'The Nature of the Firm' after sixty years

B.D. Elzas

Research Memorandum 1998-32

May 1998
‘The Nature of the Firm’ after sixty years

by B.D. Elzas

Vakgroep Algemene Economie

Onderzoekgroep Geschiedenis van het Economisch Denken

May 1998
"The Nature of the Firm" after sixty years
by B.D. Elzas'

Introduction

It is a well-known fact that the overwhelming majority of the new publications in any scientific discipline are no more than contributions within an already existing approach in the discipline concerned. Only a tiny fraction of all scientific books and articles may be considered - at least in retrospect - as attempts to examine an existing problem from another viewpoint than the traditional ones within the discipline. One or two of these come as a bombshell, almost immediately leading to heated debates in the leading scientific journals and growing into something like a new 'orthodoxy' within a few years. An example of this course of events in economics is Keynes's "General Theory".

However, this is a very exceptional course of events. Even if a new approach is eventually adopted in the discipline, the normal course of events is that its success fails to materialize for many years or even decades on end. For instance, Gossen's book of 1854 had no appreciable effect on the economic insights of his contemporaries. Only when the "Marginal Revolution" of the 1870's had already taken place, was Gossen's work rediscovered. When at last Jevons was made acquainted with it, he honestly declared that, as far as principles were concerned, he had only reinvented Gossen's wheel (Jevons (1879), 1970, Preface, p. 61).

The adventures of the new approach in Coase's article "The Nature of the Firm" (1937) were even more spectacular than those of Gossen and they were much more gratifying for its author in at least two respects. Notwithstanding the fact that this approach began to arouse attention only twenty-five years after its publication, Coase himself was able to enjoy it for more than three decades already. Thus, he was enabled to take an active part in the belated debates around it. Gossen was

*Erasmus University Rotterdam. This paper was funded by the Faculty of Economics and Econometrics of the Free University of Amsterdam as part of its research program "The Development of the New Institutional Economics".
years after his pioneering publication. Secondly, although in Coase's case (just as in Gossen's) his 'wheel' was more or less reinvented by others, it has become known in the profession at large under his name. As a consequence of his taking an active part in the later discussions Coase now is regarded not merely as a precursor, but as the father of a subsequent development in economic theory.

In this essay a number of questions concerning this pathbreaking publication by Coase are tentatively answered. Historical enquiries seldomly lead to definitive answers, because in principle there is always room left for different interpretations of 'the facts' and their combinations. In other words: history is no 'hard' science. So, for clarity's sake, let us state the most ambitious questions in the most straightforward and simple form:

1. How did Coase discover this new approach and what made him decide to take it?
2. Why did it last so long before this approach, which eventually was acknowledged by a majority in the profession to be very useful, was noticed at all?
3. What caused its ultimate breakthrough into the profession's interest?

1. The discovery

1.1 The message

In case of a lucky strike (be it in a theoretical or an applied field of work) after the discovery has become part of the generally accepted insights of the profession, it often is not so easy to understand why the discovery was not made earlier. The insight has then become so familiar to the profession that its members have difficulty in feeling the originality of the discovery any more. This also is the case with the discovery reported by Coase in "The Nature of the Firm" (Coase 1937). In order to let the reader undergo this sensation of naturalness, Coase's message will be reproduced briefly here:

According to Neoclassical price theory a market economy provides, automatically and without any control from
outside, for a perfect match between size and composition of demand on the one hand and the allocation of the factors of production on the other. In this theory it is assumed that the direction of resources is dependent directly on the price mechanism, that is: without deliberate central control. If this is a correct representation, why is it then, that firms, particles of centrally guided economy, exist, as "islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk." (Robertson, 1928, p. 85; cited by Coase, 1937, 1988, p. 35)?

Coase's answer to this question is that this is caused by the fact that economic agents cannot use the market costlessly: to use the market means incurring transaction costs. Consequently it is always a question of weighing up at the margin whether it will be cheaper to buy an extra unit of a resource on the market, or produce it 'in-house'. If there were no organization costs within the firm whatsoever, to be balanced against the costs of using the market, the cheapest way of allocating in the economy would be to have one big firm only between the markets for production factors on the one hand, and the markets for final products on the other. This big and only firm would then act as the buying party and the selling party respectively, of all transactions occurring in the economy. However, allocation cannot occur free of cost within the firm either, it always involves some organization costs. It follows that there will be more than one firm in any long-run equilibrium of the economy. Moreover, normally there will be markets for at least some intermediate products too.

Presumably no contemporary economist, upon reading the above summary of this general theory on the existence and the demarcation of firms, would be impressed by the originality of the message. It has taken root by now. At the time of publication it also did not make any impression, but for a different reason. At that time the profession was just falling under the spell of the Lausanne variant of the neoclassical approach. In that approach the firm is simply assumed to exist as a decision making unit. Therefore Coase's theory was by far too exotic and too original to be noticed.
In the beginning of the 'thirties Coase (born on 29 December 1910) was introduced to the neoclassical approach, but studying for a degree in commerce instead of economics, he was not immersed in it. It may be presumed that this helped him to persevere in regarding this approach as only an 'outsider' could have done, i.e. in a detached and critical way. But certainly this was not the only important factor, for many a person with a less inquisitive, critical and bright mind would have accomplished little with this independence. In observing this we meet with one of the least assessable causes of Coase's discovery: his exceptional intelligence. This, in combination with a strong urge to investigate real-world phenomena, at least twice led to highly original investigations and results. Of course, the combination of qualities, ascribed above to the young Coase cannot be established directly now, more than sixty years later. But it is possible to reconstruct it from his early publications and from his later career.

