

1 Elements of the theory of the firm

This chapter reviews several basic issues in the theory of the firm, and indicates how they are addressed in subsequent chapters. Section 1.1 provides a critical review of thought on the *raison d'être* of the firm as an organization. An organization is a coalition of people characterized by diversity of (and most likely, conflict among) their interests and by coordination of their activities. Among the works presented are Coase's (1937) seminal view on the nature of the firm, Arrow's (1969) concept of transaction cost, Alchian and Demsetz's (1972) monitoring problem, and Williamson's (1975, 1985) various more specific types of transaction cost. It is concluded that at least two issues should be addressed in order to analyze the firm's activities: modelling the effective incentives of each economic agent, and postulating the cooperative behavioral principle. The remark on pp. 11–12 briefly compares the cooperative approach of this monograph (analysis of the situation in which economic agents interact based on the cooperative behavioral principle) with the principal-agent approach. Section 1.2 reviews empirical observations of the separation of ownership (stockholding) and control (management) of a firm in the modern capitalistic economy. The pioneering work by Berle and Means (1932), and the more recent observations by Larner (1966) are presented. The managerial utility hypotheses by Baumol (1959), by Marris (1964) and by Williamson (1964) are briefly discussed. The need for developing a theory which gets to the heart of the basic theory of the firm, in particular the thought on the *raison d'être* of the firm, and at the same time can describe the separation of ownership and control is emphasized. Section 1.3 first points out that there are many modes of socialistic economies that have actually prevailed in this century. The market-cooperative socialism is one of them. It has attracted much attention from economic theorists, due to the initial work by Ward (1958). Institutional studies by Jan Vanek (1972) and by Prasnikar and Svejnar (1988) are presented. Their findings are fundamentally important in developing a theory of market-cooperative socialism. It is concluded that

2 Cooperative nature of the firm

the essential difference between the manager-controlled firm in the capitalistic economy and the firm in the idealized market-cooperative socialism is that while the former can take advantage of financial markets, in particular the stock market, the latter has no access to such markets. Section 1.4 shows that the cooperative approach resolves many difficulties inherent in the classical definitions of the entrepreneur. Section 1.5 discusses two further issues related to Section 1.1: the internal labor “market” of Kerr (1954) and Doeringer and Piore (1971), and the Keynesian unemployment problem. This chapter is based on Ichiishi (1985), itself a revised version of a part of the paper, Ichiishi, “Management Versus Ownership,” GSIA Working Paper No. 46–76–77 (revised: August 1979), Carnegie-Mellon University, which had been circulated since 1979.

1.1 The firm as a cooperative resource allocation mechanism

A firm is an organization of human-resource holders which engages in production activities. This section provides a critical review of thought on the *raison d'être* of the firm given by basic literature in the theory of the firm.

Before proceeding to such a review, clarification of the term “organization” is in order. A customary definition of organization given in organizational science is that it is a system of activities of two or more persons, consciously coordinated and controlled by the executive in order to achieve a set of objectives. Empowered with authority, the executive enforces coordinated activities. Barnard (1938) argued that there is a gap between the organizational objectives pursued by the executive and the personal objectives held by the individual members, so that successful organizational performance would require both effectiveness of activities and satisfaction of individual motives. Underlying Barnard’s argument is his perception that an organization consists of people having diverse and conflicting interests, yet each person gains from interacting with others, in particular through their coordinated activities. Throughout the present monograph, an *organization* is characterized simply by (1) diversity of (and most likely, conflict among) the interests of its members, and (2) coordination of the members’ activities. How coordination is agreed upon and enforced depends upon a decision-making rule in each organization. The top-down decision-making rule based on the authority of the executive, that is characteristic in most of the organizational science literature (such as Max Weber (1922), Barnard (1938), March and Simon (1958), Radner (1975)), is actually one of many rules. Unwritten practice in a partnership is another rule. Other rules include the unanimity or majority decision rules in workers’ idealized profit-sharing cooperatives without a hierarchy. The present organization concept does not depend upon specificities of these

Theory of the firm

3

rules, so the economic theory to be developed in Chapters 3–5 has broad applicability. See Remark 3.1.8.

In the neoclassical economics a firm is not conceived as an organization, but is merely identified with its production set (Definition 3.3.3). The only organization is a nation-wide market system in which, guided by the prices observed in the markets, individuals unconsciously coordinate their consumption activities and production activities. Suppose economic agent A has a plant as his initial endowment and wants to use it for his production activity. All he has to do is, for example, go to the perfectly competitive labor market, ask laborer X to supply labor in his factory at 8:00 a.m. at the wage rate established in the market, ask laborer Y to come to the factory at 9:00 a.m. to take over X 's work and perform another task, again at the established market wage rate, and so on. Since agent A can find any type and any quantity of human resources in the perfectly competitive labor market at any moment, there is no need for him to organize a firm; he has only to purchase all the necessary labor on a freelance basis.

