The Nature of the Firm

Origins, Evolution, and Development

Edited by

OLIVER E. WILLIAMSON

SIDNEY G. WINTER

New York Oxford
OXFORD UNIVERSITY PRESS 1993

Kenney, Roy, and Benjamin Klein. 1985. "The Law and Economics of Contractual Flexibility," UCLA Department of Economics Working paper no. 388.

Klein, Benjamin. 1983. "Contracting Costs and Residual Claims: The Separation of Ownership and Control," 26 *Journal of Law and Economics* 367.

———, Robert Crawford, and Armen Alchian. 1978. "Vertical Integration, Appropriable Rents and the Competitive Contracting Process," 21 *Journal of Law and Economics* 297.

Monteverde, K. M., and David J. Teece. 1982. "Appropriable Rents and Quasi-Vertical Integration," 25 *Journal of Law and Economics* 321.

Nelson, Richard R., and Sidney G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge: Harvard University Press.

Prescott, Edward C., and Michael Visscher. 1980. "Organizational Capital," 88 *Journal of Political Economy* 466.

Rosen, Sherwin. 1972. "Learning by Experience as Joint Production," 86 *Quarterly Journal of Economics* 366.

Stigler, George. 1951. "The Division of Labor Is Limited by the Extent of the Market," 59 *Journal of Political Economy* 185.

Williamson, Oliver E. 1979. "Transaction Cost Economics: The Governance of Contractual Relations," 22 *Journal of Law and Economics* 233.

1991 Nobel Lecture: The Institutional Structure of Production

R. H. COASE

In my long life I have known some great economists but I have never counted myself among their number nor walked in their company. I have made no innovations in high theory. My contribution to economics has been to urge the inclusion in our analysis of features of the economic system so obvious that, like the postman in G. K. Chesterton's Father Brown tale, "The Invisible Man," they have tended to be overlooked. Nonetheless, once included in the analysis, they will, as I believe, bring about a complete change in the structure of economic theory, at least in what is called price theory or microeconomics. What I have done is to show the importance for the working of the economic system of what may be termed the institutional structure of production. In this lecture I shall explain why, in my view, these features of the economic system were ignored and why their recognition will lead to a change in the way we analyse the working of the economic system and in the way we think about economic policy, changes which are already beginning to occur. I will also speak about the empirical work that needs to be done if this transformation in our approach is to increase our understanding. In speaking about this transformation, I do not wish to suggest that it is the result of my work alone. Oliver Williamson, Harold Demsetz, Steven Cheung, among others, have made outstanding contributions to the subject and without their work and that of many others, I doubt whether the significance of my writings would have been recognized. While it has been a great advantage of the creation of the Prize in Economic Sciences in Memory of Alfred Nobel that, by drawing attention to the significance of particular fields of economics, it encourages further research in them, the highlighting of the work of a few scholars, or, in my case, one scholar, tends to obscure the importance of the contributions of other able scholars whose researches have been crucial to the development of the field.

I will be speaking of that part of economics which has come to be called industrial organization but, to understand its present state, it is necessary to say

something about the development of economics in general. During the two centuries since the publication of *The Wealth of Nations*, the main activity of economists, it seems to me, has been to fill the gaps in Adam Smith's system, to correct his errors and to make his analysis vastly more exact. A principal theme of *The Wealth of Nations* was that government regulation or centralised planning were not necessary to make an economic system function in an orderly way. The economy could be co-ordinated by a system of prices (the "invisible hand") and, furthermore, with beneficial results. A major task of economists since the publication of *The Wealth of Nations*, as Harold Demsetz has explained,[1] has been to formalize this proposition of Adam Smith. The given factors are technology and the tastes of consumers, and individuals, who follow their own interest, are governed in their choices by a system of prices. Economists have uncovered the conditions necessary if Adam Smith's results are to be achieved and where, in the real world, such conditions do not appear to be found, they have proposed changes which are designed to bring them about. It is what one finds in the textbooks. Harold Demsetz has said rightly that what this theory analyses is a system of extreme decentralization. It has been a great intellectual achievement and it throws light on many aspects of the economic system. But it has not been by any means all gain. The concentration on the determination of prices has led to a narrowing of focus which has had as a result the neglect of other aspects of the economic system. Sometimes, indeed, it seems as though economists conceive of their subject as being concerned only with the pricing system and anything outside this is considered as no part of their business. Thus, my old chief and wonderful human being, Lionel Robbins, wrote, in *The Nature and Significance of Economic Science*, about the "glaring deficiencies" of the old treatment of the theory of production with its discussion of peasant proprietorships and industrial forms: "It suggests that from the point of view of the economist 'organisation' is a matter of internal industrial (or agricultural) arrangement—if not internal to the firm, at any rate internal to 'the' industry. At the same time it tends to leave out completely the governing factor of all productive organisation—the relationship of prices and cost. . . ."[2] What this comes down to is that, in Robbins' view, an economist does not interest himself in the internal arrangements within organisations but only in what happens on the market, the purchase of factors of production and the sale of the goods that these factors produce. What happens in between the purchase of the factors of production and the sale of the goods that are produced by these factors is largely ignored. I do not know how far economists today share Robbins' attitude but it is undeniable that microeconomics is largely a study of the determination of prices and output, indeed this part of economics is often called price theory.

