

## 2

### The Nature of the Firm (1937)

R. H. COASE

Economic theory has suffered in the past from a failure to state clearly its assumption. Economists in building up a theory have often omitted to examine the foundations on which it was erected. This examination is, however, essential not only to prevent the misunderstanding and needless controversy which arise from a lack of knowledge of the assumptions on which a theory is based, but also because of the extreme importance for economics of good judgment in choosing between rival sets of assumptions. For instance, it is suggested that the use of the word "firm" in economics may be different from the use of the term by the "plain man."<sup>1</sup> Since there is apparently a trend in economic theory towards starting analysis with the individual firm and not with the industry,<sup>2</sup> it is all the more necessary not only that a clear definition of the word "firm" should be given but that its difference from a firm in the "real world," if it exists, should be made clear. Mrs. Robinson has said that "the two questions to be asked of a set of assumptions in economics are: Are they tractable? and: Do they correspond with the real world?"<sup>3</sup> Though, as Mrs. Robinson points out, "More often one set will be manageable and the other realistic," yet there may well be branches of theory where assumptions may be both manageable and realistic. It is hoped to show in the following paper that a definition of a firm may be obtained which is not only realistic in that it corresponds to what is meant by a firm in the real world, but is tractable by two of the most powerful instruments of economic analysis developed by Marshall, the idea of the margin and that of substitution, together giving the idea of substitution at the margin.<sup>4</sup> Our definition must, of course, "relate to formal relations which are capable of being *conceived* exactly."<sup>5</sup>

It is convenient if, in searching for a definition of a firm, we first consider the economic system as it is normally treated by the economist. Let us consider the description of the economic system given by Sir Arthur Salter.<sup>6</sup> "The normal eco-

economic system works itself. For its current operation it is under no central control, it needs no central survey. Over the whole range of human activity and human need, supply is adjusted to demand, and production to consumption, by a process that is automatic, elastic and responsive." An economist thinks of the economic system as being co-ordinated by the price mechanism and society becomes not an organization but an organism.<sup>7</sup> The economic system "works itself." This does not mean that there is no planning by individuals. These exercise foresight and choose between alternatives. This is necessarily so if there is to be order in the system. But this theory assumes that the direction of resources is dependent directly on the price mechanism. Indeed, it is often considered to be an objection to economic planning that it merely tries to do what is already done by the price mechanism.<sup>8</sup> Sir Arthur Salter's description, however, gives a very incomplete picture of our economic system. Within a firm, the description does not fit at all. For instance, in economic theory we find that the allocation of factors of production between different uses is determined by the price mechanism. The price of factor *A* becomes higher in *X* than in *Y*. As a result, *A* moves from *Y* to *X* until the difference between the prices in *X* and *Y*, except in so far as it compensates for other differential advantages, disappears. Yet in the real world, we find that there are many areas where this does not apply. If a workman moves from department *Y* to department *X*, he does not go because of a change in relative prices, but because he is ordered to do so. Those who object to economic planning on the grounds that the problem is solved by price movements can be answered by pointing out that there is planning within our economic system which is quite different from the individual planning mentioned above and which is akin to what is normally called economic planning. The example given above is typical of a large sphere in our modern economic system. Of course, this fact has not been ignored by economists. Marshall introduces organization as a fourth factor of production; J. B. Clark gives the co-ordinating function to the entrepreneur; Professor Knight introduces managers who co-ordinate. As D. H. Robertson points out, we find "islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk."<sup>9</sup> But in view of the fact that it is usually argued that co-ordination will be done by the price mechanism, why is such organization necessary? Why are there these "islands of conscious power"? Outside the firm, price movements direct production, which is co-ordinated through a series of exchange transactions on the market. Within a firm, these market transactions are eliminated and in place of the complicated market structure with exchange transactions is substituted the entrepreneur-co-ordinator, who directs production.<sup>10</sup> It is clear that these are alternative methods of co-ordinating production. Yet, having regard to the fact that if production is regulated by price movements, production could be carried on without any organization at all, well might we ask, why is there any organization?

Of course, the degree to which the price mechanism is superseded varies greatly. In a department store, the allocation of the different sections to the various locations in the building may be done by the controlling authority or it may be the result of competitive price bidding for space. In the Lancashire cotton industry, a weaver can rent power and shop-room and can obtain looms and yarn on credit.<sup>11</sup>

This co-ordination of the various factors of production is, however, normally carried out without the intervention of the price mechanism. As is evident, the amount of "vertical" integration, involving as it does the supersession of the price mechanism, varies greatly from industry to industry and from firm to firm.

It can, I think, be assumed that the distinguishing mark of the firm is the supersession of the price mechanism. It is, of course, as Professor Robbins points out, "related to an outside network of relative prices and costs,"<sup>12</sup> but it is important to discover the exact nature of this relationship. This distinction between the allocation of resources in a firm and the allocation in the economic system has been very vividly described by Mr. Maurice Dobb when discussing Adam Smith's conception of the capitalist: "It began to be seen that there was something more important than the relations inside each factory or unit captained by an undertaker; there were the relations of the undertaker with the rest of the economic world outside his immediate sphere . . . the undertaker busies himself with the division of labour inside each firm and he plans and organises consciously," but "he is related to the much larger economic specialisation, of which he himself is merely one specialised unit. Here, he plays his part as a single cell in a larger organism, mainly unconscious of the wider rôle he fills."<sup>13</sup>

