Theories of the firm

Dr Udo C Braendle
Strategy, Economic Policy and Legal Affairs (Kommunalkredit Austria)
Honorary Visiting Fellow (School of Law, University of Manchester)
Email: udo.braendle@manchester.ac.uk

Keywords: Transaction cost theory, firm vs market, incomplete contracts, agency theory, team production, management, positive theory of the firm

Abstract:
Coase (1937) was one of the first scholars who asked why firms exist and what precisely a firm was. Both questions are fundamental to understand corporate governance. Before the 1930s the firm was often seen as a “black box” which was assumed to behave like any other self-interested utility maximizing economic actor. Although Adam Smith already cited the problems like the separation of ownership and control in firms, it took more than 150 years before economists such as Coase and Williamson put theories around these questions. In the meantime catchwords like agency theory, corporate social responsibility and corporate governance made this research area very prominent. This chapter examines the different theories of the firm, legal and economic ones, how they are connected and what they mean for the corporate governance discussion. But this chapter shall also show the limitations of these theories and present some outlook for new theories of the firm.
“The directors of such companies [joint stock companies] however being the managers rather of other peoples’ money than of their own, it cannot well be expected, that they should watch over it with the same anxious vigilance [as if it were their own].” (Adam Smith, 1776)

1. Introduction

Theories of the firm are ways of conceptualising the firm. The answers to the questions why firms exist and what precisely a firm is are fundamental for the understanding of corporate governance. Theories of the firm not only try to answer why businesses are organised in firms but how the relationships within the firm as well as between the firm and society at large look like.

Since the seminal article of Ronald Harry Coase on the nature of the firm (Coase, 1937), these questions were brought to the attention of a large number of economists and a growing number of lawyers, mainly in the area of Law and Economics. Many economists (still) use the tool of neoclassical economics to explain why business activities are carried out with the structure of a firm and to develop policy implications in corporate governance.

Company law and corporate governance proposals are based on particular understandings of what firms are for and whose interests they should serve. Therefore the theory of the firm is an indispensable starting point for corporate governance studies.

Before the 1930s the firm was very often seen as a “black box” which was assumed to behave like any other self-interested utility maximizing economic actor. This view was based on the belief about the firm’s ability to almost instantaneously adjust itself to a changing environment. Consequently resources of a firm were assumed to be put to their most efficient use without having a look “inside” the firm. It was treated as an entity competing with other firms in the market. Although the limitations of this macroeconomic view have already been cited by authors like Adam Smith (1723-1790), the contemporary legal concept of separate legal personalities of companies supports this theory. Only this broad and abstract perspective of firms can identify problems such as monopolies and oligopolies, where one or a group of firms are able to drive competitors out of the market. Antitrust law is a response to these sorts of market failures.

But at the same time firms are as well a collection of individuals – all having their own preferences and values - tied together in legal relationships within the legally constructed black box of the firm. Although this microeconomic view of firms is less abstract than the black box perspective, it does not explain why these individuals prefer organising themselves into business structures rather than remaining independent and making contracts on their own.
In a free market environment the answer must lie in significant economic benefits of organisation within the structure of a firm compared with contractual relationship on a market. In his seminal 1937 article, Coase developed a theory of the firms which was against the mainstream literature at this time and put emphasis on these relationships within the firm. Coase challenged economists and was the starting point of various theories of the firm. What was originally a research area for a few economists, has in the meantime become a playing field for lawyers and economists interested in corporate governance. Reviews of the theories of the firm and catchwords like the separation of ownership and control, agency theory and corporate social responsibility make this research area very prominent.

This chapter will, starting from Coase, present the transaction cost economics of the firm (section 2). Both theories look at the firm as an alternative to markets. Oliver Williamson, an institutional economist, further developed Coase’s theory of the firm through a deeper analysis of different forms of contracts (section 3). As most theories of the firm are based on the idea of a firm being a nexus of incomplete contracts (which will be dealt with in section 4), opportunism may influence the relationships within firms. The latter are based on asymmetric information, where one party like the directors of a company can misuse the informational advantage to exploit another party like the shareholders. Incentive systems and monitoring are two measures to reduce this agency problem which will all be discussed in section 5. Section 6 has a closer look at theories around the management and collective production in companies.