This neglect of other aspects of the system has been made easier by another feature of modern economic theory—the growing abstraction of the analysis, which does not seem to call for a detailed knowledge of the actual economic system or, at any rate, has managed to proceed without it. Holmstrom and Tirole writing on "The Theory of the Firm" in the recently published *Handbook of Industrial Organization*, conclude at the end of their article of 63 pages that "the evidence/theory ratio . . . is currently very low in this field."[3] Peltzman has written a scathing review of the *Handbook* in which he points out how much of the discussion in it is theory without any empirical basis.[4] What is studied is a system which lives in the minds of economists but not on earth. I have called the result "blackboard economics." The firm and the market appear by name but they lack any substance. The firm in mainstream economic theory has often been described as a "black box." And so it is. This is very extraordinary given that most resources in a modern economic system are employed within firms, with how these resources are used dependent on administrative decisions and not directly on the operation of a market. Consequently, the efficiency of the economic system depends to a very considerable extent on how these organisations conduct their affairs, particularly, of course, the modern corporation. Even more surprising, given their interest in the pricing system, is the neglect of the market or more specifically the institutional arrangements which govern the process of exchange. As these institutional arrangements determine to a large extent what is produced, what we have is a very incomplete theory. All this is beginning to change and in this process I am glad to have played my part. The value of including such institutional factors in the corpus of mainstream economics is made clear by recent events in Eastern Europe. These ex-communist countries are advised to move to a market economy, and their leaders wish to do so, but without the appropriate institutions no market economy of any significance is possible. If we knew more about our own economy, we would be in a better position to advise them.

What I endeavoured to do in the two articles cited by the Royal Swedish Academy of Sciences was to attempt to fill these gaps or more exactly to indicate the direction in which we should move if they are ultimately to be filled. Let me start with "The Nature of the Firm." I went as a student to the London School of Economics in 1929 to study for a Bachelor of Commerce degree, specialising in the Industry group, supposedly designed for people who wished to become works managers, a choice of occupation for which I was singularly ill-suited. However, in 1931 I had a great stroke of luck. Arnold Plant was appointed Professor of Commerce in 1930. He was a wonderful teacher. I began to attend his seminar in 1931, some five months before I took the final examinations. It was a revelation. He quoted Sir Arthur Salter: "The normal economic system works itself." And he explained how a competitive economic system co-ordinated by prices would lead to the production of goods and services which consumers valued most highly. Before being exposed to Plant's teaching, my notions on how the economy worked were extremely woolly. After Plant's seminar I had a coherent view of the economic system. He introduced me to Adam Smith's "invisible hand." As I had taken the first year of University work while still at High School, I managed to complete the requirements for a degree in two years. However, University regulations required three years of residence before a degree could be granted. I had therefore a year to spare. I then had another stroke of luck. I was awarded a Cassel travelling scholarship by the University of London. I decided to spend the year in the United States, this being treated as a year's residence at the London School of Economics, the regulations being somewhat loosely interpreted.

I decided to study vertical and lateral integration of industry in the United States. Plant had described in his lectures the different ways in which various industries were organised but we seemed to lack any theory which would explain these differences. I set out to find it. There was also another puzzle which, in my mind, needed to be solved and which seemed to be related to my main project. The view of the pricing system as a co-ordinating mechanism was clearly right but there were aspects of the argument which troubled me. Plant was opposed to all schemes, then very fashionable during the Great Depression, for the co-ordination of industrial production by some form of planning. Competition, according to Plant, acting through a system of prices, would do all the co-ordination necessary. And yet we had a factor of production, management, whose function was to co-ordinate. Why was it needed if the pricing system provided all the co-ordination necessary? The same problem presented itself to me at that time in another guise. The Russian Revolution had taken place only fourteen years earlier. We knew then very little about how planning would actually be carried out in a communist system. Lenin had said that the economic system in Russia would be run as one big factory. However, many economists in the West maintained that this was an impossibility. And yet there were factories in the West and some of them were extremely large. How did one reconcile the views expressed by economists on the role of the pricing system and the impossibility of successful central economic planning with the existence of management and of these apparently planned societies, firms, operating within our own economy?[5]