1.2 On Coase's personal background

In 1987, fifty years after "The Nature of the Firm" was first published (Economica, New Series, Vol. 4, pp. 386-405), a conference was organized to celebrate this anniversary. The organizers found themselves in the lucky circumstances that the author, at that time 76 years old, was able not only to be present at the conference, but also to give three lectures, and thus to contribute substantially to the proceedings. In his lectures he successively discussed the "origin", "meaning", and "influence" of his article (Coase 1993 [1988]). Especially for discussing its origin he was qualified as no one else could have been. In this section we shall use his expositions under these three headings.

In order to answer our first question in the Introduction ("how did Coase invent this new approach...?") it is necessary to know something about Coase's personal background and the state of economics at that time. Firstly, in terms of his higher education Coase proves to be no typical general economist at all. His higher education could be typified as that of a business economist with a remarkably broad range of subjects, a little bit of general economics being one of them. Moreover, this bit he learnt already at secondary school where
he followed a course of two years "...preparing for the external intermediate degree examinations of the University of London (the equivalent of the first year of university work)." (Origin, p.36). After making a false start for a science degree he soon switched to studying for a commerce degree.

"This move was undoubtedly made easier by the fact that I was at that time a socialist and the interest in social problems that this implies made the prospect of studying economics (a requirement for a commerce degree) attractive. During the next two years I studied economics, geography, French, English economic history and accounting..." (p. 36).

Having passed the examinations at eighteen, he went to the London School of Economics (LSE) to continue his studies for a commerce degree. In order to get an idea of his educational background it seems proper to quote extensively from Origin:

"In 1929-1930, I worked for Part I of the final examination, taking courses in British foreign trade, the principles of currency, the theory of production, industrial relations, the economic development of the overseas dominions, India and the tropical dependencies, commercial law, statistical method, the organization of transport and psychology. In 1930-1931, I decided to take group D (industry) in Part 2 of the Bachelor of Commerce, a group which was recommended for those wishing to enter the engineering and metal trades, distributing trades (in certain instances), and generally for those engaged in works and factory management. It was a choice of occupation for which I was singularly ill-suited (...) In studying for Part 2 in 1930-1931, I took courses in French, accounting, business administration, works and factory accounting and cost accounts, the raw materials of industry and trade, problems of modern industry, the financing of industry, industrial law, and industrial psychology." (p. 36)

Coase adds that the variety of subjects studied was very great and therefore little time (let alone systematic reading) could be devoted to each separately. Moreover, it shows that his studies during these two years did not contain a course in
economics at all. His education was in a great measure that of a practical business economist. Thus it seems that for his later excellence in price theory other factors must be held responsible. There was a little bit more of economics, and especially of price-theoretical insights in his education, however, than can be gleaned from the account given above. In 1930-1931, the lectures in business administration at the LSE for the first time were given by Arnold Plant. In *Origin* Coase explicitly states that it was Plant who provided him with his basic approach to economics (p. 37). Basically, this was that the price system, when combined with competition, has big advantages over alternative coordination devices.

After passing for his final examination, Coase could not yet be awarded the B. Corn. because three years' residence was required for a degree. So in principle he had to study a year longer at the LSE. However, he had been awarded a Cassel travelling scholarship for the academic year 1931-1932 and his study proposal, to be realized in the United States, was accepted as a sufficient substitute for a year's further residence. The accepted subject of study was the vertical and lateral integration in industry. Fifty-six years later he explains his choice for this subject as follows:

"My choice of subject (and it was my choice) undoubtedly resulted from the fact that Plant had referred to the different ways in which industries were organized. What stimulated my interest was that we seemed to lack any theory which would explain why those industries were organized in the way they *were.*" (*Origin*, p. 38).

Plant's influence shows not only in the choice of subject, but also in the fact that Coase was puzzled by two of Plant's publications of 1931, in which he opposed government coordination of industries, arguing that competition would provide all the coordination needed. Coase did not disagree with this, but had the feeling that something was missing in his argument. With regard to the economic planning discussion, kindled by the new planning experiment in the Soviet Union, Coase had a similar, and indeed related, feeling. Some maintained "...that to run the economy as one big factory was an *impossibility.*" (p. 39; Mises and Hayek belonged to these 'some'). "And yet there were factories in England and America.
How did one reconcile the impossibility of running Russia as one big factory with the existence of factories in the western world?" (ibid.).

Thus we see that Coase, only twenty years of age at the time, was a real outsider in economics from the start. He really believed what he had learned from Plant about coordination by competition, yet he did not believe it completely. Our proposition is, that the strong business economic accent in his education was one of the causal factors for this 'heresy'. Of course there were other factors, like intellect and character, contributing to his dissent. But the study of business economic subjects means that one has to put the management of the firm in the central place. Whether the subject is 'finance', 'cost accounting', 'profit assessment', or 'management of personnel', they all require a view of the firm from within. So, the firm cannot be dealt with as an indivisible unit (or 'black box'), as it is in neoclassical price theory.