In view of the fact that while economists treat the price mechanism as a co-ordinating instrument, they also admit the co-ordinating function of the "entrepreneur," it is surely important to enquire why co-ordination is the work of the price mechanism in one case and of the entrepreneur in another. The purpose of this paper is to bridge what appears to be a gap in economic theory between the assumption (made for some purposes) that resources are allocated by means of the price mechanism and the assumption (made for other purposes) that this allocation is dependent on the entrepreneur-co-ordinator. We have to explain the basis on which, in practice, this choice between alternatives is effected.<sup>14</sup>

## II

Our task is to attempt to discover why a firm emerges at all in a specialized exchange economy. The price mechanism (considered purely from the side of the direction of resources) might be superseded if the relationship which replaced it was desired for its own sake. This would be the case, for example, if some people preferred to work under the direction of some other person. Such individuals would accept less in order to work under someone, and firms would arise naturally from this. But it would appear that this cannot be a very important reason, for it would rather seem that the opposite tendency is operating if one judges from the stress normally laid on the advantage of "being one's own master."<sup>15</sup> Of course, if the desire was not to be controlled but to control, to exercise power over others, then people might be willing to give up something in order to direct others; that is, they would be willing to pay others more than they could get under the price mechanism in order to be able to direct them. But this implies that those who direct pay in order to be able to do this and are not paid to direct, which is clearly not true in the

majority of which are pro one would ex firms are to l involved.

The main there is a cost production th prices are.<sup>18</sup> T of specialists ing a separate must also be changes, a tec eliminated. It are greatly rec make a series firm, as wouk the working o At this stage, enters that is c certain remun tions of an en should only st he can therefo

There are nism. It may t or service. Th period, instea will be avoide prefer to make ty of forecasti commodity or person purcha may well be a modity which service or con courses he wil provided is exp All that is sta commodity or to do is not sta direction of res buyer in this w is likely therefo be unsatisfact

majority of cases.<sup>16</sup> Firms might also exist if purchasers preferred commodities which are produced by firms to those not so produced; but even in spheres where one would expect such preferences (if they exist) to be of negligible importance, firms are to be found in the real world.<sup>17</sup> Therefore there must be other elements involved.

The main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism. The most obvious cost of "organizing" production through the price mechanism is that of discovering what the relevant prices are.<sup>18</sup> This cost may be reduced but it will not be eliminated by the emergence of specialists who will sell this information. The costs of negotiating and concluding a separate contract for each exchange transaction which takes place on a market must also be taken into account.<sup>19</sup> Again, in certain markets, e.g., produce exchanges, a technique is devised for minimizing these contract costs; but they are not eliminated. It is true that contracts are not eliminated when there is a firm but they are greatly reduced. A factor of production (or the owner thereof) does not have to make a series of contracts with the factors with whom he is co-operating within the firm, as would be necessary, of course, if this co-operation were as a direct result of the working of the price mechanism. For this series of contracts is substituted one. At this stage, it is important to note the character of the contract into which a factor enters that is employed within a firm. The contract is one whereby the factor, for a certain remuneration (which may be fixed or fluctuating), agrees to obey the directions of an entrepreneur *within certain limits*.<sup>20</sup> The essence of the contract is that it should only state the limits to the powers of the entrepreneur. Within these limits, he can therefore direct the other factors of production.

There are, however, other disadvantages — or costs — of using the price mechanism. It may be desired to make a long-term contract for the supply of some article or service. This may be due to the fact that if one contract is made for a longer period, instead of several shorter ones, then certain costs of making each contract will be avoided. Or, owing to the risk attitude of the people concerned, they may prefer to make a long rather than a short-term contract. Now, owing to the difficulty of forecasting, the longer the period of the contract is for the supply of the commodity or service, the less possible, and indeed, the less desirable it is for the person purchasing to specify what the other contracting party is expected to do. It may well be a matter of indifference to the person supplying the service or commodity which of several courses of action is taken, but not to the purchaser of that service or commodity. But the purchaser will not know which of these several courses he will want the supplier to take. Therefore, the service which is being provided is expressed in general terms, the exact details being left until a later date. All that is stated in the contract is the limits to what the persons supplying the commodity or service is expected to do. The details of what the supplier is expected to do is not stated in the contract but is decided later by the purchaser. When the direction of resources (within the limits of the contract) becomes dependent on the buyer in this way, that relationship which I term a "firm" may be obtained.<sup>21</sup> A firm is likely therefore to emerge in those cases where a very short-term contract would be unsatisfactory. It is obviously of more importance in the case of services —

labor—than it is in the case of the buying of commodities. In the case of commodities, the main items can be stated in advance and the details which will be decided later will be of minor significance.

We may sum up this section of the argument by saying that the operation of a market costs something and by forming an organization and allowing some authority (an “entrepreneur”) to direct the resources, certain marketing costs are saved. The entrepreneur has to carry out his function at less cost, taking into account the fact that he may get factors of production at a lower price than the market transactions which he supersedes, because it is always possible to revert to the open market if he fails to do this.

The question of uncertainty is one which is often considered to be very relevant to the study of the equilibrium of the firm. It seems improbable that a firm would emerge without the existence of uncertainty. But those, for instance, Professor Knight, who make the *mode of payment* the distinguishing mark of the firm—fixed incomes being guaranteed to some of those engaged in production by a person who takes the residual, and fluctuating, income—would appear to be introducing a point which is irrelevant to the problem we are considering. One entrepreneur may sell his services to another for a certain sum of money, while the payment to his employees may be mainly or wholly a share in profits.<sup>22</sup> The significant question would appear to be why the allocation of resources is not done directly by the price mechanism.