I found the answer by the summer of 1932. It was to realise that there were costs of using the pricing mechanism. What the prices are have to be discovered. There are negotiations to be undertaken, contracts have to be drawn up, inspections have to be made, arrangements have to be made to settle disputes, and so on. These costs have come to be known as transaction costs. Their existence implies that methods of co-ordination alternative to the market, which are themselves costly and in various ways imperfect, may nonetheless be preferable to relying on the pricing mechanism, the only method of co-ordination normally analysed by economists. It was the avoidance of the costs of carrying out transactions through the market that could explain the existence of the firm in which the allocation of factors came about as a result of administrative decisions (and I thought it did). In my 1937 article I argued that in a competitive system there would be an optimum of planning since a firm, that little planned society, could only continue to exist if it performed its co-ordination function at a lower cost than would be incurred if it were achieved by means of market transactions and also at a lower cost than this same function could be performed by another firm. To have an efficient economic system it is necessary not only to have markets but also areas of planning within organizations of the appropriate size. What this mix should be we find as a result of competition. This is what I said in my article of 1937. However, as we know from a letter I wrote in 1932 which has been preserved, all the essentials of this argument had been presented in a lecture I gave in Dundee at the beginning of October, 1932.[6] I was then twenty-one years of age and the sun never ceased to shine. I could never have imagined that these ideas would become

some 60 years later a major justification for the award of a Nobel Prize. And it is a strange experience to be praised in my eighties for work I did in my twenties.

There is no doubt that the recognition by economists of the importance of the role of the firm in the functioning of the economy will prompt them to investigate its activities more closely. The work of Oliver Williamson and others has led to a greater understanding of the factors which govern what a firm does and how it does it. And we can also hope to learn much more in future from the studies of the activities of firms which have recently been initiated by the Center for Economic Studies of the Bureau of the Census of the United States. But it would be wrong to think that the most important consequence for economics of the publication of "The Nature of the Firm" has been to direct attention to the importance of the firm in our modern economy, a result which, in my view, would have come about in any case. What I think will be considered in future to have been the important contribution of this article is the explicit introduction of transaction costs into economic analysis. I argued in "The Nature of the Firm" that the existence of transaction costs leads to the emergence of the firm. But the effects are pervasive in the economy. Businessmen in deciding on their ways of doing business and on what to produce have to take into account transaction costs. If the costs of making an exchange are greater than the gains which that exchange would bring, that exchange would not take place and the greater production that would flow from specialisation would not be realised. In this way transaction costs affect not only contractual arrangements but also what goods and services are produced. Not to include transaction costs in the theory leaves many aspects of the working of the economic system unexplained, including the emergence of the firm, but much else besides. In fact, a large part of what we think of as economic activity is designed to accomplish what high transaction costs would otherwise prevent or to reduce transaction costs so that individuals can freely negotiate and we can take advantage of that diffused knowledge of which Hayek has told us.

I know of only one part of economics in which transaction costs have been used to explain a major feature of the economic system and that relates to the evolution and use of money. Adam Smith pointed out the hindrances to commerce that would arise in an economic system in which there was a division of labour but in which all exchange had to take the form of barter. No one would be able to buy anything unless he possessed something that the producer wanted. This difficulty, he explained, could be overcome by the use of money. A person wishing to buy something in a barter system has to find someone who has this product for sale but who also wants some of the goods possessed by the potential buyer. Similarly, a person wishing to sell something has to find someone who both wants what he has to offer and also possesses something that the potential seller wants. Exchange in a barter system requires what Jevons called "this double coincidence." Clearly the search for partners in exchange with suitable qualifications is likely to be very costly and will prevent many potentially beneficial exchanges from taking place. The benefit brought about by the use of money consists of a reduction in transaction costs. The use of money also reduces transaction costs by facilitating the drawing up of contracts as well as by reducing the quantity of goods that need

to be held for purposes of exchange. However, the nature of the benefits secured by the use of money seems to have faded into the background so far as economists are concerned and it does not seem to have been noticed that there are other features of the economic system which exist because of the need to mitigate transaction costs.