A second remarkable fact appears from the story of his investigations in the United States. What he read there, were not

...reports of the Federal Trade Commission and books describing the organization of different industries (...)
I also read trade periodicals and used more unusual sources (for an economist), such as the yellow pages of the telephone directory, where I was fascinated to find so many specialist firms operating within what we thought of as a single industry as well as such interesting combinations of activities as those represented by coal and ice companies." (p. 39)

The remarkable thing is that he studied these kinds of practical sources. Again, this was a rather unusual thing to do for an economist, but not so for a business economist. In the course of his studies, moreover, he did not only visit universities, but many firm establishments as well. As he remarks in 1987 (Meaning, p. 52):

"...I tried to find the reason for the existence of the firm in factories and offices rather than in the writings
of economists (..) (M)y aim was to discover a "realistic assumption."

Perhaps this ambition should also be ascribed to his education as a business economist, or it could be more related to his character. Anyway, it shows a great measure of open-mindedness on Coase's part beside an intense interest in 'pedestrian' facts. He has maintained this interest, as is evident from many of his publications from the mid- 'thirties till the early 'nineties.

One more fact merits signalling here. He relates in Origin (p. 34/35) that he had completed the framework of the theory that was published in "The Nature of the Firm" already by the autumn of 1932, when he was only twenty-one. By that time he already gave lectures on it at the University of Dundee, where he had just been appointed assistant lecturer. This means that he had produced a novelty in economic theory as a mere youngster. From the later adventures of his novelty it will become clear that it was far ahead of its time.

Possibly less remarkable is the fact, that no less than five years went by before the publication of his discovery. A draft of the article was completed by the early summer of 1934. (p. 48). As far as Coase remembers, only "very" few changes were in fact made before the article was published in Economica in 1937. "(ibid.) Even more strongly: "The article published in 1937 was essentially the same as the draft of 1934." (ibid.) Why then the three years delay? His reply to this question lists a combination of factors (Meaning, p. 49-50), among them a reluctance "...to rush into print...", but also pressure of work in preparing his lectures in different courses at the universities of Dundee, Liverpool and the LSE successively.

1.3 More on Coase's characteristics as a scientist

Answering our first question from the Introduction could possibly be helped by searching for common traits in Coase's early publications. What kind or kinds of publications were they? By giving us a picture of Coase as a scientist in his early period, they could help us to understand better why he was able to make his discovery.
We restrict our inquiry to the period before the four-years gap in 
Coase's publication sequence caused by the Second World War. His first professional publication dates back to 1935 and the last one before the war gap appeared in 1940. Seven of the eleven publications he authored or co-authored in these six years seven will be typified here.  

His first publication was a purely theoretical one: "The Problem of Duopoly Reconsidered" (1935). The object was "... to examine the problem of duopoly with the aid of the cost curve apparatus developed by Professor Pigou, Mr. Harrod, Mrs. Robinson, and others." (p. 137). This article demonstrates that Coase, though originally not educated as a general economist, by 1935 already was able to take part in a contemporary discussion on price theory. Only two years before the books on imperfect competition by Mrs. Robinson and Chamberlin had given new inspiration to the discussion of imperfect competition models. Its subject matter also confirms something we learnt already from his own story in Origin, more than fifty years later: he was deeply interested in the relations between entrepreneurial decisions and market influence.  

In the same year, 1935, the first of a series of three articles by Coase and R.H. Fowler on price expectations held by producers appeared in print: "Bacon Production and the Pig Cycle in Great Britain". It was inspired by the current popularity of the idea that the 'cobweb theorem' provided a good explanation of the so-called 'pig cycle'. With the latter term a certain regularity of price fluctuations on pig markets or on pork product markets was meant. The cobweb theorem, then only recently developed (e.g. by Tinbergen, 1930), explained regularities in price fluctuations by assuming the combination of price-takership and 'static' price expectations on the supply side of the market. The stated object of the paper by Coase and Fowler is: "... to discover by an examination of the post-war experience of the British pig industry whether this argument is one which is really valid." (p. 144). Here they

\footnote{Of the other four one is his now famous 1937 article, one is only a rejoinder, and two, both co-authored by R.S. Edwards and R.F. Fowler, I didn't succeed to get hold of.}
did a thorough debunking job. They showed, firstly, that static price expectations are not in accordance with the facts on the British bacon market. Secondly, they argued that these facts could very well be in accordance with much more sensible behaviour on the part of the pig breeders than was attributed to them by the assumption of static price expectations.

The other two articles of this 'series' (Coase and Fowler 1937, 1940) were provoked by comments on the first one. Their first paper had been restricted to demonstrating that static price expectations did not make sense of the statistical facts. In the second they attempted an explanation of these facts, and in the third they even made an attempt at discovering what the producers' expectations on this market really could be. In the last paper there are two warnings (p. 286, p. 288), that it would be wise not to copy blindly the assumptions made by the authors. If one should try to discover producer's expectations on other markets, it would be better to study the particulars of the production structure in the industry concerned first.

These warnings point to two lasting traits in Coase's oeuvre: firstly, he time and again pleads with his fellow-economists to study the details of individual cases before generalizing, and not to indulge in 'armchair' theorizing. Secondly, he practises what he preaches.

In "Some Notes on Monopoly Price" (1937a) his skill in handling price-theoretical shows up again. Now the objective is not a purely technical one anymore: "This paper aims at making monopoly analysis more useful by introducing certain of the more important modifications which have to be made if Mrs. Robinson's theory is to be of use in increasing our understanding of the working of the actual economic system." (p. 17). This paper may be typified as an attempt to translate statical monopoly theory into a more realistic theory, that...
hopefully would throw light on the market behaviour of a monopolistic firm 'of flesh and blood'. One of the essential differences between such a firm and the one of static price theory is that the real firm normally does not know the demand curve for its product and, so, has to guess in its attempts to maximize profits. Here we encounter Coase's interest in imperfect information again, as in his more famous paper published in that year. Here this notion is not used to explain why firms exist, but to transform static monopoly theory into a tool for describing real-life firm behaviour on its product market.