Another factor that should be noted is that exchange transactions on a market and the same transactions organized within a firm are often treated differently by Governments or other bodies with regulatory powers. If we consider the operation of a sales tax, it is clear that it is a tax on market transactions and not on the same transactions organized within the firm. Now since these are alternative methods of “organization”—by the price mechanism or by the entrepreneur—such a regulation would bring into existence firms which otherwise would have no *raison d'être*. It would furnish a reason for the emergence of a firm in a specialized exchange economy. Of course, to the extent that firms already exist, such a measure as a sales tax would merely tend to make them larger than they would otherwise be. Similarly, quota schemes, and methods of price control which imply that there is rationing, and which do not apply to firms producing such products for themselves, by allowing advantages to those who organize within the firm and not through the market, necessarily encourage the growth of firms. But it is difficult to believe that it is measures such as have been mentioned in this paragraph which have brought firms into existence. Such measures would, however, tend to have this result if they did not exist for other reasons.

These, then, are the reasons why organizations such as firms exist in a specialized exchange economy in which it is generally assumed that the distribution of resources is “organized” by the price mechanism. A firm, therefore, consists of the system of relationships which comes into existence when the direction of resources is dependent on an entrepreneur.

The approach which has just been sketched would appear to offer an advantage in that it is possible to give a scientific meaning to what is meant by saying that

a firm gets its  
(which could  
are organized  
zation of such  
study the force  
not organize  
sor Knight co

the relative  
theory, b  
sonality a  
But the q  
offers a p  
firm, whi  
decrease  
if even be

Professor Knight  
ly the determi  
developed abo

It was su  
existence of n  
apart from th  
organizing on  
are there any  
one big firm?

First, as  
function, that  
may rise.<sup>25</sup> Na  
transaction w  
transaction in  
neur. Secondl  
entrepreneur  
greatest, that  
point must be  
marketing cos  
transaction w  
or more of th  
small firm are  
the expansion  
factors mentio  
the economist

The poin  
expand until  
equal to the c

a firm gets larger or smaller. A firm becomes larger as additional transactions (which could be exchange transactions co-ordinated through the price mechanism) are organized by the entrepreneur and becomes smaller as he abandons the organization of such transactions. The question which arises is whether it is possible to study the forces which determine the size of the firm. Why does the entrepreneur not organize one less transaction or one more? It is interesting to note that Professor Knight considers that:

the relation between efficiency and size is one of the most serious problems of theory, being, in contrast with the relation for a plant, largely a matter of personality and historical accident rather than of intelligible general principles. But the question is peculiarly vital because the possibility of monopoly gain offers a powerful incentive to *continuous and unlimited* expansion of the firm, which force must be offset by some equally powerful one making for decreased efficiency (in the production of money income) with growth in size, if even boundary competition is to exist.<sup>23</sup>

Professor Knight would appear to consider that it is impossible to treat scientifically the determinants of the size of the firm. On the basis of the concept of the firm developed above, this task will now be attempted.

It was suggested that the introduction of the firm was due primarily to the existence of marketing costs. A pertinent question to ask would appear to be (quite apart from the monopoly considerations raised by Professor Knight), why, if by organizing one can eliminate certain costs and in fact reduce the cost of production, are there any market transactions at all?<sup>24</sup> Why is not all production carried on by one big firm? There would appear to be certain possible explanations.

First, as a firm gets larger, there may be decreasing returns to the entrepreneur function, that is, the costs of organizing additional transactions within the firm may rise.<sup>25</sup> Naturally, a point must be reached where the costs of organizing an extra transaction within the firm are equal to the costs of organizing in carrying out the transaction in the open market, or, to the costs of organizing by another entrepreneur. Secondly, it may be that as the transactions which are organized increase, the entrepreneur fails to place the factors of production in the uses where their value is greatest, that is, fails to make the best use of the factors of production. Again, a point must be reached where the loss through the waste of resources is equal to the marketing costs of the exchange transaction in the open market or to the loss if the transaction was organized by another entrepreneur. Finally, the supply price of one or more of the factors of production may rise, because the "other advantages" of a small firm are greater than those of a large firm.<sup>26</sup> Of course, the actual point where the expansion of the firm ceases might be determined by a combination of the factors mentioned above. The first two reasons given most probably correspond to the economists' phrase of "diminishing returns to management."<sup>27</sup>

The point has been made in the previous paragraph that a firm will tend to expand until the costs of organizing an extra transaction within the firm become equal to the costs of carrying out the same transaction by means of an exchange on

the open market or the costs of organizing in another firm. But if the firm stops its expansion at a point below the costs of marketing in the open market and at a point equal to the costs of organizing in another firm, in most cases (excluding the case of "combination"<sup>28</sup>), this will imply that there is a market transaction between these two procedures, each of whom could organize it at less than the actual marketing costs. How is the paradox to be resolved? If we consider an example the reason for this will become clear. Suppose *A* is buying a product from *B* and that both *A* and *B* could organize this marketing transaction at less than its present cost. *B*, we can assume, is not organizing one process or stage of production, but several. If *A* therefore wishes to avoid a market transaction, he will have to take over all the processes of production controlled by *B*. Unless *A* takes over all the processes of production, a market transaction will still remain, although it is a different product that is bought. But we have previously assumed that as each producer expands he becomes less efficient; the additional costs of organizing extra transactions increase. It is probable that *A*'s cost of organizing the transactions previously organized by *B* will be greater than *B*'s costs of doing the same thing. *A* therefore will take over the whole of *B*'s organization only if his cost of organizing *B*'s work is not greater than *B*'s cost by an amount equal to the costs of carrying out an exchange transaction on the open market. But once it becomes economical to have a market transaction, it also pays to divide production in such a way that the cost of organizing an extra transaction in each firm is the same.