I now turn to that other article cited by the Swedish Academy, "The Problem of Social Cost," published some 30 years ago. I will not say much here about its influence on legal scholarship, which has been immense, but will mainly consider its influence on economics, which has not been immense, although I believe that in time it will be. It is my view that the approach used in that article will ultimately transform the structure of microeconomics—and I will explain why. I should add that in writing this article I had no such general aim in mind. I thought that I was exposing the weaknesses of Pigou's analysis of the divergence between private and social products, an analysis generally accepted by economists, and that was all. It was only later, and in part as a result of conversations with Steven Cheung in the 1960's that I came to see the general significance for economic theory of what I had written in that article and also to see more clearly what questions needed to be further investigated.

Pigou's conclusion and that of most economists using standard economic theory was (and perhaps still is) that some kind of government action (usually the imposition of taxes) was required to restrain those whose actions had harmful effects on others (often termed negative externalities). What I showed in that article, as I thought, was that in a regime of zero transaction costs, an assumption of standard economic theory, negotiations between the parties would lead to those arrangements being made which would maximise wealth and this irrespective of the initial assignment of rights. This is the infamous Coase Theorem, named and formulated by Stigler, although it is based on work of mine. Stigler argues that the Coase Theorem follows from the standard assumptions of economic theory. Its logic cannot be questioned, only its domain.[7] I do not disagree with Stigler. However, I tend to regard the Coase Theorem as a stepping stone on the way to an analysis of an economy with positive transaction costs. The significance to me of the Coase Theorem is that it undermines the Pigovian system. Since standard economic theory assumes transaction costs to be zero, the Coase Theorem demonstrates that the Pigovian solutions are unnecessary in these circumstances. Of course, it does not imply, when transaction costs are positive, that government actions (such as government operation, regulation or taxation, including subsidies) could not produce a better result than relying on negotiations between individuals in the market. Whether this would be so could be discovered not by studying imaginary governments but what real governments actually do. My conclusion: let us study the world of positive transaction costs.

If we move from a regime of zero transaction costs to one of positive transaction costs, what becomes immediately clear is the crucial importance of the legal system in this new world. I explained in "The Problem of Social Cost" that what are traded on the market are not, as is often supposed by economists, physical entities but the rights to perform certain actions and the rights which individuals possess are established by the legal system. While we can imagine in the hypothetical world of zero transaction costs that the parties to an exchange would negotiate to change any provision of the law which prevents them from taking whatever steps are required to increase the value of production, in the real world of positive transaction costs, such a procedure would be extremely costly, and would make unprofitable, even where it was allowed, a great deal of such contracting around the law. Because of this, the rights which individuals possess, with their duties and privileges, will be, to a large extent, what the law determines. As a result, the legal system will have a profound effect on the working of the economic system and may in certain respects be said to control it. It is obviously desirable that these rights should be assigned to those who can use them most productively and with incentives that lead them to do so and that, to discover (and maintain) such a distribution of rights, the costs of their transference should be low, through clarity in the law and by making the legal requirements for such transfers less onerous. Since this can come about only if there is an appropriate system of property rights (and they are enforced), it is easy to understand why so many academic lawyers (at least in the United States) have found so attractive the task of uncovering the character of such a property rights system and why the subject of "law and economics" has flourished in American law schools. Indeed, work is going forward at such a pace that I do not consider it overoptimistic to believe that the main outlines of the subject will be drawn within five or ten years.

Until quite recently most economists seem to have been unaware of this relationship between the economic and legal systems except in the most general way. Stock and produce exchanges are often used by economists as examples of perfect or near-perfect competition. But these exchanges regulate in great detail the activities of traders (and this quite apart from any public regulation there may be). What can be traded, when it can be traded, the terms of settlement and so on are all laid down by the authorities of the exchange. There is, in effect, a private law. Without such rules and regulations, the speedy conclusion of trades would not be possible. Of course, when trading takes place outside exchanges (and this is almost all trading) and where the dealers are scattered in space and have very divergent interests, as in retailing and wholesaling, such a private law would be difficult to establish and their activities will be regulated by the laws of the State. It makes little sense for economists to discuss the process of exchange without specifying the institutional setting within which the trading takes place since this affects the incentives to produce and the costs of transacting. I think this is now beginning to be recognized and has been made crystal-clear by what is going on in Eastern Europe today. The time has surely gone in which economists could analyse in great detail two individuals exchanging nuts for berries on the edge of the forest and then feel that their analysis of the process of exchange was complete, illuminating though this analysis may be in certain respects. The process of contracting needs to be studied in a real world setting. We would then learn of the problems that are encountered and of how they are overcome and we would certainly become aware of the richness of the institutional alternatives between which we have to choose.