In a series of twelve shorter pieces in The Accountant in 1938, Coase tried to explain to business men how they could profit from knowing as much as possible about their opportunity cost function before making their decisions (Coase 1973 [1938])4. This argument was wholly in line with the Austrian view of 'costs' as subjective opportunity costs. It was shared by a small group of economists at the LSE, under Arnold Plant (Coase, 1973, "Introductory Note", pp. 97-8). Together with R.S. Edwards and R.F. Fowler (the other members of this group) Coase published two papers in this period (Coase, Edwards, Fowler 1938, 1939).

In the 1939 volume of Economica Coase published a case study in economic history "Rowland Hill and the Penny Post", about the postal reform in Britain in 1840. This reform led to an early example of a 'uniform pricing system', a subject on which Coase published more after the war (1947).

To conclude this survey, two general traits in Coase's oeuvre can already be shown in his early work. Firstly, abstract 'theory of the firm' he only accepts as a first step towards explaining real-firm behaviour on markets. It means that he strives for greater realism in the description of the relevant circumstances for decision making by the entrepreneur. (See Coase and Fowler 1935, 1937 and 1940, plus Coase 1937 and 1937a).

---

4 'As I am only acquainted with the shortened one-piece 1973 version of this series, I count it as only one publication, out of a total of eleven publications by Coase up to 1940.'
A second and related trait is **Coase's sceptical** attitude towards general statements without sufficient empirical evidence. 'Armchair theorizing' is rejected if it is not followed by practical investigations inspired by the theory (see also Coase 1972, 1982, and 1991). Repeatedly both traits led him to investigations with debunking results: his work (with Fowler) on the pig cycle and his 1937 article on the nature of the firm are pre-war examples.*

2. Profound silence

During the first twenty years after its publication **Coase's** 1937 article was little noticed. In this period it did not meet with **any** reaction whatsoever in the professional journals. For many years there were no assenting or repudiating references to it. The article is included in the anthology *Readings in Price Theory* (1952), edited by Stigler and Boulding on behalf of the American Economic Association, and in the course of time more and more footnotes refer to it. However, the analysis supplied by it is not really applied or developed further. In 1972 Coase articulates this tersely: "... much cited and little used." (Coase (1972), 1988, p. 62; quoted in *Meaning*, p. 51).

By the time he makes this observation, however, a change in attitude by the profession had at last gotten into stride. This change was presumably partially caused by a later pathbreaking publication by Coase (1960), in which he applies the notion of transaction costs to another sphere. (See section 3).

Here we shall give a tentative explanation of the long period of silence about Coase's article of 1937, in accordance with the second question in our Introduction. For a start, we shall summarize some obvious factors contributing to the lack of attention by the profession. To begin with, 1937 belongs to

* 'Post-war examples are his 1960 paper on externalities, and his 'Lighthouse' paper (1974).

**The** first reference to it that I found in the journals dates from 1955, in the QJE. Coase himself remarks that some references to the article were made in the 1940's, and more in the 1950's.
the years of the Great Depression and obviously much attention was paid to that problem. This factor must have been very strong indeed in 1937, for Keynes' *General Theory* had just appeared. It immediately attracted enormous attention and caused intense debates, as is evidenced by the contents of British and American professional journals at the time. (See the Appendix, table 1).

However true this may be, it certainly does not mean that no attention whatsoever was left for other questions, such as those in price theory and what was later called "industrial organization". The *General Theory* did not prevent that in the next few years many articles on these micro-economic subjects were published too. It was only from 1941 onwards that the number of articles per volume of the leading periodicals began to diminish greatly in Great Britain (but much less so in the United States). (See the Appendix, table 2). By then also a shift towards subjects connected with the war effort occurred. Obviously, the war years cannot explain the pre-war part of the period of silence on Coase's article.

As far as these pre-war years are concerned, a development within price theory itself seems to be of some importance as an explaining factor. In the course of the 'thirties price theory underwent a rigorous formalization, under the influence of the approach of the Lausanne School. Until then price theory in Great Britain had been under the spell of Marshall. This meant that theoretical questions concerning the functioning of markets and the role of firms were approached by the Marshallian partial-equilibrium method. By this method it was still possible to get glimpses of entrepreneurs of flesh and blood, be it in a stylized way (as is necessarily the case in all theoretical approaches).' In the course of the 'thirties the Walrasian method strongly gained in influence, as is witnessed by, e.g., the following publications: Hicks & Allen (1934), Lerner (1934), Lange (1936/1937), Hicks (1939),

\[\text{On the Continent price theory had been more under the influence of the Austrian School, but in this rival approach the entrepreneur was sketched more or less as an active actor too, albeit in a different way from Marshall's.}\]
and, from a methodological point of view, Robbins (1938). The more formalized and tautologized price theory became, the further the notion of a 'firm' it employed moved away from the real-life entity whose management has to make complicated policy decisions.

With the help of a number of simplifying assumptions the tasks of management were stylized into the well-known marginal cost rule. Thus, the behaviour of the neoclassical 'entrepreneur' (if we can still use that expression) was reduced to that of an automaton. The decision making process was completely stylized away and the firm was reduced to a 'black box'.