In his seminal paper, Coase (1937) asked why there is a firm as an organization. In the presence of the market system, some of the neoclassical hypotheses on the system have to be replaced by an alternative hypothesis which would explain the formation of firms. Coase argued first that the neoclassical hypothesis of free complete information on the labor market has to be replaced by the hypothesis of the costliness of obtaining information on the labor market. Economic agent A in the example of the preceding paragraph does not know whether the wage rate he settles at after negotiation with laborer X is really the market rate, or whether X actually has the intention or ability to supply as high quality labor as promised in the contract. In reality, agent A finds it costly to identify the right labor and relevant wage rate each time he tries to use the labor market. Based on this hypothesis of incomplete information on the labor market, Coase then argued that an entrepreneur (like agent A of the example) does not fully rely on the labor market when assigning labor for his production activity, as such labor allocation would be costly if done properly, but instead forms an organization of human-resource holders, called a firm, and within this organization dictates labor assignment each time he undertakes production planning. Rather than detailed specification of labor, obedience up to certain limits, or more generally, willingness to coordinate labor supply to specific instances during the laborer's tenure, is the object of the contract between the entrepreneur and a laborer. Thus, a firm is formed as a human-resource allocation mechanism which is a superior alternative to the market mechanism in default of complete information. In Max Weber's (1922) view, the market as a place of exchange and the firm (organization) as a place of conscious coordination provide two of the basic complementary

4 Cooperative nature of the firm

functions in the modern society. Coase (1937), on the other hand, attributed the formation of firms to a market failure.

At this point, it is necessary to fix the following terminology, in order to facilitate further discussion in this chapter: A *resource allocation mechanism* is a set of specified forces in a society which together transform *a priori* given initial resources of the society, and allocate the transformed resources among its members. The *social outcome* of a resource allocation mechanism is the final allocation of the transformed resources. The market mechanism is a particular resource allocation mechanism that works in the organization called the market system, and the competitive equilibrium is its social outcome.

Arrow (1969) classified the possible causes for Pareto non-optimality of social outcome into (1) purely technological conditions (such as increasing returns to scale) and (2) specific modes of economic organization (such as the market failure), and discussed the latter causes. He viewed the market failure as the situation in which instituting the perfect market system to achieve Pareto optimum at its competitive equilibrium is not worthwhile, as it would be too costly. In the presence of externalities (specifically, in the presence of a public good), for example, the exclusion principle and hence Pareto optimality of the equilibrium could be restored, if [the good held by agent i and enjoyed by j] and [the same good held by i and enjoyed by another agent j due to the externality] were distinguished as different commodities, that is, if sufficiently many markets were introduced. To introduce such an abundance of markets and to enforce the price-taking behavior as opposed to the free-riding behavior would be too costly, however, so in reality there are fewer markets than necessary at the sacrifice of the exclusion principle, hence Pareto non-optimality of the social outcome of the resulting imperfect market system. Another example of market failure, in the presence of uncertainty, is a situation caused by asymmetric information among the agents. The competitive equilibrium realized in the complete set of state-contingent markets given the *ex ante* information structures is Pareto optimal relative to the same information structures, but it is in general not Pareto optimal relative to the *ex post* information structures obtained by each agent's observing the other agents' behavior (Radner, 1968). Arrow then defined the generalized concept of *transaction cost* as the cost of using a resource allocation mechanism for achieving Pareto optimality of its social outcome. The market failure, in his view, is the extreme case in which the market mechanism involves too high transaction cost to justify its use. In this case, Arrow argued, there will arise pressure in the society to overcome it by introducing additional resource allocation mechanisms that possibly have lower transaction costs. He was aware of the nature of the firm *à la* Coase as an alternative resource allocation mechanism (Arrow, 1969, p. 20), but he mainly (1) discussed

problems inherent in another mechanism, decision-making in the government, and (2) proposed one force that can be internalized in the market system, norms of social behavior.