But this chapter will not only present theories of the firm but try to highlight the relevancy of each of these theories for the corporate governance debate and what practical importance they have for policy makers. In addition, this chapter will also show the limitations of these theories and present some outlook for new theories of the firm (section 7).

2. The theory of the firm – the theory of transaction costs?

The theory of the firm was traditionally one branch of Microeconomics which studied the supply of goods by profit-maximising agents. In this theory, production costs played a crucial role. Coase (1937) was one of the first pointing out that in addition to production costs of the usual sort, one must also consider transaction costs in explaining institutions like the firm. He focused on the comparative transaction costs of alternative organizational structures, such as firms and markets. This theory was later extended by Oliver Williamson and became widely...
known as transaction-cost economics (Williamson, 1985) or more broadly the economics of organization.

Transaction costs are costs (e.g. in terms of money or time) incurred when making an economic exchange. If we extend this term, transaction costs do not only include bilateral transactions but subsume contractual relationships between individuals. In general, transaction costs symbolise “friction losses”, i.e. the lost resources for the involved parties, but which are inevitable to reach certain goals. In firms, transaction costs may include the costs of organising business activity over time, planning the future and limiting as well as allocating risks which may arise in the future. It therefore includes the elements of uncertainty and opportunism, which are both indispensable for debates in corporate governance.

Coase argued in his 1937 article that transaction costs explain both the existence of firms and their optimal size. In “The Nature of the Firm” he identified certain transactions which are prohibitively costly if the parties involved could only deal with instant market transactions. In order to carry out a market transaction it is necessary to identify the party one wishes to deal with, establishing terms and conditions, conducting negotiations and concluding a contract. After the conclusion of the contract monitoring is needed to make sure that all terms and conditions are fulfilled. If slight changes are wished, the whole transaction process needs to be initiated again. Or, to put it in other words, Coase emphasised that making contracts and purchasing assets and other property in markets incurred costs that were not accounted for by the “price mechanism”. Individuals would therefore organise firms and maintain them when the organisational entity provided implicit savings in terms of assembling resources, assets, and labour internally.

This describes situations in which market transactions would show their relative inflexibility to re-contracting when changes in the existing relationship arise. Regularly recurring transactions and long-term transactions might be good examples. In such situations longer, incomplete contracts, which are typical for firms, provide much more flexibility for the parties in a world of uncertainty. These contracts can be left open to be flexible in case of a changing environment.

On the other hand dissimilarities of transactions, the probability of changes in the market prices for the relevant resources as well as the spatial distribution of the relevant resources and transactions highlight factors which increase the costs of using a firm.

One might argue in this context that transaction costs would be minimised in a world without transactions. This could be achieved if rights and duties would initially be assigned in the “right” way.
Based on this idea Armen Alchian and Harold Demsetz built their theory of property rights. Property can be tangible (e.g. equipment in a firm) and intangible (intellectual property), and property rights theory argues that the ownership, which includes residual rights to the benefits of ownership, of productive assets provides a foundation for explaining firms. According to Oliver Hart, one of the leading scholars in this area, a firm without property is just a phantom (Hart, 1995). In situations where ordinary contractual relationships fail, firms arise and the ownership of capital assets puts (collection of) persons in the position to organise production through the purchase of economic factors, including labour (Hart, 1995).

Applied to corporate governance, this theory provides a supplement to contract theories. The theory claims that legal systems should assign and secure property rights and additionally explains that those who invest in or own productive property and capital of the firm, have a privileged position as legal agents to bargain with other parties such as directors, employees, suppliers, and other constituencies.

Coase, in his theory of the firm, built a connection between the above discussed property rights and transaction costs. In a world without the latter, the initial assignment of property rights would be irrelevant as each “error” in the assignment of these rights could easily be rescinded by additional transactions. This is the idea of the Coase Theorem for property rights. The applicability of the Coase Theorem to companies is questionable as the idea rests on the assumptions that there are no legal, strategic and informational barriers to bargaining. But in modern firms all these barriers normally exist and make transactions more expensive, i.e. incur transaction costs.

