs in focusing attention on the role of physical, or at least nonhuman, assets in a contractual relationship.

In order to understand this approach, consider an economic relationship of the type analyzed by Williamson, where relationship-specific investments are important, and also, for transaction cost reasons, it is impossible to write a comprehensive long-term contract to govern the terms of the relationship (i.e., the initial contract is incomplete). Consider the non-human assets which, in the post-investment stage, make up this relationship. Given that the initial contract has gaps or missing provisions (or is

¹⁰ This literature owes a lot to the earlier property rights literature on the efficiency of private property in an externality-free world. See, e.g., Demsetz (1967).

ambiguous), states of the world will typically occur in which some aspects of the use of these assets are not specified: e.g., a contract between GM and Fisher might leave open certain aspects of maintenance policy for Fisher machines, or might neglect to 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, i.e., the owner has the right to use that asset in any way which is not inconsistent with a prior contract, a custom, or any law. So in the examples above, 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.¹¹

Finally, identify a firm with all the non-human assets that belong to it (i.e., that the firm's owners possess by virtue of being owners of the firm). Included in this category are machines, inventories, buildings or locations, and cash; and also 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 are not included, however. The reason is that, while one can argue that when GM buys up Fisher it gets control over all Fisher's physical or non-human assets, one cannot make

¹¹ This view of ownership seems consistent with the standard one adopted by lawyers. For example, Oliver Wendell Holmes [(1881/1946) page 246] writes: "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 but him."

the same argument for human assets: given anti-slavery laws, Fisher 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 agents to invest in that relationship. Hence, when contracts are incomplete, the boundaries of firms matter. 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 fall. In addition, the shift in control may lower the investment incentives of workers in the acquired firm (as we shall see). In some cases these reductions in investment will be sufficiently great that nonintegration is preferable to integration.¹²

It is also worth noting that integration in which firm A buys firm B is not the same as integration in which firm B buys firm A. The reason is that, in the former case, all the residual control rights shift to (owner-)manager A, and he can use these to hold up manager B and firm B's workers; whereas, in the latter case, all the residual control

¹² 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 it is the possession of control rights which 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 via a contract. It is only if firm A wants to acquire control over firm B's assets that it needs to integrate.

rights shift to (owner-)manager B, and he can use these to hold up (owner-) manager A and firm A's workers. In other words, in assessing the effects of integration, one needs to know not only the characteristics of the firms merging, but also who will own the merged company.

It will be helpful to illustrate these ideas in the context of the Fisher Body-General Motors relationship. Suppose these companies have an initial contract which specifies that Fisher must supply a certain number of car bodies a week. Imagine that demand for GM cars now rises and GM wants to increase the quantity Fisher supplies. Suppose also that the initial contract is silent about this possibility (say, because it was hard to predict Fisher's costs of increasing supply in advance). Then, with Fisher a separate company, GM presumably needs to get 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 so that the extra bodies are supplied; 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, when 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, when 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.¹³

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 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 them 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, Fisher management's bargaining and threat power is reduced.

Anticipating the way surplus is divided, GM will typically be much more prepared to invest in machinery which is specifically geared to Fisher bodies if it owns Fisher than if Fisher is independent, since the threat of expropriation is reduced.¹⁴ The incentives

¹³ 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.

¹⁴ 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.

for Fisher may be quite the opposite, however. Fisher management will generally be much more willing to come up with cost-saving or quality-enhancing innovations if they are an independent firm than if they are part of GM. The reason is that Fisher management is more likely to see a return on its activities if it is independent, since it can then 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 to the extent that the innovation is asset-specific rather than management-specific, and GM can threaten to hire a management team to incorporate the innovation.¹⁵

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, e.g., by learning better some aspect of the production process. Suppose further that this improvement in car body quality is specifically of interest to GM, i.e., none of

¹⁵ Under some conditions expropriation or hold-up 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 will be an insufficient tool 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, non-monetary rewards such as perks, and monetary returns which can be diverted so that they do not show up in the firm's accounts (i.e., are not verifiable).