Arrow's concept of transaction cost is a general notion: It means whatever cost of using an organization for allocation of resources, and includes specifically Coase's notion of the cost of obtaining information on the imperfect labor market. To be precise, the *transaction cost* concept should be re-defined (as done implicitly in the literature) relative to the set of agents who would bear it. In the preceding paragraph it is defined relative to all members of the society, so achieving a Pareto optimal social outcome is the natural criterion. To choose a resource allocation mechanism that has lower transaction cost then becomes an issue of normative economics. In the example of the third and fourth paragraphs of this section, the transaction cost is borne solely by agent *A* (at least, that is how Coase formulated his model), and is defined as the cost of achieving his optimal state in this context, i.e., the cost of identifying the right labor and the right wage rate in the imperfect labor market. Agent *A* will go ahead to create an alternative resource allocation mechanism (firm), in an attempt to replace the so-defined transaction cost by the lower transaction cost of using the firm. To study how these firms are formed and how their mechanisms work in conjunction with the market mechanism is an issue of descriptive economics.

Malmgren (1961) emphasized the advantages to a firm of controlling (in particular, of processing) information, and of monopolizing better information. Arrow (1974) pointed out several issues pertaining to the information processing within a firm.

The subsequent development of thought on the *raison d'être* of the firm took place basically (1) by making precise the contents of the transaction cost associated with any organization, the imperfect market system or another form of organization; (2) by clarifying the specific mode (i.e., specific internal organization) of the firm as an organization; (3) by comparing firms having different specific modes (e.g., Chandler's (1962) unitary structure versus multidivisional structure); and (4) by studying the workings of these resource allocation mechanisms by specifying the behavioral principle and (conflicting) motives of the members of an organization. To appreciate this development, it is necessary to see the precise contents of the transaction cost that Coase originally had in mind. Recall again the example of entrepreneur *A* in the third and fourth paragraphs of this section. The Coasian transaction cost is best expressed in the following passage:

In order to carry out a market transaction it is necessary to discover who it is that one wishes to deal with, to inform people that one wishes to deal and on what terms, to conduct negotiations leading up to a bargain, to draw up the contract, to

6 Cooperative nature of the firm

undertake the inspection needed to make sure that the terms of the contract are being observed, and so on (Coase, 1960, p. 15).

Coase contended that the firm is created in order to reduce these specific costs (provided that the transaction cost of using the firm is lower); specifically, the firm (1) internalizes the cost of obtaining information on the available labor, and (2) replaces a succession of short-term complete contracts by a single long-term incomplete contract which essentially stipulates only the limits to A 's authoritative power, thereby avoiding the costs of bargaining and concluding too many short-term contracts. Here is a minor logical gap: It does not follow from his specification of the transaction cost (listed above) why the organization with this specific mode and this specific hierarchical structure has the lowest transaction cost and satisfies A the most. There might be another mode such that some of the costs listed above may not be minimal, yet the overall transaction cost is minimal because other costs not listed above are sufficiently cut down. Agent A might choose yet another mode which does not minimize the total transaction cost, because it brings in sufficiently large revenue and hence the maximal utility. As possible other modes, it suffices to think of use of a subcontractor rather than an employee, and formation of a partnership rather than a strict hierarchy. Even if this logical gap is somehow filled, there still remain unsolved more serious and fundamental problems; these are the issues introduced in the next paragraph.

The central thesis of Coase (1937) is that a firm is formed as a human-resource allocation mechanism which is a superior alternative to the market mechanism in default of complete information. The central drawback of Coase (1937, 1988) is that he has not clarified what precisely this alternative human-resource allocation mechanism is. Economists want to capture the laws prevailing in the economy as properties of a social outcome, indeed Coase wants to "understand the working of the economic system, to analyze many of its problems in a useful way, or to have a basis for determining policy" (Coase, 1988, p. 6, final three lines), but the very mechanism that gives rise to the social outcome is not pinpointed here. There are at least two issues that one has to address (and Coase did not) in formulating the required resource allocation mechanism. The first issue is the effective incentives of each economic agent. Even in the authoritative organization that Coase suggested as a firm, entrepreneur A cannot make a successful decision without having a good sense of what interest laborer X may pursue or under what constraint X may try to realize his best interest. One may postulate, as in the neoclassical economics, that the ultimate interest of X is represented by the utility level of his consumption bundle. The constraint of X is, on the other hand, influenced by other economic