3. Beyond the firm and market dichotomy

Institutional economists like Williamson have further developed Coase’s theory of the firm through a deeper analysis of different forms of contracts. He developed governance structures which can be seen as a spectrum of contracts between the extremes of Coase’s market versus firm analysis. His theory of transaction costs provides answers why firms change organisational structures over time.

Modern phenomena such as competition law, merger control as well as outsourcing can be explained with the help of this theory. Concerning the latter, a firm may find that it cheaper in terms of direct and administrative costs as well as legal liability to conclude an arms-length contract for labour services rather than bearing the cost of hiring employees directly. This influences the size of a firm (which won’t grow) but the outsourced services will expand the
general market for those services. In a broader context this has implications on merger control. By suggesting circumstances in which integration (or outsourcing) might be efficient, transaction costs theory influences decisions of firms. This in turn has implications on competition law as vertical as well as horizontal mergers may disrupt competitive markets and lead to market structures which need regulation to enforce competition.

In his theory, Williamson presented the firm and instant market transactions as lying on a spectrum of forms of organisations (“governance structures”) rather than as simple alternatives. The three forms of intermediate structures are the classical, neo-classical and relational contract. The classical contract is characterised by full contingency, i.e. the contract should be as complete as possible, looking for regulations for each uncertainty. The contract should provide a complete guide to the terms of all consequential transactions.

A neo-classical contract, however, leaves some terms open. Here the goal of the parties is not completeness of contracts, but to establish mutually acceptable mechanisms to minimise the costs of leaving contractual terms open. To put it in other words, leaving parts of a contract about an uncertain future open should allow efficient re-negotiations at a later stage. But this incompleteness of contracts opens room for exploitation of one of the contractual parties.

To assure flexibility whilst minimising the risk of opportunism, special legal devices were developed. These include referees and arbitrators as third parties to resolve disputes as well as external reference points such as a stock indices or interest rates. All these devices have one common feature, namely reducing the risk of exploitation of one of the parties.

The relational contract as a third governance structure stands for future decision making rather than attempting to make decisions in advance. For the parties involved, the goal is a trade-off in terms of maximising the gains of the transaction whilst trying to minimise transaction costs. The contract itself prescribes the governance structure, its operations and may restrict and place boundaries on its discretion. To avoid a wide degree of discretion, some of the neo-classical safeguards mentioned above may be used.

Compared to Coase and his extremes of firms and markets, Williamson’s governance structures can be seen as a spectrum between these boundaries. Given a certain set of transactions, Williamson identified factors for the optimal choice of governance structure. This is influenced by the frequency of transactions as well as the transaction-specificity. Concerning frequency there are likely to be efficiency gains in making long term arrangements if one party needs a particular resource. A transaction-specific investment arises when one party requires an unusual or idiosyncratic resource. Consequently alternative sources of supply are unlikely to be readily available. As one or both parties make greater
investments in the transaction, they are more vulnerable to opportunism by the other party and therefore will try to make satisfactory arrangements covering the future. This problem is in the literature well known as the *hold-up problem* and refers to the above mentioned question of integration. Just think about a typical market example: A seller S can either produce a product demanded by a buyer with a machine B, which is cheaper to acquire (and maybe used to manufacture other products) but which may incur high variable costs (due to set-up costs) in producing the product for a buyer. On the other hand, S could acquire a more expensive machine A, which is more specific and therefore incurs less variable cost. In acquiring machine A, S would put herself in a position where she is at the mercy of the buyer. If the buyer decides to quit the relationship, S would possess a machine which cannot be used for other customers. Although this transaction-specific investment (acquiring machine A) would be beneficial for both parties, a lot of these mutually beneficial transactions do not take place as one party fears being exploited. A solution to this hold-up problem would be integration. The buyer could integrate the seller into her firm and thereby eliminate the risk of being exploited. The above mentioned merger control and competition law are consequences of such developments.

