

## An economist's perspective on the theory of the firm

OLIVER HART

Oliver Hart was born in London, England, in 1948. He received a Ph.D. in economics from Princeton University in 1974. When this article was published, he was Professor of Economics at the Massachusetts Institute of Technology. Since 1993, he has taught at Harvard University, where he is currently the Andrew E. Furer Professor of Economics.

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.<sup>1</sup> 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.

Excerpted from Oliver Hart, "An Economist's Perspective on the Theory of the Firm." This article originally appeared at 89 *Columbia Law Review* 1757 (1989). Reprinted by permission.

<sup>1</sup> Several recent surveys provide other perspectives on this material. See, e.g., Holmstrom and Tirole (1989), Milgrom and Roberts (1988b), Williamson (1988).

## **I. Established theories**

...

### *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, 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. 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.

...

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. 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. 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.

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 under way.

...

In Williamson's view, bringing a transaction from the market into the firm – the phenomenon of integration – mitigates... 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.

...

## 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. 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.

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 own human capital both before and after any 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.<sup>2</sup> 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.<sup>3</sup>

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

<sup>2</sup> 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 infranotes 58–59 [ed.: notes 5 and 6, this version] and accompanying text.

<sup>3</sup> 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.

rights in the new firm. *A* can then use those rights to hold up the managers and workers of firm *B*. Should the 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*.

...

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 holdup 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 holdups without any compensating gains.<sup>4</sup>

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 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.

<sup>4</sup> 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 Hart and Moore (1988), p. 19.

### *An economist's perspective on the theory of the firm*

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**

...

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. 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.<sup>5</sup> In addition, some of the shareholders' rights sift 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.<sup>6</sup>

### **Conclusion**

This article began with the observation that the portrayal of the firm in neo-classical 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, a caricature, but perhaps not such an unreasonable one. One promising sign is that the different approaches

<sup>5</sup> See, e.g., Clark (1985), Easterbrook and Fischel (1983), Fama and Jensen (1983b).

<sup>6</sup> See, e.g., Grossman and Hart (1988), Harris and Raviv, (1988), Aghion and Bolton (1988), Kahn and Huberman (1988).

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.