variables, like the willingness of another entrepreneur B to make a new employment contract with X , so it should be endogenously determined as a part of the social outcome. The second issue is the behavioral principle of the economic agents. Coase was aware that the concept of organization is something new to the neoclassical economics. He brought in without questioning, however, the neoclassical noncooperative behavioral principle as the behavioral principle of the members in an organization. Coase's error in this respect is illustrated in the following discussion, again using the example of the third and fourth paragraphs of this section: It is a one-sided view to assume that the transaction cost is borne solely by entrepreneur A . Just as A finds it costly to identify the right labor and wage rate, laborer X also finds it costly to identify the right capital to work with and the right wage rate to request. There is a host of literature in the search theory; see, e.g., Lippman and McCall (1976a, 1976b). Rather than fully relying on the labor market to obtain information on its demand side each time he wants to supply labor (say, once every day), agent X joins a firm and accepts his daily work assignment, or better still (in case his labor is sufficiently differentiated from the others'), makes every effort to have the assignment made to his best interest. Without joiners, like X , a firm cannot be formed. The firm's ownership is a separate issue. Thus, an organization of agents $\{A, X, Y, \dots\}$ is formed as a firm to minimize the transaction cost borne by the coalition $\{A, X, Y, \dots\}$. Here, the concept of organization should be understood broadly as defined in the second paragraph of this section. Formation of a firm should be viewed as a consequence of cooperative rather than of noncooperative behavior.

For a clear understanding of the difference between the two types of behavior, cooperative versus noncooperative, a brief digression is in order. The distinction lies in the very definitions of the behavioral principles: *Cooperative behavior* takes into account joint decisions made by more than one agent about their strategy choice, while *noncooperative behavior* does not. Consider a manufacturer as an organization of agents $\{A, X, Y, Z, \dots\}$, in which A is the sole executive, and suppose that a specific coordination of activities is to be enforced according to the top-down decision-making rule. If A 's order is too harsh for worker Z to take, Z will quit the organization. In this case, Z decides by himself (i.e., noncooperatively) to deviate from the organizational joint strategies. If two engineers, X and Y , find out that by putting their efforts together they can establish a new engineering company which will bring higher welfare levels to both of them, then they will create their own firm $\{X, Y\}$. In this case, X and Y decide jointly (i.e., cooperatively) to deviate from the strategies of the original organization. This joint deviation is characteristic only of the cooperative behavioral principle. (Chapter 2 reviews formal analysis of various types of descriptive

8 **Cooperative nature of the firm**

cooperative behavior. Specifically, the fourth through sixth paragraphs of Section 2.1 contain the formal definition of a cooperative game, and Remark 2.3.6 presents the view that contrary to the popular conception, the noncooperative “foundation” of cooperative game theory is actually not a foundation of the cooperative behavioral principle. Indeed, the noncooperative behavior is a special case of the cooperative behavior in which only singletons are allowed to form.) The postulate of cooperative behavior becomes more appropriate, if one studies a manufacturer $\{A, B, X, Y, Z, \dots\}$, in which two executives, A and B , consult with each other in making major corporate decisions. Noncooperative decision-making by A and B (say, the passive decision-making formulated by Cournot (1838) and Nash (1950)) is not appropriate within one organization. The need for the cooperative behavioral principle is more compelling in situations in which several firms are involved. In the above case of Z 's quitting the firm $\{A, X, Y, Z, \dots\}$, Z does so typically because he finds a better job in another firm, say in $\{A', B', \dots, X', Y', Z', \dots\}$. What is going on here is that the new firm $\{A', B', \dots, X', Y', Z', \dots, Z\}$ is formed (although it may well keep the same name as that of $\{A', B', \dots, X', Y', Z', \dots\}$). Forming this new organization is a cooperative behavior (i.e., the employer(s) and the new employee agree on an employment contract), and this coordinated strategy-choice stipulates in part Z 's resignation from his old firm $\{A, X, Y, Z, \dots\}$.

Alchian and Demsetz's (1972) study was the first explicit treatment of the role of incentives of every member in a firm. Laborer X wants to pursue his self-interest, so rather than put all his effort into his firm as promised in the contract, he may devote some of his effort to his personal enjoyment at the sacrifice of the firm's productivity. As long as his shirking behavior is not discovered, he will shirk. More generally, to the extent that his colleagues' observation of his specific labor supply (both quantity and quality) is consistent with the one stipulated in the contract, he will behave in his best interest. Agent X is a decision-maker himself just as his employer is, although the ranges of the decision variables are quite different. Alchian and Demsetz identified this monitoring cost as the transaction cost. A firm is formed to reduce this monitoring cost by appointing a monitor. But the monitor himself may shirk. To avoid this last problem, he must be given an incentive for monitoring such that the more accurately he monitors, the higher reward he receives. He is thus made a residual-claimant. This is Alchian and Demsetz's reasoning for the formation of a specific type of firm, called a classical firm. They also extended implications of this transaction cost (monitoring cost) to other modes of organization, such as profit-sharing firms, socialist firms, big corporations in the capitalistic economy in which the major decision-makers are not stockholders, etc.