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 only work to the extent that either GM's investment can be contracted on or Fisher can make part of the investment on GM's behalf. Otherwise, GM can use an upfront payment from Fisher to make a non-relationship-specific investment.

Fisher's other customers cares about it. There are a number of ways in which the worker might be rewarded for this, but one important one is likely to be 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, in the example given, the worker's ability to do this is likely to be greater if GM controls the assets than if Fisher does. The reason is that, in the former case, the worker will bargain directly with the party who benefits from the worker's increased skill: GM.¹⁶ In the latter case, the worker will bargain with a party--Fisher--who only receives a fraction of these benefits, since it must in turn bargain with GM to parlay these benefits into dollars (by assumption, the benefits are specific to GM). In consequence, the worker's share of the surplus will typically be lower, and so therefore will be his incentive to make the improvement in the first place.

In other words, given that the worker is liable to be held up no matter who owns the Fisher assets (assuming that he, himself, does not), it is better in terms of his incentives 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

¹⁶ This is not quite correct since the worker will actually bargain with GM management rather than with GM shareholders. 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 (see, however, footnote 19 and the comments in Section 6).

product.¹⁷ As a result, we might expect the worker to get, say, a third of his increased marginal product (equal division with Fisher and GM). With GM management in control of the Fisher assets, the potential hold-ups are reduced to one, since the power to deny the worker his increased marginal product is concentrated in one agent's hands. As a result, we might expect the worker in this case to get, say, half of his increased marginal product (equal division with GM).¹⁸

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, e.g., the worker learns how to reduce Fisher management's costs of making car bodies, regardless of Fisher's final customer (but the cost reduction cannot 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 return bargain with Fisher management to benefit from the reduction in costs.

So far we have taken it as a given that GM management will control GM assets. However, this need not be the case; in some situations (maybe unrealistic ones) it might make more sense for Fisher management to control these assets, i.e., for Fisher to buy

¹⁷ We are assuming that the initial contract was incomplete in the sense that no payment was specified for the improved product.

¹⁸ For a formal analysis of this, see Hart and Moore (1988).

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. The reason is that, 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. Giving control of these assets to two different management teams is therefore bound to be bad, since it increases the number of parties with hold-up power (for details, see Hart and Moore (1988)). This result confirms Klein, Crawford and Alchian's insight that when lock-in effects are extreme, integration will dominate non-integration; Klein, Crawford and Alchian do not provide a formal justification for their conclusions, however.

These ideas can be used to build up a theory of the firm's boundaries. First, as we have seen, highly complementary assets should be owned in common, which may provide minimum size to the firm. Second, as the firm grows beyond a certain point, the manager at the center will become less and less important as far as operations at the periphery are concerned; 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 there without any compensating gains. It should also be clear from this line of argument that, in the absence of significant lock-in effects, non-integration is always better than integration, i.e., it is optimal to do things through the market. Again, the reason is that integration only increases the number of

potential hold-ups without any compensating gains.¹⁹

To conclude this section, note that the property rights approach has the following important implication: the purchase of physical assets leads to control over human assets. To see this, consider again the GM-Fisher example. There it was shown that someone working with Fisher assets is more likely to improve Fisher's output in a way that is specifically of value to GM if GM owns these assets than if Fisher does (the reverse conclusion holds with respect to improvements specifically of value to Fisher). This result can be expressed more informally as follows: a worker will put more weight on an agent's objectives if that agent is the worker's boss, i.e., if that agent 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 obtains the conclusion by assuming that a boss can tell a worker what to do; in contrast, the property rights approach obtains 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

¹⁹ 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 (i.e., vote) in the asset and adopt majority rule. For details, see Hart and Moore (1988).

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).