4. Incomplete contracts

The previous hold-up example explains a situation which is typical for many business relationships. One might now argue now that this problem could easily be avoided by writing a *complete contract* between buyer and seller. A complete contract is a concept in contract theory that describes an agreement that would specify the respective rights and duties for every possible future state of the world. As it would be prohibitively expensive to write such a complete contract, contracts in the real world are usually incomplete. Just think about a typical employment contract of a chief executive in a company. It would be impossible to write a contract which makes clear what this executive should do in each future situation. Furthermore, not only that the future remains uncertain, even if a detailed contract could be written, it would be very costly.

To avoid confusion, it is necessary to point out that this concerns a contract from an economic perspective. As should be evident from our contract law classes, a contract is a binding agreement between two or more parties to perform certain obligations, whereas a complete contract is one which includes all *essentialia negotii*. From an economic perspective, a
contract is everything which pictures observable states, or, to put it in other words, an instrument which facilitates an exchange of property rights.

It is fairly obvious that the definition of an economic contract is much broader and therefore a complete contract more difficult to achieve. It would need explicit regulations to cover every possible state in the future, and to highlight the rights and duties of each party. A complete contract therefore remains a fiction.

Basically there are six reasons why a contract (in economic terms) remains incomplete: Firstly, ambiguous diction within the contract. In addition, the parties may have forgotten about regulations for certain points of the contract. Thirdly, despite the awareness of the parties, the costs of bargaining for some issues might be prohibitively high. Asymmetric information is another reason for incompleteness, an area we will discuss in more detail in the next section. Bounded rationality is a concept based on information economics which gives an explanation why contracts are incomplete. It rests on the idea that economic actors do not have perfect information and/or may not be able to process all available data adequately. Therefore the amount of information plays a crucial role. Think again about the employment contract with the chief executive of a company. Even if it would be possible to foresee the future and write a complete contract for every possible state that may occur, it would be impossible to write such a voluminous contract in time and no party would be able to process all of the information in the contract.

Bounded rationality asserts that decision makers are intended to be rational, i.e. they are goal-oriented and adaptive, but because of human cognitive and emotional architecture, they sometimes fail to be so, even when it comes to important decision making. Finally, the heterogeneity of the market serves an explanation for incompleteness of contracts.

If we assume that the firm is a nexus of contracts, i.e. emphasising the network of different kind of contract made by individuals to compose a firm, the incompleteness of contracts has substantial consequences. As contracts are incomplete, parties have to agree on the allocation of control rights for all situations that are not specified in the contract. The resulting institutional design constitutes what we describe as corporate governance. It assigns control rights for the use of the firm’s assets.

In the “nexus of contract” theory of the firm contractual relationships are the essences of the firm. The firm is simply one form of legal fiction which serves as a nexus for contracting relationships (Jensen and Meckling, 1976). As a consequence, corporate law is merely enabling. This theory also presents a perspective that emphasises contracts which can explain
why phenomena like outsourcing as well as entering into joint ventures or strategic alliances as alternatives to direct growth take place. Unfortunately these contractual relationships are characterised by fundamental problems which are described by agency theory.

5. Agency theory

One of the key elements of agency theory is opportunism, a point stressed heavily by Williamson. If one party (the agent) has discretion which she is supposed to exercise for the benefit of another (the principal), she may exercise it to maximise her own utility instead. This is inefficient where the resulting loss to the principal exceeds the benefits to the agent. If the agent is rewarded by the principal on a basis which does not correlate her effort to the reward, the agent may not have the incentive to exercise the highest effort. The costs resulting from this agency problem includes both the loss of potential benefits and the costs of measures designed to reduce the loss of potential benefits. Michael Jensen and William Meckling (1976) identified these costs and termed them agency costs.

Applied to corporate governance, legal protection against fraud and other forms of dishonesty may provide some protection. But economic analysis suggests that internalising some of these market transactions into a firm may substantially reduce the risks of opportunism. But despite reducing some of the costs of opportunism in the market, the special structure of firms creates other forms of opportunism in those entrusted with economic responsibility to manage the firm.