Up to now it has been assumed that the exchange transactions which take place through the price mechanism are homogeneous. In fact, nothing could be more diverse than the actual transactions which take place in our modern world. This would seem to imply that the costs of carrying out exchange transactions through the price mechanism will vary considerably as will also the costs of organizing these transactions within the firm. It seems therefore possible that quite apart from the question of diminishing returns the costs of organizing certain transactions within the firm may be greater than the costs of carrying out the exchange transactions in the open market. This would necessarily imply that there were exchange transactions carried out through the price mechanism, but would it mean that there would have to be more than one firm? Clearly not, for all those areas in the economic system where the direction of resources was not dependent directly on the price mechanism could be organized within one firm. The factors which were discussed earlier would seem to be the important ones, though it is difficult to say whether "diminishing returns to management" or the rising supply price of factors is likely to be the more important.

Other things being equal, therefore, a firm will tend to be larger:

- a. the less the costs of organizing and the slower these costs rise with an increase in the transactions organized.
- b. the less likely the entrepreneur is to make mistakes and the smaller the increase in mistakes with an increase in the transactions organized.
- c. the greater the lowering (or the less the rise) in the supply price of factors of production to firms of larger size.

Apart from different sized mistakes will be organized, in the relevant would appear different place decrease as a function nearer to the firm.<sup>30</sup> Character of organizing improve ma-

It should be used to give. There is a cost or more entries involves the entrepreneur ways. The ordinary tec-

III

The problem entirely neglected given above preferred to

It is so the division adopted and increasing cost differentiation would economy the ment is an exists in the economic scization must inadmissible neur) should

The mc have been given *Uncertainty*

Apart from variations in the supply price of factors of production to firms of different sizes, it would appear that the costs of organizing and the losses through mistakes will increase with an increase in the spatial distribution of the transactions organized, in the dissimilarity of the transactions, and in the probability of changes in the relevant prices.<sup>29</sup> As more transactions are organized by an entrepreneur, it would appear that the transactions would tend to be either different in kind or in different places. This furnishes an additional reason why efficiency will tend to decrease as the firm gets larger. Inventions which tend to bring factors of production nearer together, by lessening spatial distribution, tend to increase the size of the firm.<sup>30</sup> Changes like the telephone and the telegraph which tend to reduce the cost of organizing spatially will tend to increase the size of the firm. All changes which improve managerial technique will tend to increase the size of the firm.<sup>31/32</sup>

It should be noted that the definition of a firm which was given above can be used to give more precise meanings to the terms "combination" and "integration."<sup>33</sup> There is a combination when transactions which were previously organized by two or more entrepreneurs become organized by one. This becomes integration when it involves the organization of transactions which were previously carried out between the entrepreneurs on a market. A firm can expand in either or both of these two ways. The whole of the "structure of competitive industry" becomes tractable by the ordinary technique of economic analysis.

### III

The problem which has been investigated in the previous section has not been entirely neglected by economists and it is now necessary to consider why the reasons given above for the emergence of a firm in a specialized exchange economy are to be preferred to the other explanations which have been offered.

It is sometimes said that the reason for the existence of a firm is to be found in the division of labor. This is the view of Professor Usher, a view which has been adopted and expanded by Mr. Maurice Dobb. The firm becomes "the result of an increasing complexity of the division of labour. . . . The growth of this economic differentiation creates the need for some integrating force without which differentiation would collapse into chaos; and it is as the integrating force in a differentiated economy that industrial forms are chiefly significant."<sup>34</sup> The answer to this argument is an obvious one. The "integrating force in a differentiated economy" already exists in the form of the price mechanism. It is perhaps the main achievement of economic science that it has shown that there is no reason to suppose that specialization must lead to chaos.<sup>35</sup> The reason given by Mr. Maurice Dobb is therefore inadmissible. What has to be explained is why one integrating force (the entrepreneur) should be substituted for another integrating force (the price mechanism).

The most interesting reasons (and probably the most widely accepted) which have been given to explain this fact are those to be found in Professor Knight's *Risk, Uncertainty and Profit*. His views will be examined in some detail.



Professor Knight starts with a system in which there is no uncertainty:

acting as individuals under absolute freedom but without collusion men are supposed to have organised economic life with the primary and secondary division of labour, the use of capital, etc., developed to the point familiar in present-day America. The principal fact which calls for the exercise of the imagination is the internal organisation of the productive groups or establishments. With uncertainty entirely absent, every individual being in possession of perfect knowledge of the situation, there would be no occasion for anything of the nature of responsible management or control of productive activity. Even marketing transactions in any realistic sense would not be found. The flow of raw materials and productive services to the consumer would be entirely automatic.<sup>36</sup>

Professor Knight says that we can imagine this adjustment as being "the result of a long process of experimentation worked out by trial-and-error methods alone," while it is not necessary "to imagine every worker doing exactly the right thing at the right time in a sort of 'pre-established harmony' with the work of others. There might be managers, superintendents, etc., for the purpose of co-ordinating the activities of individuals," though these managers would be performing a purely routine function, "without responsibility of any sort."<sup>37</sup>

Professor Knight then continues:

With the introduction of uncertainty — the fact of ignorance and the necessity of acting upon opinion rather than knowledge — into this Eden-like situation, its character is entirely changed. . . . With uncertainty present doing things, the actual execution of activity, becomes in a real sense a secondary part of life; the primary problem or function is deciding what to do and how to do it.<sup>38</sup>

This fact of uncertainty brings about the two most important characteristics of social organization.