This reduction was no drawback in view of the aim, on behalf of which the neoclassical theory was reformed in this way. After all, the Walrasian aim was the analysis of the co-ordinating abilities of the market system as a whole, in the case of perfectly decentralized decisionmaking by 'economic agents'.

Seen in the light of this aim only, all unnecessary complications could be left out for simplicity's sake. By assuming 'perfect information' for all agents and 'profit maximization' in a timeless setting as the only objective of all firms, the entrepreneur is reduced to a pure calculator, a truly "one-dimensional man". The other agents (consumers and suppliers of production factors) are assumed to have the same one-dimensional objective as the real and final one of the entrepreneur: 'utility maximization'.

---

*Whereas the first edition of Robbins's famous methodological Essay (1932) still had a strong Austrian flavour, the 'classical' second edition (1935) had much less. Apparently his insights had changed in the direction of formalized theory. See for instance Demsetz 1982.

"Profit maximization in fact is a derived objective. It would imply utility maximization for the entrepreneur in this theory, because profit is all the firm can offer the entrepreneur in terms of his utility, on-the-job consumption being assumed away by the reduction of the firm to a black box (an indivisible particle). See Demsetz 1995, p. 11.

In the case of the supplier of a production factor this means, that he maximizes the difference between the utility he can 'buy' with his total money remuneration on the one hand, and the total disutility of losing his command over the sold
The argument of Coase (1937) goes in a completely different direction. Instead of reducing Marshall's firm to a Walrasian black box, the notion of transaction costs asks us to taking a closer look at the internal organization of the firm. Thus it is no wonder that the article did not attract the attention it merited, at least from 'neoclassical circles'. To be tractable, the analysis of the market system as a coordination device required 'bloodless' firms. As far as the realism of this approach is concerned, a great part of the profession continued to believe that, notwithstanding this increasing formalization, its implications captured some essential characteristics of the market system in the real world. The latter did not satisfy all the assumptions made, but, as Robbins argued in both editions of his influential Essay (1932, 1935), this did not matter as long as the fundamental assumptions accorded with reality. He argued that this was indeed the case. Thus belief that market-directed allocation had a tendency to come into accordance with consumption preferences in the economy (the latter weighted by each consumer's purchasing power), remained strong in the 'thirties and 'forties. In this respect there was no fundamental change from the four preceding decades.12

As a stonesober outsider in the midst of this formalization movement, Coase asked to what purpose firms were still presumed to exist in this neoclassical world? If "(t)he normal economic system works itself" (Sir Arthur Salter; quoted in Coase (1937), 1988, p. 34), these particles of centrally commanded economy have no function whatsoever. The continued existence of the firm as a species in a market system with competition suggests that it does perform a function, that somehow this function is complementary to that performed by the market system itself. Within neoclassical economics this function remains in the dark. Though the firm, seen as an organization controlled by the entrepreneur, is assumed to exist, coordination in the economy is assumed to be performed

units of the production factor on the other.

12For instance, the much less formalized Austrian approach to the market system had also been accepted by its adherents on the basis of this heroic assumption, though normally the assumption itself was kept implicit. An exception to this lighthearted attitude will be mentioned shortly.
entirely without it. From this point of view the firm seems to be superfluous.

It is along this line of thinking that Coase makes his discovery: firms can have a reason for being only if they cause a decrease in costs, as compared to market coordination of actions by individuals only. In neoclassical theory such a decrease is impossible, because all market exchange is performed costlessly, in consequence of the assumption of perfect information. In reality this clearly is not true. Coase goes on to describe some types of transaction cost:

"The most obvious cost of "organizing" production through the price mechanism is that of discovering what the relevant prices are. 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. 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 a direct result of the working of the price mechanism. For this series of contracts is substituted one." (Coase (1937), pp. 38-9).

Contract costs may also be reduced by the substitution of one long-term contract for a series of short-term contracts (p. 39). Or, people may prefer a long-term contract on risk grounds (ibid.).

"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. A firm is likely, therefore, to emerge in those cases where a very short-term contract would be unsatisfactory." (p. 40).
In order to be able to incorporate all these types of transaction costs one must drop the assumption of perfect information.

Evidently it was asking too much from the refiners of neoclassical theory to appreciate this comment on their work. This becomes even more understandable if two more things are taken into account. Firstly, dropping the assumption of perfect information at that time would hopelessly have complicated the formalization work-in-progress of neoclassical theory. Secondly (and in our view more importantly) Coase's aim differed from the neoclassical endeavour of analysing the principles of market coordination. At the time his ultimate objective at the time was an explanation of vertical and lateral integration. For this reason his interest lay in the firm as an organization. There is no reason why this matter would have interested the neoclassical formalizers, as theirs was a radically different objective.

Related to this formalization trend, another factor of importance for the neglect of Coase's 1937 article is to be detected by returning to Marshall's price theory which was constructed from a partial-equilibrium approach. This means that not the economy, but the industry is scrutinized. This type of analysis demands that the industry is to a certain degree isolated from the rest of the economy, which Marshall did by applying the ceteris paribus assumption. As often happens with a well thought-out theory by a genius, Marshall's followers ran away with his subtle system. In some respects they stylized the construction further, thereby simplifying it. One of the simplifications was to restrict their attention to the case of perfect competition. (Andriessen, 1965, p. 41). It is well-known that in this market model an industry can only reach a long-run equilibrium situation if the average total cost curve of all the firms has a U-form. In a general-equilibrium approach this shape of the curve may be assumed without running into difficulties. However, Marshall's partial-equilibrium approach runs into the difficulty that rising long-run average cost must be caused by rising prices of production factors. The impact of the latter is not re-

---

17The following exposition is based on Andriessen, third impression, 1965.
stricted to the industry concerned, but is also felt by all other industries. But then the ceteris paribus assumption becomes implausible and the partial-equilibrium approach runs aground.