Agency theory is based on the incompleteness of contracts and the separation of ownership (shareholders) and control (management), which is the main characteristic of corporations nowadays. Though the resulting problems were already mentioned by Adam Smith in the 18th century, they were prominently highlighted by Adolf Berle and Gardiner Means in the 1930s. They argued that that a company does not behave in accordance with the classical model, which assumed that despite the management of companies by agents, these agents act in the best interest of the owners of the firm. As a consequence the owners would ensure that the agents do act in their interest. But this idea, the stewardship theory, assumes that managers do not necessarily work only in their personal interest. Managers are seen as “honorable wealth builders” and can therefore be instructed with the management of corporations.

Berle and Means argued that the interests of managers and shareholder may diverge and that shareholders would not act as owners, exacerbating the agency problem. Concerning the first argument the stockholders (principals) want their agents (managers) to maximize the value of
the shares. But the manager may be better off pursuing some other strategy and therefore acting opportunistically. Despite the law asserting that an agent has a fiduciary duty to serve the principal’s interest, the agent will also tend to serve their own self-interests. If we think about staff, sales, size of the firm, remuneration, etc., being an input-vector, we can assume that the vector which maximizes the profits of the firm does not necessarily maximize the manager’s wealth. We can assume that most managers will choose the vector which maximizes their own wealth. The recent financial scandals with stories about greedy managers all over the world confirm this picture.

On the other hand, the second point of Berle and Means should not be forgotten. Due to the shareholders’ perceived “limited liability” and the shareholders’ inability in practice to control the management, the agency conflict is exacerbated. Limited liability means that that a company is responsible for its own debts and liabilities. Shareholders are only liable to the company to pay up their share capital. In other words, they are sharing the company’s profits, but they are not responsible for all of its losses. Limited liability, so the argument goes, shifts the risk of business failure from the company’s shareholders to its creditors. Both, the companies’ owners and managers therefore may have too much of an incentive to take risks, as the creditors would be the party which would suffer most in case of a bankruptcy. This could result in an inefficient use of resources.

The diversity and large number of shareholders in a typical public company cannot or will not exert effective control over the management for various reasons such as the existence of a coordination problem. This includes problems of different interests of shareholders as well as bringing shareholders with the same beliefs together.

In general we refer to the collective action problem, where it might be rational for each of the shareholders not to engage in control. Just think about yourself being one of many small shareholders of a public company. Your incentive for attending the general meeting of the company or exercising control in any other way is minimised as the transaction costs of controlling (in monetary terms as well as time) exceed the benefits. You need to spend time in reading bulky reports of the company, the general meeting might not be close to your home, you have you spend a whole day at the meeting, all of these costs might exceed the benefits you get from voting your rights. It is therefore rational not to engage in the control of the company. With the imminent problem of free riding, i.e. each shareholder wants to avoid the costs of control by hoping that the other shareholders are exercising the necessary control, this leads to a collective action problem. Each of the shareholders is acting rational, when not exercising control. But this leads to a situation where nobody controls the management at all.
All this means that management’s incentive as to how they exercise their powers of management may not be aligned with the interests of the company’s shareholders, giving rise to manifold economic problems, such as various forms of opportunism. Due to the consequent danger of the inefficient use of resources there is a justification for correction. To reinforce the classical model of the company where the interests of the owners and managers of the company are aligned, regulatory measures – mainly in the form of laws and codes – are used. These include strengthening shareholders’ voting rights, e.g. with the help minority shareholder rights. In addition the accountability of the management to shareholders is achieved by imposing penalties on managers when they behave wrongly. Furthermore, enforced publicity and disclosure should reduce the asymmetric information between the parties and therefore lead to better control. All of these measures are reflected in the pattern of UK corporate governance reform and will be critically evaluated within this book. As it is impossible to write complete contracts between the different parties, a first best contract does not exist and leaves a gap which should be filled by corporate governance. All control forms involve agency costs, therefore corporate governance revolves around finding the control which is minimises agency costs.