In the first place, goods are produced for a market, on the basis of entirely impersonal prediction of wants, not for the satisfaction of the wants of the producers themselves. The producer takes the responsibility of forecasting the consumers' wants. In the second place, the work of forecasting and at the same time a large part of the technological direction and control of production are still further concentrated upon a very narrow class of the producers, and we meet with a new economic functionary, the entrepreneur. . . . When uncertainty is present and the task of deciding what to do and how to do it takes the ascendancy over that of execution the internal organisation of the productive groups is no longer a matter of indifference or a mechanical detail. Centralisation of this deciding and controlling function is imperative, a process of "cephalisation" is inevitable.<sup>39</sup>

The m

the sys  
insure  
in retu  
we kno  
to ano  
direct  
self un  
of this  
of indi  
tainty.

These  
uncertainty  
special clas  
guaranteed  
dence in on

Profes  
grounds. F  
better judg  
income fro  
advice or k  
imagine a s  
possible to  
part in pr  
merchant b  
illustrates t  
that certair  
Professor K  
cable or ve  
latter's acti  
rect. A larg  
guaranteed  
involve any  
been chang  
tion.<sup>42</sup> The  
place himse  
to the probl  
the case of  
considers tl  
routine fun  
of any sort  
that nowhe  
be supersed

The most fundamental change is:

the system under which the confident and venturesome assume the risk or insure the doubtful and timid by guaranteeing to the latter a specified income in return for an assignment of the actual results. . . . With human nature as we know it it would be impracticable or very unusual for one man to guarantee to another a definite result of the latter's actions without being given power to direct his work. And on the other hand the second party would not place himself under the direction of the first without such a guarantee. . . . The result of this manifold specialisation of function is the enterprise and wage system of industry. Its existence in the world is the direct result of the fact of uncertainty.<sup>40</sup>

These quotations give the essence of Professor Knight's theory. The fact of uncertainty means that people have to forecast future wants. Therefore, you get a special class springing up who direct the activities of others to whom they give guaranteed wages. It acts because good judgment is generally associated with confidence in one's judgment.<sup>41</sup>

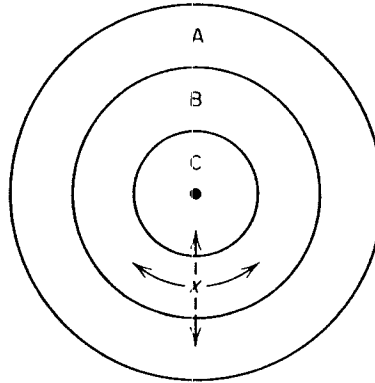
Professor Knight would appear to leave himself open to criticism on several grounds. First of all, as he himself points out, the fact that certain people have better judgment or better knowledge does not mean that they can only get an income from it by themselves actively taking part in production. They can sell advice or knowledge. Every business buys the services of a host of advisers. We can imagine a system where all advice or knowledge was bought as required. Again, it is possible to get a reward from better knowledge or judgment not by actively taking part in production but by making contracts with people who are producing. A merchant buying for future delivery represents an example of this. But this merely illustrates the point that it is quite possible to give a guaranteed reward providing that certain acts are performed without directing the performance of those acts. Professor Knight says that "with human nature as we know it, it would be impracticable or very unusual for one man to guarantee to another a definite result of the latter's actions without being given power to direct his work." This is surely incorrect. A large proportion of jobs are done to contract, that is, the contractor is guaranteed a certain sum providing he performs certain acts. But this does not involve any direction. It does mean, however, that the system of relative prices has been changed and that there will be a new arrangement of the factors of production.<sup>42</sup> The fact that Professor Knight mentions that the "second party would not place himself under the direction of the first without such a guarantee" is irrelevant to the problem we are considering. Finally, it seems important to notice that even in the case of an economic system where there is no uncertainty Professor Knight considers that there would be co-ordinators, though they would perform only a routine function. He immediately adds that they would be "without responsibility of any sort," which raises the question by whom are they paid and why? It seems that nowhere does Professor Knight give a reason why the price mechanism should be superseded.

IV

It would seem important to examine one further point and that is to consider the relevance of this discussion to the general question of the "cost-curve of the firm."

It has sometimes been assumed that a firm is limited in size under perfect competition if its cost curve slopes upward,<sup>43</sup> while under imperfect competition, it is limited in size because it will not pay to produce more than the output at which marginal cost is equal to marginal revenue.<sup>44</sup> But it is clear that a firm may produce more than one product and, therefore, there appears to be no prima facie reason why this upward slope of the cost curve in the case of perfect competition or the fact that marginal cost will not always be below marginal revenue in the case of imperfect competition should limit the size of the firm.<sup>45</sup> Mrs. Robinson<sup>46</sup> makes the simplifying assumption that only one product is being produced. But it is clearly important to investigate how the number of products produced by a firm is determined, while no theory which assumes that only one product is in fact produced can have very great practical significance.