There are two conceivable alternatives to the U-form of the average cost curve: it could be continuously descending or horizontal. The first of these alternatives cannot be reconciled with the notion of an industry equilibrium under perfect competition. Thus only the other alternative remains: constant average total cost as a function of firm output. This alternative is not very attractive either, because a priori there are no reasons why average total cost should as a rule be constant. From this Sraffa (1926) concluded that there were only two ways out of this difficulty: either it has to be accepted that monopoly is a more usual outcome than perfect competition (in case average total cost is a decreasing function of output), or the partial-equilibrium approach must be given up.

After this conclusion was made two camps arose amongst the adherents of the partial-equilibrium approach. The first comprised those who wanted to maintain Marshall's model. Following him, they advanced the limitation of the life span of the individual firm as an argument against monopolization of the industry as the inevitable static outcome of the dynamic competition process. The other camp tried to keep the monopolization of the industry out of their theory with the aid of assuming special conditions on the demand side of the firm. This attempt to evade Sraffa's conclusion culminated in 1933 in the publication of two influential books on "monopolistic competition" by Chamberlin and Joan Robinson. Under their influence the further development of theoretical industry analysis was directed towards oligopoly. Both authors maintained the assumption of perfect information, but dropped the assumption of homogenous products. As was to be expected, the admissibility of the assumption of heterogeneity took a central place in the debate brought about by this new theoretical model, whereas 'perfect information' stayed out of the discussion completely. We conclude that here is another reason for the absolute neglect of Coase's approach. Not only was it not in the mainstream, but it went directly against its current.
In view of all the factors mentioned it seems quite understandable that, in view of general theoretical developments, time really was not ripe yet for a warm welcome or a sharp repudiation of Coase's innovation. More telling in particular is the attitude of two rising stars in economics at the time, because they were members of the same LSE staff as Coase: Robbins and Hayek. Coase relates that neither of them ever commented positively or negatively on his publication "... although my relations with both of them were quite cordial . . ." (Meaning, p. 51). With respect to Robbins, Coase concludes, that attention from him for an article entitled "The Nature of the Firm" was not to be expected, because Robbins "...was devoted to high theory...", i.e. the formalization of general equilibrium theory. In this regard it is remarkable that several decades later Robbins still did not pay any attention to Coase's innovation (Robbins 1971), although he normally had an open mind towards innovations.

Hayek's silence, not only at the time, but for many decades on end, seems even more remarkable. Coase does not go into that in his retrospective, but it certainly is remarkable in view of the fact that Hayek himself had touched upon the assumption of 'perfect knowledge' only the year before. In a speech of 1936 he pointed out a number of difficulties one has to face if one wants to apply the very stylized general equilibrium theory to the market system of the real world. The speech was published in 1937 in the same volume of Economica as Coase's article - under the title "Economics and Knowledge". Hayek's argument focused on the differences between the roles played by 'knowledge' in 'high theory' on the one hand, and in the real world on the other. He tried to establish the minimal content of 'perfect knowledge' needed for enabling a model of the market system to reach general equilibrium.

On the face of it this subject seems to be very strongly related to the notion of imperfect information, on which Coase constructed his transaction cost explanation of the firm. Moreover, Hayek certainly must have taken notice of the contents of Economica on a regular basis, because his own speech was published in this 'natural' outlet for LSE authors, as many more of his articles in the 'thirties and the first half of the 'forties were. So it seems strange that he did
not discuss Coase's contribution with his young colleague at all. However, when both publications are contemplated a little longer, important differences may be discerned between the two viewpoints. Hayek's problem is the robustness of the connection between the knowledge possessed by the agents in the market economy and a tendency to general equilibrium. Coase tries to reconcile the notions of coordination by the market economy and the firm. Besides, the two approaches are differently directed as regards the notion of a general equilibrium. In Hayek's approach it is viewed to be the situation only theoretically and ultimately to be reached by way of the market process. That starts from a disequilibrium situation and functions in 'real' time: it is a dynamic process, possibly a 'converging' one. In Coase's approach, however, the notion of general equilibrium, including the usual assumption of perfect knowledge, is accepted as a starting point, be it only in order to demonstrate that this leaves no room for the firm as a viable organization.

Possibly an explanation for Hayek's complete neglect of Coase's 1937 article must be searched for in this direction: he may not have recognized it as a discussion related to his own inquiry. This notwithstanding the fact that expressions of the idea of imperfect knowledge are to be found at least five times in the text plus notes of the article. Besides, an argument adopting the idea that the market system does not work costlessly could hardly have been experienced by Hayek as helpful to his cause in his current debate on socialist calculation with Lange et al.