Agency theory relies on the idea of writing a contract to align the incentives of both parties involved. In other words, it tries to strike a balance between providing incentives for managers and guaranteeing a maximisation of shareholder value. To achieve this goal, monitoring refers on the one hand to strategies of managerial supervision and oversight to improve performance. This explains the existence of board systems and other external monitoring such as rating agencies and institutional investors. On the other hand high-powered incentive contracts such as shares and stock-options to remunerate directors were implemented in most companies over the last years. In chapter X we will have a closer look at the pitfalls of these remuneration schemes. For the moment being, two problems of asymmetric information in firms shall be presented. We distinguish between moral hazard (hidden action), as well as adverse selection (hidden information).

5.1. Adverse selection
Adverse selection describes an agency problem where asymmetric information exists before the transaction occurs, leading to an inefficient allocation of resources. Originally the term adverse selection came from the insurance industry. It describes a situation of private information where the insured are more likely to suffer a loss than the uninsured. As an
example, think of two groups of car drivers. On the one hand the careful and risk-averse drivers and on the other hand the reckless, risk-taking drivers. An insurer selling car policies cannot easily assign the different drivers to these groups, so each of them pays the same premium. But reckless, risk-taking car drivers are much more likely causing damages than the careful, risk-averse ones. Consequently the car insurance is a better deal for the beneficiaries of the reckless drivers. As the insurance companies do not want to loose money, they will set the premiums accordingly high. This in turn may result in a situation where careful drivers may want to go uninsured (as the premium is too high for them) and only the high-risk drivers remain on the market, as they benefit from the insurance.

Economists like George Akerlof applied the concept of adverse selection into markets other than insurance, where similar asymmetries of information may exist. Applied to corporate governance, investors very often face problems of adverse selection. When buying a share an investor usually has incomplete information about the management of the company in which she intends to invest.

In his seminal 1973 article Michael Spence (born 1943) proposed a way how two parties could get around the problem of adverse selection. The idea is having the informed party sending a signal that would reveal some piece of relevant information to the uninformed party. The latter interprets the signal and adjusts her purchasing behavior accordingly. In our example the disclosure of information from companies could be interpreted as signals to the stock market.

Another way to overcome adverse selection is screening. Here the uninformed party takes actions to separate different types of informed parties. In the example the investor could try to collect data about the risk of each company and adjust the investment accordingly.

5.2. Moral hazard

Compared to adverse selection, moral hazard describes an agency problem which exists after a transaction is made, leading to an inefficient allocation of resources. The term moral hazard as well has its origin in the insurance industry.

Applied to corporate governance, the shareholder-manager agency relationship is a good example of moral hazard. Directors of a company may, after signing their employment contract, start acting in a way which benefits themselves but not the shareholders. This leads to all of the problems discussed above with managers acting opportunistically.

Monitoring and incentive compatible contracts are two ways to overcome moral hazard. Former builds on the straightforward idea that opportunistically acting managers will stop
doing so if they are detected and penalised. But monitoring itself is on the one hand very
costly and consequentially reduces the flexibility of companies, a fact which is very often
neglected. Incentive compatible contracts, however, are a cheaper way to reduce the moral
hazard problem. With the help of incentives, managers should be motivated to align their
interests to the ones of their shareholders. Bonuses and stock options are just two measures to
achieve this goal. Unfortunately these measures are themselves subject to new problems like
short-term profit maximisation of managers to cash-in their bonuses and stocks, without
thinking to much about the longer term. Another problem, which became especially immanent
during the financial scandals, is that managers used their creativity in accounting to reach
their bonus threshold.
Although most of the literature puts emphasis on this “classical” agency problem between the
suppliers of equity – the shareholders - and management, agency conflicts can arise between
various groups of stakeholders in a firm.