It might be replied that under perfect competition, since everything that is produced can be sold at the prevailing price, then there is no need for any other product to be produced. But this argument ignores the fact that there may be a point where it is less costly to organize the exchange transactions of a new product than to organize further exchange transactions of the old product. This point can be illustrated in the following way. Imagine, following von Thunen, that there is a town, the consuming center, and that industries are located around this central point in rings. These conditions are illustrated in the following diagram in which *A*, *B*, and *C* represent different industries.



Imagine an entrepreneur who starts controlling exchange transactions from *x*. Now as he extends his activities in the same product (*B*), the cost of organizing increases until at some point it becomes equal to that of a dissimilar product which is nearer. As the firm expands, it will therefore from this point include more than one product (*A* and *C*). This treatment of the problem is obviously incomplete,<sup>47</sup> but it is

necessary  
give a list  
of perfect  
vious.

To d  
is, the co  
entrepreneur  
each firm  
Mr. Shov  
Mrs. Rob  
would se

V

Only one  
has been  
the quest  
ship nor  
The esser

(1) tl  
mast  
tract

perso  
ence  
servi  
the t  
and  
play  
the c  
work  
give

We thus s  
"employe  
oped abo

That  
ence  
but r

We can th  
mates clo.

necessary to show that merely proving that the cost curve turns upwards does not give a limitation to the size of the firm. So far we have only considered the case of perfect competition; the case of imperfect competition would appear to be obvious.

To determine the size of the firm, we have to consider the marketing costs (that is, the costs of using the price mechanism), and the costs of organizing the different entrepreneurs and then we can determine how many products will be produced by each firm and how much of each it will produce. It would, therefore, appear that Mr. Shove<sup>48</sup> in his article on "Imperfect Competition" was asking questions which Mrs. Robinson's cost curve apparatus cannot answer. The factors mentioned above would seem to be the relevant ones.

#### V

Only one task now remains; and that is, to see whether the concept of a firm which has been developed fits in with that existing in the real world. We can best approach the question of what constitutes a firm in practice by considering the legal relationship normally called that of "master and servant" or "employer and employee."<sup>49</sup> The essentials of this relationship have been given as follows:

(1) the servant must be under the duty of rendering personal services to the master or to others on behalf of the master, otherwise the contract is a contract for sale of goods or the like.

(2) The master must have the right to control the servant's work, either personally or by another servant or agent. It is this right of control or interference, of being entitled to tell the servant when to work (within the hours of service) and when not to work, and what work to do and how to do it (within the terms of such service) which is the dominant characteristic in this relation and marks off the servant from an independent contractor, or from one employed merely to give to his employer the fruits of his labour. In the latter case, the contractor or performer is not under the employer's control in doing the work or effecting the service; he has to shape and manage his work so as to give the result he has contracted to effect.<sup>50</sup>

We thus see that it is the fact of direction which is the essence of the legal concept of "employer and employee," just as it was in the economic concept which was developed above. It is interesting to note that Professor Batt says further:

That which distinguishes an agent from a servant is not the absence or presence of a fixed wage or the payment only of commission on business done, but rather the freedom with which an agent may carry out his employment.<sup>51</sup>

We can therefore conclude that the definition we have given is one which approximates closely to the firm as it is considered in the real world.

Our definition is, therefore, realistic. Is it manageable? This ought to be clear. When we are considering how large a firm will be the principle of marginalism works smoothly. The question always is, will it pay to bring an extra exchange transaction under the organizing authority? At the margin, the costs of organizing within the firm will be equal either to the costs of organizing in another firm or to the costs involved in leaving the transaction to be "organized" by the price mechanism. Business men will be constantly experimenting, controlling more or less, and in this way, equilibrium will be maintained. This gives the position of equilibrium for static analysis. But it is clear that the dynamic factors are also of considerable importance, and an investigation of the effect changes have on the cost of organizing within the firm and on marketing costs generally will enable one to explain why firms get larger and smaller. We thus have a theory of moving equilibrium. The above analysis would also appear to have clarified the relationship between initiative or enterprise and management. Initiative means forecasting and operates through the price mechanism by the making of new contracts. Management properly merely reacts to price changes, rearranging the factors of production under its control. That the business man normally combines both functions is an obvious result of the marketing costs which were discussed above. Finally, this analysis enables us to state more exactly what is meant by the "marginal product" of the entrepreneur. But an elaboration of this point would take us far from our comparatively simple task of definition and clarification.

## NOTES

1. Joan Robinson, *Economics Is a Serious Subject* (1932), 12.
2. See N. Kaldor, "The Equilibrium of the Firm," 44 *The Economic Journal* (1934) 60-76.
3. Op. cit., 6.
4. J. M. Keynes, *Essays in Biography* (1933), 223-24.
5. L. Robbins, *Nature and Significance of Economic Science* (1935), 63.
6. This description is quoted with approval by D. H. Robertson, *Control of Industry* (1923), 85, and by Professor Arnold Plant, "Trends in Business Administration," 12 *Economica* (1932) 45-62. It appears in *Allied Shipping Control*, pp. 16-17.
7. See F. A. Hayek, "The Trend of Economic Thinking," 13 *Economica* (1933) 121-37.
8. See F. A. Hayek, op. cit.
9. Op. cit., 85.
10. In the rest of this paper I shall use the term entrepreneur to refer to the person or persons who, in a competitive system, take the place of the price mechanism in the direction of resources.
11. *Survey of Textile Industries*, 26.
12. Op. cit., 71.
13. *Capitalist Enterprise and Social Progress* (1925), 20. Cf., also, Henderson, *Supply and Demand* (1932), 3-5.
14. It is easy to see when the State takes over the direction of an industry that, in planning it, it is doing something which was previously done by the price mechanism. What is

usually not r  
ments is also  
is therefore p  
economic pla  
tive system.  
[1936] 676-90  
imposed on  
method of or  
planning!