6. Remarks

The property rights approach has features in common with each of the approaches described previously. It is based on maximizing behavior (like the neoclassical approach); emphasizes incentive issues (like the principal-agent approach); emphasizes contracting costs (like the transaction cost approach); treats the firm as a standard form contract (as in Jensen and Meckling (1976));²⁰ and relies on the idea that a firm's owner has the right to alter membership of the firm, i.e., the owner has the right to decide who uses the firm's assets and who doesn't (as in Alchian and Demsetz (1972)). It has the advantage over these other approaches, however, in being able to explain both the costs and the benefits of integration; in particular, it shows how incentives change when one firm buys up another one.

One reaction that some people have to the property rights approach is a scepticism that a firm can be characterized completely by the non-human assets under its control. 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. For

²⁰ The firm is shorthand for a collection of assets and ownership is shorthand for the possession of residual rights of control over these assets.

example, Klein (1988), in his analysis of GM's decision to acquire Fisher Body in 1926, argues that getting control over Fisher's organizational assets rather than their physical capital was the crucial factor. Klein writes:

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 best to do something (that is, to make automobile bodies).

Klein's conclusion is in no way inconsistent with the property rights approach, since, as we have seen, that approach shows that control of physical capital can lead to control of human assets, i.e., organizational capital (the observation that the whole of organizational capital is typically greater than the sum of its parts is equivalent to the observation that total output of a group of workers typically exceeds the sum of the workers' individual outputs, to the extent that there are complementarities). However, Klein appears to argue that his conclusion would hold true even if physical assets were irrelevant. 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.

This conclusion is reasonable when physical capital is important, since it would be hard (impossible?) 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 yet again, 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,²¹ and the degree of worker inertia--factors which do not seem to have anything 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.²²

Before we conclude, an important lacuna in the property rights approach should be mentioned. As it stands, the approach makes no distinction between ownership and control. That is, throughout we have talked as if those who own assets also manage them. In most of the formal models which have been developed, such an arrangement

²¹ See Mailath and Postlewaite (1988).

²² 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 (1988). Since employees do in practice leave to form new firms, and the courts often do not enforce covenants not to compete even when they are explicit, it is unclear how important this factor could have been in the GM-Fisher acquisition.

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 open corporations.

In principle, it ought to be possible to extend the existing models to public corporations, too. A public corporation is still usefully thought of as a collection of assets, with ownership providing control rights over these assets. Now, however, the picture is more complicated since, although owners (i.e., shareholders) typically retain some control rights, e.g., the right to replace the board of directors, in practice they delegate many others to management, at least on a day-to-day basis.²³ In addition, some of the shareholders' rights shift to creditors during periods of financial distress. Developing a formal model of the firm with all these features--including the determination 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.²⁴

²³ See, e.g., Clark (1985), Easterbrook and Fischel (1983), and Fama and Jensen (1983).

²⁴ See, e.g., Aghion and Bolton (1988), Harris and Raviv (1988), Huberman and Kahn (1988), and Grossman and Hart (1988).

7. Conclusions

I began this article by observing that the portrayal of the firm in neoclassical economics is a caricature of the modern firm. I then went on to discuss some other approaches which try to develop a more realistic picture. The end product to date is still, in many ways, a caricature, but perhaps not quite such an unreasonable one. One promising sign is that the different approaches which economists have used to address this issue--neoclassical, principal-agent, and transaction cost--appear to be converging. The hope is 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; one that captures the salient features of modern corporations as well as owner-managed firms, and is illuminating to economists and lawyers alike.