Oliver Williamson has ascribed the non-use or limited use of my thesis in "The Nature of the Firm" to the fact that it has not been made "operational," by which he means that the concept of transaction costs has not been incorporated into a general theory. I think this is correct. There have been two reasons for this. First, incorporating transaction costs into standard economic theory, which has been based on the assumption that they are zero, would be very difficult and economists who, like most scientists, as Thomas Kuhn has told us, are extremely conservative in their methods, have not been inclined to attempt it. Second, Williamson has also pointed out that although I was correct in making the choice between organization within the firm or through the market the center piece of my analysis, I did not indicate what the factors were that determined the outcome of this choice and thus made it difficult for others to build on what is often described as a "fundamental insight." This also is true. But the interrelationships which govern the mix of market and hierarchy, to use Williamson's terms, are extremely complex and in our present state of ignorance it will not be easy to discover what these factors are. What we need is more empirical work. In a paper written for a conference of the National Bureau of Economic Research I explained why I thought this was so. This is what I said: "An inspired theoretician might do as well without such empirical work, but my own feeling is that the inspiration is most likely to come through the stimulus provided by the patterns, puzzles and anomalies revealed by the systematic gathering of data, particularly when the prime need is to break our existing habits of thought."[8] This statement was made in 1970. I still think that in essentials it is true today. Although much interesting and important research was done in the seventies and eighties and we certainly know much more than we did in 1970, there is little doubt that a great deal more empirical work is needed. However, I have come to the conclusion that the main obstacle faced by researchers in industrial organization is the lack of available data on contracts and the activities of firms. I have therefore decided to do something about it.

Believing that there is a great deal of data on contracts and the activities of firms in the United States available in government departments and agencies in Washington, D. C., and that this information is largely unknown to economists, I organized a conference at the University of Chicago Law School in the summer of 1990 at which government officials presented papers in which they described what data was available and how to get access to it and also reported on some of the research being carried out within their departments. The audience consisted of academic economists. It was, as a colleague remarked, a case of supply meeting demand. The proceedings of this conference will be published in a special issue of the *Journal of Law and Economics*. Another development with which I am associated is the establishment of the Center for the Study of Contracts and the Structure of Enterprise at the Business School of the University of Pittsburgh. This Center will make large-scale collections of business contracts and will prepare databases which will be made available to all researchers, whatever their institution. Nor should we forget the work now getting started at the Center for Economic Studies of the Bureau of the Census. This greater availability of data and the encouragement given to all researchers working on the institutional structure of production by the award to me of the Nobel Prize should result in a reduction in that elegant but sterile theorizing so commonly found in the economics literature on industrial organization and should lead to studies which increase our understanding of how the real economic system works.

My remarks have sometimes been interpreted as implying that I am hostile to the mathematization of economic theory. This is untrue. Indeed, once we begin to uncover the real factors affecting the performance of the economic system, the complicated interrelations between them will clearly necessitate a mathematical treatment, as in the natural sciences, and economists like myself, who write in prose, will take their bow. May this period soon come.

I am very much aware that many economists whom I respect and admire will not agree with the opinions I have expressed and some may even be offended by them. But a scholar must be content with the knowledge that what is false in what he says will soon be exposed and, as for what is true, he can count on ultimately seeing it accepted, if only he lives long enough.

NOTES

1. Harold Demsetz, *Ownership, Control and the Firm* Volume 1, page 145.
2. Lionel Robbins, *The Nature and Significance of Economic Science* (1932), page 70.
3. Richard Schmalensee and Robert D. Willig (editors), *Handbook of Industrial Organization*, page 126.
4. Sam Peltzman, "The Handbook of Industrial Organization: A Review Article," *Journal of Political Economy*, February, 1991, pages 201–217.
5. A fuller account of these events will be found in Oliver E. Williamson and Sidney G. Winter (editors), *The Nature of the Firm, Origins, Evolution and Development*, pages 34–47.
6. *Ibid*, pages 34–35.
7. George J. Stigler, "Two Notes on the Coase Theorem," *Yale Law Journal*, December, 1989, pages 631–633.
8. R.H. Coase, *The Firm, the Market and the Law*, page 71.