15. Cf.  
*nal* (1934) 84  
better paid o  
independent"

16. Noi  
earn less than

17. G. I  
*Journal* (193:  
he gives is alr

18. Acc  
librium," 17  
static theory  
true of the re

19. Thi  
capitalism. H  
waste of ener  
not develop t

20. It w  
would be vol  
(1933), 18, su

21. Of  
there is a firm  
whether there  
cussion of thi

22. The

23. *Ris*

ics Series of F

24. The  
"consumers"  
might be so l  
obtained was

25. Thi

homogeneous

26. For  
firms of vary

It is sometime  
increases beca  
heads of dep  
Macgregor, *I*

usually not realized is that any business man in organizing the relations between his departments is also doing something which could be organized through the price mechanism. There is therefore point in Mr. Durbin's answer to those who emphasize the problems involved in economic planning that the same problems have to be solved by business men in the competitive system. (See "Economic Calculus in a Planned Economy," 46 *The Economic Journal* [1936] 676-90.) The important difference between these two cases is that economic planning is imposed on industry while firms arise voluntarily because they represent a more efficient method of organizing production. In a competitive system, there is an "optimum" amount of planning!

15. Cf. Harry Dawes, "Labour Mobility in the Steel Industry," 44 *The Economic Journal* (1934) 84-94, who instances "the trek to retail shopkeeping and insurance work by the better paid of skilled men due to the desire (often the main aim in life of a worker) to be independent" (86).

16. None the less, this is not altogether fanciful. Some small shopkeepers are said to earn less than their assistants.

17. G. F. Shove, "The Imperfection of the Market: a Further Note," 44 *The Economic Journal* (1933) 113-24, n. 1, points out that such preferences may exist, although the example he gives is almost the reverse of the instance given in the text.

18. According to N. Kaldor, "A Classificatory Note of the Determinateness of Equilibrium," 1 *The Review of Economic Studies* (1934) 122-36, it is one of the assumptions of static theory that "All the relevant prices are known to all individuals." But this is clearly not true of the real world.

19. This influence was noted by Professor Usher when discussing the development of capitalism. He says: "The successive buying and selling of partly finished products were sheer waste of energy." (*Introduction to the Industrial History of England* (1920), 13.) But he does not develop the idea nor consider why it is that buying and selling operations still exist.

20. It would be possible for no limits to the powers of the entrepreneur to be fixed. This would be voluntary slavery. According to Professor Batt, *The Law of Master and Servant* (1933), 18, such a contract would be void and unenforceable.

21. Of course, it is not possible to draw a hard and fast line which determines whether there is a firm or not. There may be more or less direction. It is similar to the legal question of whether there is the relationship of master and servant or principal and agent. See the discussion of this problem below.

22. The views of Professor Knight are examined below in more detail.

23. *Risk, Uncertainty and Profit*, Preface to the Re-issue, London School of Economics Series of Reprints, No. 16 (1933).

24. There are certain marketing costs which could only be eliminated by the abolition of "consumers' choice" and these are the costs of retailing. It is conceivable that these costs might be so high that people would be willing to accept rations because the extra product obtained was worth the loss of their choice.

25. This argument assumes that exchange transactions on a market can be considered as homogeneous; which is clearly untrue in fact. This complication is taken into account below.

26. For a discussion of the variation of the supply price of factors and production to firms of varying size, see E. A. G. Robinson, *The Structure of Competitive Industry* (1932). It is sometimes said that the supply price of organizing ability increases as the size of the firm increases because men prefer to be the heads of small independent businesses rather than the heads of departments in a large business. See Jones, *The Trust Problem* (1921), 531, and Macgregor, *Industrial Combination* (1935), 63. This is a common argument of those who

advocate Rationalization. It is said that larger units would be more efficient, but owing to the individualistic spirit of the smaller entrepreneurs, they prefer to remain independent, apparently in spite of the higher income which their increased efficiency under Rationalization makes possible.

27. This discussion is, of course, brief and incomplete. For a more thorough discussion of this particular problem, see N. Kaldor, "The Equilibrium of the Firm," 44 *The Economic Journal* (1934) 60-76, and E. A. G. Robinson, "The Problem of Management and the Size of the Firm," 44 *The Economic Journal* (1934) 242-57.

28. A definition of this term is given below.

29. This aspect of the problem is emphasized by N. Kaldor, *op. cit.* Its importance in this connection had been previously noted by E. A. G. Robinson, *The Structure of Competitive Industry* (1932), 83-106. This assumes that an increase in the probability of price movements increases the costs of organizing within a firm more than it increases the cost of carrying out an exchange transaction on the market—which is probable.

30. This would appear to be the importance of the treatment of the technical unit by E. A. G. Robinson, *op. cit.*, 27-33. The larger the technical unit, the greater the concentration of factors and therefore the firm is likely to be larger.

31. It should be noted that most inventions will change both the costs of organizing and the costs of using the price mechanism. In such cases, whether the invention tends to make firms larger or smaller will depend on the relative effect on these two sets of costs. For instance, if the telephone reduces the costs of using the price mechanism more than it reduces the costs of organizing, then it will have the effect of reducing the size of the firm.