REFERENCES

- Aghion, P. and P. Bolton (1986), "An 'Incomplete Contracts' Approach to Bankruptcy and the Optimal Financial Structure of the Firm," Harvard University mimeo.
- Alchian, A. and H. Demsetz (1972), "Production Information Costs and Economic Organization," *American Economic Review*, 777-95.
- Bishop, R.L. (1968), "The Effects of Specific and ad Valorem Taxes," *Quarterly Journal of Economics*, 82: 198-218.
- Calvo, G. and S. Wellisz (1978), "Supervision, Loss of Control and the Optimal Size of the Firm," *Journal of Political Economy*, 86: 943-52.
- Clark, R. (1985), "Agency Costs vs. Fiduciary Duties," in *Principals and Agents: the Structure of Business*, J. Pratt and R. Zeckhauser (eds.), Harvard Business School Press: Cambridge MA.
- Coase, R. (1937), "The Nature of the Firm," *Economica*, 4, 386-405, reprinted in G. Stigler and K. Boulding (eds.) (1952), *Readings in Price Theory*, Richard D. Irwin: Homewood, IL.
- Demsetz, H. (1967), "Toward a Theory of Property Rights," *American Economic Review*, 57: 347-48.
- Easterbrook, F. and D. Fischel (1983), "Voting in Corporate Law," *Journal of Law and Economics*, 26: 395-428.
- Goldberg, V. (1976), "Regulation and Administered Contracts," *Bell Journal of Economics*, 7: 426-48.
- Grossman, S. and O. Hart (1986), "The Costs and Benefits of Ownership: a Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94: 691-719.
- Grossman, S. and O. Hart (1988), "One Share-One Vote and the Market for Corporate Control," *Journal of Financial Economics*, 20: 175-202.
- Harris, M. and A. Raviv (1988), "Corporate Governance: Voting Rights and Majority Rules," *Journal of Financial Economics*, 20: 203-235.
- Hart, O. and J. Moore (1988), "Property Rights and the Nature of the Firm," MIT Discussion Paper.

- Hart, O. (1988), "Incomplete Contracts and the Theory of the Firm," *Journal of Law, Economics, and Organization*, 4:1, 119-139.
- Hart, O. and B. Holmstrom (1987), "The Theory of Contracts," in *Advances in Economic Theory*, T. Bewley (ed.), Fifth World Congress, Cambridge: Cambridge University Press.
- Holmes, O.W., *The Common Law*, 1881 Reprint, Boston: Little Brown, 1946.
- Holmstrom, B. (1979), "Moral Hazard and Observability," *Bell Journal of Economics*, 10: 74-91.
- Holmstrom, B., and J. Tirole (1989), "The Theory of the Firm," in the *Handbook of Industrial Organization*, R. Schmalensee and R. Willig (eds.), Amsterdam: North-Holland.
- Huberman, G. and C. Kahn (1988), "Default, Foreclosure, and Strategic Renegotiation," Mimeo, University of Chicago.
- Jensen, M. and W. Meckling (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3, 305-360.
- Klein, B. (1988), "Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited," *Journal of Law, Economics, and Organization*, 4:1, 199-213.
- Klein, B., R. Crawford and A. Alchian (1978), "Vertical Integration, Appropriable Rents and the Competitive Contracting Process," *Journal of Law and Economics*, 21: 297-326.
- Mailath, G. and A. Postlewaite (1988), "Workers Versus Firms: Bargaining Over a Firm's Value," University of Pennsylvania Working Paper #88-11.
- Masten, S. (1988), "A Legal Basis for the Firm," *Journal of Law, Economics, and Organization*, 4: 181-98.
- Milgrom, P. and J. Roberts (1988), "Economic Theories of the Firm: Past, Present, and Future," *Canadian Journal of Economics*, XXI:3, 444-458.
- Shavell, S. (1979), "Risk Sharing and Incentives in the Principal and Agent Relationship," *Bell Journal of Economics*, 10: 55-73.
- Simon, H. (1951), "A Formal Theory of the Employment Relation," *Econometrica*, 19: 293-305.

Tirole, J. (1988), *The Theory of the Industrial Organization*, MIT Press, Cambridge, MA.

Varian, H. (1978), *Microeconomic Analysis*, W.W. Norton and Company, New York.

Williamson, O. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*, Free Press: New York.

Williamson, O. (1985), *The Economic Institutions of Capitalism*, Free Press: New York.

Williamson, O. (1988), "The Logic of Economic Organization," *Journal of Law, Economics, and Organization*, 4:1, 65-94.