In no later part of Hayek's oeuvre is anything done with Coase's insight, and only one reference to Coase 1937 was found 14. This fact is remarkable, because Hayek in many of his publications since 1937 devoted his mental powers repeatedly to speculating on the consequences of the absence of the 'perfection' of information for the market system 15. As far as could be verified, Coase on his part did not do anything with Hayek's approach either. This mutual ignoring can

14Hayek, 1988, p. 36.
15See: Hayek 1940; 1944, Ch. 4; 1945; 1946; 1962; 1968; 1976, notably Ch. 9 and 10; 1979, notably Ch. 15.
possibly be explained by the following combination of circumstances:

Firstly, Coase emphasises the cost of using the market mechanism. Hayek, conversely, draws attention to its 'yield': by using the market system the participants of the economic process are enabled to make use of specific knowledge which is present in the economy, but dispersed, and part of which is even untransferable to people not on the same spot. Agents are enabled to take advantage of it without having this specific knowledge themselves. Thereby they enhance economic efficiency in the economy without necessarily intending so.

Secondly, notwithstanding the assumption of transaction cost, Coase continues to make use of the method of comparative statics. For instance, in the final section of his 1937 article he concludes that in a market economy in general equilibrium firms must exist with a view to economic efficiency. From here he infers the consequences of an exogenous change in the level of transaction costs for the size and number of firms in the new equilibrium. In contrast, Hayek refuses to take general equilibrium as a useful guideline for studying the consequences of exogenous changes. So, it is not unthinkable that - if he ever read Coase's final section - he was deterred by it at once (and for decades afterwards) from taking an interest in Coase's approach. But on this we can of course only speculate.

3. The end of silence

After 1960 interest in Coase's approach increased rapidly and substantially. To this momentum he contributed himself with his article The Problem of Social Cost (Coase 1960), which later became at least as famous as his 1937 article. The neglect of transaction cost in traditional price theory is again taken to task in this new paper. In 1937 Coase had demonstrated that in traditional price theory firms lead a completely fictitious 'life', as a consequence of the assumption of costless markets. In 1960 he demonstrated that this assumption makes "externalities" to a complete fiction too. Pigou's expositions concerning government measures to correct
for externalities (Pigou 1912, 1920)\(^{16}\) are demonstrated to be floating in the air. Moreover, even the law is shown to be completely superfluous in a world without transaction costs.

One event in the history leading to this new publication in an 'old' approach is interesting enough to relate here, because it makes the rapid gaining of influence of Coase's approach after 1960 more comprehensible. In the Journal of Law & Economics of 1959 an article by Coase was published concerning the allocation of radio and television frequencies (Coase 1959). In it he gave his fundamental critique of neoclassical price theory in a shortened form. The Journal was established in 1958, under auspices of the University of Chicago Law School. Coase came into contact with its editor Aaron Director, who was professor of economics here. As related in Kitch (ed., 1983), p. 220, Director informed Coase that one of the referees wanted to see the passage containing the critique skipped, because he held it for erroneous\(^{17}\). In his autobiography Stigler (1985) describes the amazement Coase's critique aroused in him and his colleagues at the University of Chicago:

"... Chicago economists could not understand how so fine an economist as Coase could make so obvious a mistake." (p. 75).

Coase persisted, and, when he was invited by Stigler to come to Chicago and "... give a paper or present something at the Industrial Organization workshop" (ibid.), he accepted on condition that he could also on some occasion present his views on this other question (Kitch 1983, p.220). "Some twenty economists from the University of Chicago and Ronald Coase assembled one evening at the home of Aaron Director." (Stigler 1985, p. 75). Among them were Stigler, Director, Milton Friedman, Reuben Kessel, John McGee, Lloyd Mintz, Arnold Harberger, Gregg Louis, and Martin Bailey (Kitch 1983, p. 221). In his talk Coase succeeded in convincing his audience that externalities would not exist in a world without transaction costs. From this follows compellingly: "In a regime of zero

\(^{16}\)See Coase 1960, p. 28, note 35.

\(^{17}\)As related in Kitch (ed.), 1983, p. 220.
transaction costs, lawyers would perish." (Stigler 1985, p. 76).

"But Ronald asked us also to believe a second proposition about this world without transaction costs: Whatever the assignment of legal liability for damages, or whatever the assignment of legal rights of ownership, the assignments would have no effect upon the way economic resources would be used! We strongly objected to this heresy. Milton Friedman did much of the talking, as usual. He also did much of the thinking, as usual. In the course of two hours of argument the vote went from twenty against and one for Coase to twenty-one for Coase. What an exhilarating event!" (ibid.)

From this eyewitness report two things emerge which are important for answering our third question ("What caused its breakthrough into the profession's interest at last?"). In the first place, at that time first-class minds even were still not ripe for accepting Coase's approach without help. Secondly, these influential Chicago economists (who were no Walrasians in their orientation) were turned around in one evening's time, somewhere in the first half of 1960. This helps to explain why Coase's approach gained influence so quickly after 1960. At the invitation of Director, Coase wrote out his argument for the 1960 issue of the Journal of Law & Economics, which appeared sometime before May 1961 (Kitch, p. 221). This article gained rather soon a reputation. This is shown by the fact that the Coase Theorem" (the name Stigler invented for the bogus existence of the law in a neoclassical world without transaction costs) had already made its way into intermediate textbooks on price theory by the mid-1970's.

As the method of Coase's (1960) is the same as that of Coase (1937) i.e., the assumption of positive transaction costs and an investigation of its consequences compared with those of neoclassical theory. The following hypothesis seems to be warranted: The relatively quick general acceptance of the

\[\text{For instance: Hirshleifer's (1976). Stigler, of course, was very early; he incorporated it in the third edition of his textbook "Theory of Price", which was published in 1966, already.}\]
fundamental insight offered by the later article helped in clearing the way for giving attention to the fundamental insight offered by the old article. Coase shares this hypothesis (Influence, p. 62).