32. An illustration of these dynamic forces is furnished by Maurice Dobb, *Russian Economic Development* (1928), 68. "With the passing of bonded labour the factory, as an establishment where work was organised under the whip of the overseer, lost its *raison d'être* until this was restored to it with the introduction of power machinery after 1846." It seems important to realize that the passage from the domestic system to the factory system is not a mere historical accident, but is conditioned by economic forces. This is shown by the fact that it is possible to move from the factory system to the domestic system, as in the Russian example, as well as vice versa. It is the essence of serfdom that the price mechanism is not allowed to operate. Therefore, there has to be direction from some organizer. When, however, serfdom passed, the price mechanism was allowed to operate. It was not until machinery drew workers into one locality that it paid to supersede the price mechanism and the firm again emerged.

33. This is often called "vertical integration," combination being termed "lateral integration."

34. *Op. cit.*, 10. Professor Usher's views are to be found in his *Introduction to the Industrial History of England* (1920), 1-18.

35. Cf. J. B. Clark, *Distribution of Wealth* (1899), 19, who speaks of the theory of exchange as being the "theory of the organisation of industrial society."

36. *Risk, Uncertainty and Profit*, 267.

37. *Op. cit.*, 267-68.

38. *Op. cit.*, 268.

39. *Op. cit.*, 268-95.

40. *Op. cit.*, 269-70.

41. *Op. cit.*, 270.

42. This shows that it is possible to have a private enterprise system without the existence of firms. Though, in practice, the two functions of enterprise, which actually influences

the system  
casts, and  
normally c  
theory. Th

43. S  
the Firm.

44. M  
small firm

45. M  
He is foll  
*American*  
the firm is

46. E  
47. A

organizing  
48. C

115. In co  
charged by  
suburbs?"  
apparatus

49. T  
are not ide

over their l  
of the lega

50. B  
51. C

the system of relative prices by forecasting wants and acting in accordance with such forecasts, and management, which accepts the system of relative prices as being given, are normally carried out by the same persons, yet it seems important to keep them separate in theory. This point is further discussed below.

43. See Kaldor, *op. cit.*, and Robinson, *The Problem of Management and the Size of the Firm*.

44. Mr. Robinson calls this the Imperfect Competition solution for the survival of the small firm.

45. Mr. Robinson's conclusion, *op. cit.*, 249, n. 1, would appear to be definitely wrong. He is followed by Horace J. White, Jr., "Monopolistic and Perfect Competition," 26 *The American Economic Review* (1936) 645, n. 27. Mr. White states "It is obvious that the size of the firm is limited in conditions of monopolistic competition."

46. *Economics of Imperfect Competition* (1934).

47. As has been shown above, location is only one of the factors influencing the cost of organizing.

48. G. F. Shove, "The Imperfection of the Market," 43 *The Economic Journal* (1933), 115. In connection with an increase in demand in the suburbs and the effect on the price charged by suppliers, Mr. Shove asks ". . . why do not the old firms open branches in the suburbs?" If the argument in the text is correct, this is a question which Mrs. Robinson's apparatus cannot answer.

49. The legal concept of "employer and employee" and the economic concept of a firm are not identical, in that the firm may imply control over another person's property as well as over their labor. But the identity of these two concepts is sufficiently close for an examination of the legal concept to be of value in appraising the worth of the economic concept.

50. Batt, *The Law of Master and Servant*, 6.

51. *Op. cit.*, 7.



lexibili-  
Owner-  
priable  
nomics  
vertical  
nomic  
rnal of  
ournal  
et," 59  
ractual

## 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

© The Nobel Foundation 1991.

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,<sup>1</sup> 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. . . ."<sup>2</sup> 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

ev  
te  
di  
wi  
"t  
ar  
ar  
as  
in  
ar  
of  
ve  
la  
in  
tic  
ar  
in  
ha  
of  
ex  
lea  
of  
wo  
  
Ac  
the  
sta  
Ec  
the  
ma  
in  
Co  
19:  
He  
exp  
to  
Be  
we  
noi  
tak  
cor  
tio  
the  
Ca:  
yea  
Sch

evidence/theory ratio . . . is currently very low in this field.”<sup>3</sup> 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.<sup>4</sup> 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?<sup>5</sup>

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.<sup>6</sup> 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  
a stra

T  
the ro  
gate it  
to a g  
it doe  
the ac  
nomic  
wrong  
tion o  
the fir  
about  
import  
costs i  
tence  
pervas  
and or  
making  
that ex  
from s  
only c  
Not to  
of the  
much  
designe  
reduce  
advant

I  
used to  
evoluti  
that we  
but in v  
buy any  
culty, h  
buy sor  
sale bu  
Similar  
what he  
Exchan  
Clearly  
be very  
place. T  
transact  
the driv

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.<sup>7</sup> 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

pos  
cal  
to  
ste  
tra  
un]  
the  
pri  
sys  
ma  
rig  
inc  
dis  
in  
Sir  
(ar  
(at  
ch;  
no  
suc  
of  
tio  
St  
or  
tie  
W  
all  
W  
po  
all  
int  
est  
litt  
the  
inc  
be  
Eu  
gr  
an  
na  
ne  
th  
co  
to

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."<sup>8</sup> 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 encourage-

ment  
by th  
but s  
orga  
the r

the r  
uncc  
comj  
treat  
pros

not :  
ther  
he s:  
seir

NO

ganiz

Jour.

G. V  
34-4

ber,



ment 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.