Two of the key elements in Alchian and Demsetz's argument are

asymmetric information between an employer and an employee and a difference in their interests. The workings of resource allocation mechanisms that would emerge in an organization characterized by these two elements have been studied in the recent literature of the principal-agent theory, albeit under the postulate of noncooperative behavior and in most cases within oversimplified frameworks. See Remark 1.1.1 at the end of this section. Hart (1988) forcefully argued the limits of the principal-agent theory, and provided an account for the appropriateness and power of analysis of the cooperative behavioral principle.

It appears that Alchian and Demsetz were fully aware of the cooperative nature of the firm, when they wrote:

Team production . . . is production in which 1) several types of resources are used and 2) the product is not a sum of separable outputs of each cooperating resource. An additional factor creates a team organization problem – 3) not all resources used in team production belong to one person (p. 779).

Alchian and Demsetz even talked about a firm having more than one major decision-maker, like a corporation. The purpose of Chapters 3 and 4 of this monograph is to develop an axiomatic theory in order to study the workings of the human-resource allocation mechanism that will emerge in a society characterized by conflicting interests and a cooperative behavioral principle. Recall that there are at least two issues that one has to address to pinpoint an organizational resource allocation mechanism: agents' effective incentives and the cooperative behavioral principle. The axiomatic theory addresses both. In fact, Chapters 3 and 4 study the workings of the conjunction of (1) the cooperative mechanism for the allocation of human resources and (2) the neoclassical market mechanism for the allocation of non-human resources.

The theory developed in Chapters 3 and 4 is descriptive, as it is intended to describe the present-day capitalistic economy and a certain socialistic economy. An organization (firm) is formed as a result of its members' pursuit of self-interest; the members do not care about the society's overall performance. Indeed, its social outcome is in general not Pareto optimal. Chapter 5, on the other hand, develops a normative theory within a simplified framework. In the presence of the increasing returns to scale technology, a new resource allocation mechanism is proposed, so that its social outcome is Pareto optimal.

Williamson (1975, 1985) was the most prominent in identifying various types of transaction costs and their influences on the modes of realized firms. This paragraph reorganizes some of these types and presents them accordingly. Needless to say, the basic concept here is the organization characterized by conflicting interests and coordination of strategies (the

10 **Cooperative nature of the firm**

second paragraph of this section). The first characteristic (conflicting interests) combined with asymmetric information has the tendency to create the specific behavior emphasized by Williamson as “opportunism.” (The point here is that contrary to his treatment, “opportunism” should be explained as a phenomenon the extent of whose realization is endogenously determined in the model.) Williamson (1985, pp. 52–61) introduced three dimensions for transaction cost so that “the factors responsible for differences among transactions be identified and explicated” (p. 52): (1) asset specificity, (2) uncertainty, and (3) frequency. Dimension (1) refers to the job-specific nature of inputs, which does not exist before the contract but shows up continuously after the contract. Examples are site-specificity and specialized labor obtained by on-the-job training. Dimension (2) includes specifically environmental uncertainty and asymmetric information. In the above quotation from Coase (1960, p. 15), there were five contents of transaction costs; the first four are typical examples of dimension (3). In order to see the exact roles that these three dimensions play in Williamson’s argument, it is more appropriate to call them three *types* of transaction costs. Another type of transaction cost is: (4) bounded rationality. Radner (1985) contains three precise formulations of the bounded rationality concept, each interpreted differently. The preceding paragraph showed that Alchian and Demsetz emphasized asymmetric information (a specific case of uncertainty) to explain the formation of a classical firm (one mode of organization). Williamson extended the Coasian approach: Asset specificity and frequency in general lead to a long-term contract. A manufacturer’s inability to extend quality control to distributors (a specific case of asymmetric information) and hence the possibility of reputation-damage lead to forward integration into distribution. In the US railroad industry of the nineteenth century, the site specificity of operating matters (a specific case of asset specificity), like equipment utilization and maintenance, led to formation of the decentralized line-and-staff mode of organization. In the twentieth century, bounded rationality led to creation of the multidivisional mode of organization (i.e., semiautonomous profit centers organized along product, brand, or geographic lines), and furthermore to creation of conglomerates, and so on. Applicable to all these arguments are exactly the same criticisms as those addressed to Coase (1937, 1988) earlier in this section. It is true that Williamson sometimes looked carefully at the role of effective incentives of agents; the most beautiful example is his explanation of the limits of firm size (Williamson, 1985, Chapter 6). Williamson has carefully identified specificities of various types of transaction costs. Taking them into account while pinpointing and formulating a non-market resource allocation mechanism would surely yield sharper properties of the associated social outcome – but this task is beyond the scope of this monograph.