Soon after 1960 some new schools within general economics gained momentum. Authors like Stigler, Alchian, Demsetz, Cheung, Furubotn, and Pejovich published on various topics related to "property rights" and saw their papers accepted, not only by the 'maverick' Journal of Law and Economics, but by more traditional general economics journals as well. The emerging school of "law and economics", which originated from the University of Chicago Law School, gained a foothold in the general economics profession, as witnessed by, e.g., Manne 1965 in JPE, Calabresi 1970, Demsetz 1974, Becker 1968 in JPB.

Adjacent to the property rights approach the oeuvre of Douglass C. North began to expand. It can be characterized as an application of the property rights view to the development of types of economic organization of societies in history.

Last but not least, a modern theory of the firm begins to be developed. This development is to be seen as a direct (be it belated) descendant of Coase's old article. Of the many names that could be mentioned in this connection only a few follow here: Alchian and Demsetz 1972, Jensen and Meckling 1976 ("Agency Theory"), Chandler 1977, Nelson and Winter 1982 and practically the whole oeuvre of O.E. Williamson, spanning some

---


20The Property Rights approach and Law and Economics can only be distinguished from a historical point of view. Alchian, for instance, who was one of the originators of the first-mentioned approach, never was a member of the Chicago Law School, or of the University of Chicago at all, and from this point of view cannot be described as a 'Chicago economist'. Since the early 'sixties, however the two historically different approaches soon converged into one. See also Posner, 1993, p. 82, note 16.
Of course, the impact of Coase's 1960 article was not the only catalytic agent in the rapid development of all these 'Coasian' approaches to economic problems. Another important factor was the growing disappointment in the 'structure-conduct-performance' approach to "Industrial Organization". In these years it became ever more clear that this approach failed in at least two respects: it could not be integrated properly into Neoclassical price theory, and it was not able to produce an independent logical explanation for the interdependencies between technical and economic developments in an individual industry on the one hand, and market power distribution over individual firms in that industry on the other.**

4. Concluding remarks

In retrospect, we may typify the new approach taken by Coase in 1937 as an extraordinary stimulating and fruitful one, be it only after a remarkably long gestation period. Our answers to the questions posed in the Introduction above may be summarized thus:

1. Coase's discovery is to be ascribed to an exceptional combination of wits, sober scepticism, intellectual independence from current mainstream economics and a strong urge to learn the basic interactions between firm organization and market conditions in the real world.

2. The profession at large did not pick up Coase's approach until decades later, mainly because at the time its interests were developing in radically different directions.

3. The breakthrough came only after, and partly as a result of, Coase's 1960 publication. Growing disappointment with the structure-conduct-performance approach in 'Industrial Organization' also played a role in this.

---

**For a methodological critique of Williamson's choice of assumptions see Elzas, "Williamson's 'New Institutional Economics': a case for using Ockham's Razor" (forthcoming).

---

For a succinct survey of the difficulties in this respect see Demsetz (1974) 1989, Ch. 7.
Appendix

Table 1

<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>US</th>
<th>UK</th>
<th>US</th>
<th>UK</th>
<th>US</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1931</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>1932</td>
<td>8</td>
<td>4</td>
<td>12</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>1933</td>
<td>9</td>
<td>3</td>
<td>13</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>1934</td>
<td>6</td>
<td>7</td>
<td>15</td>
<td>4</td>
<td>4</td>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td>1935</td>
<td>6</td>
<td>3</td>
<td>14</td>
<td>3</td>
<td>4</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>1936</td>
<td>4</td>
<td>7</td>
<td>19</td>
<td>4</td>
<td>7</td>
<td>15</td>
<td>26</td>
</tr>
<tr>
<td>1937</td>
<td>10</td>
<td>6</td>
<td>19</td>
<td>3</td>
<td>4</td>
<td>15</td>
<td>26</td>
</tr>
<tr>
<td>1938</td>
<td>7</td>
<td>6</td>
<td>21</td>
<td>3</td>
<td>7</td>
<td>15</td>
<td>26</td>
</tr>
<tr>
<td>1939</td>
<td>13</td>
<td>4</td>
<td>24</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>33</td>
</tr>
<tr>
<td>1940</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>1941</td>
<td>8</td>
<td>6</td>
<td>16</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>25</td>
</tr>
</tbody>
</table>

'Subjects: business cycle theory, capital theory, theory of interest, growth theory.


'First volume: October 1933–June 1934.

Volumes run from November to August.

Table 2

<table>
<thead>
<tr>
<th>Period</th>
<th>UK</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1931–1941</td>
<td>22.5</td>
<td>30.1</td>
</tr>
<tr>
<td>1942–1946</td>
<td>15.5</td>
<td>26.4</td>
</tr>
<tr>
<td>1947–1960</td>
<td>23.9</td>
<td>30.7</td>
</tr>
</tbody>
</table>

'See Table 1, notes 2 and 3, for the names of the journals.
References


Coase, Ronald H. 1939, "Rowland Hill and the Penny Post", Economica, N.S., 6, 423-35.


Coase, R.H., R.S. Edwards and R.F. Fowler 1939, The Iron and Steel Industry 1926-1935: An Investigation Based on the
Accounts of Public Companies, Special Memorandum No. 49 of the London and Cambridge Economic Service.


Routledge & Kegan Paul.


Hicks, J.R. 1939, Value and Capital, Oxford University Press.


Theoretical Economics, 149, 73-87.


Robinson, Joan 1933, The Economics of Imperfect Competition, Macmillan.