5.3. Different agency conflicts
In addition to this relationship, the conflicts between majority- and minority shareholders as
well as between shareholders and other stakeholders should be mentioned. The latter became
prominent in the literature as the shareholder-stakeholder debate or Corporate Social
Responsibility.
The conflict between majority- and minority shareholders is typical for companies with a
concentrated ownership structure. The conflict revolves around the term of “private benefits
of control”, highlighted by Lucian Bebchuk in 1994. The question is to what extent the
majority shareholders of a company can enjoy additional benefits from the company which
are not shared with the other shareholders. In this context the purchase of shares below market
value or the change of a product mix should be mentioned. Mark Roe describes in his 2003
book the effects of corporate governance in social democracies and illustrates how political
ideas can be used to identify the conflicts between majority- and minority shareholders.
Traditionally some corporate governance systems have to deal with this conflict if we just
think about the power of banks or the state. Both, banks and the state as owners of companies,
may follow other interests than the rest of the shareholders do.
The third conflict is between shareholders and other interest groups such as employees and
creditors and concerns the classical shareholder versus stakeholder differentiation. The
stakeholder model claims that the firm should serve wider interests of stakeholders rather than
shareholders only. Stakeholders such as employees, creditors, suppliers, customers and local
communities have long-term relationships with the firm and therefore affect its long-term success. The company is seen as an entity embedded within the society it operates and therefore should take care of it. In the meantime there is a huge debate all over the world about whether it is the duty of company law in general and corporate governance in particular to solve this conflict between shareholders and other stakeholders or, what the Anglo-Saxon literature and practice favours, that the legislator should deal with this problems exclusively within labour law and other social legislation.

Shareholders and stakeholders favour different corporate governance structures and also monitoring mechanisms. In this context we can differ between the Anglo-Saxon (Common Law) model with its emphasis on shareholder value, compared to the German/Japanese (Civil Law) model where certain stakeholders enjoy special power. The shareholder versus stakeholder debate will be discussed in more detail in this book under the heading Corporate Social Responsibility in chapter X.

5.4. Criticism

Agency theory is a very influential research area to explain and mitigate the problems in the relationships between shareholders and management as well as other stakeholders in a company. Despite, or especially because of its huge impact, there are a number of critics. Ghoshal and Moran (1996) describe very well that the underlying logic of agency theory is “bad practice” for companies. Agency theory is based on the assumption of a certain human nature, which these authors neglect. Instead of assuming that all directors are greedy, it is much more realistic to assume that a small minority is responsible for the bad reputation of all of them. Furthermore, the assumption of greed for all managers can become a self-fulfilling prophecy and should therefore be avoided.

The contractarian view of firms has had many critics, especially from legal scholars. Millon (1990), among others, objected to the normative emphasis given to economic efficiency by those deploying these theories. His article reviews historical perspectives on the nature of the corporation and the implications of these perspectives for business and social responsibility. According to Millon, all views of the corporation that emphasise its private, internal nature leave much less room for social impact discussions than if the public nature of the corporation is emphasised.
6. Management and collective production

Whatever one might think about agency theory, the separation of the two functions, shareholder and management, makes economic sense. The shareholders on the one hand provide equity capital together with subsequent risk-bearing, whereas professional managers on the other hand do not have to possess capital and risk-bearing expertise but provide managerial expertise.

As management can be performed more efficiently in a small, cohesive body, it is better to have a professional management than let a large number of shareholders decide on the day-to-day business. The risks associated with providing equity capital, however, can be borne more efficiently when shared among a large number of people. These, in turn, can protect themselves by holding diversified portfolios of investment.

Armen Alchian and Harold Demsetz’s took up the structure of companies and drew attention to the analysis of team production (1972), itself an extension of the earlier work by Coase. According to this theory, the firm is preferred to the market because of the benefits of team production. Nevertheless, team production always involves a metering problem. Where parties are engaged in some form of collective business activity, there may be difficulties in measuring the contribution made by each participant and matching their reward to their actual contribution. The rewards of the participants of the team may be, to a greater or lesser extent, determined not by their individual effort, but by their collective effort as a team.

The metering problem therefore presents a fertile soil for a form of opportunism known as “shirking”. In a team where each member’s reward is not fully related to their actual input, the member has too little incentive to contribute fully to the team’s activities. However, the incentive to obtain utility in other ways (to shirk) is increasing. As the members enjoy the full benefit of their shirking, but share the costs with other members of the team, shirking is a common moral hazard problem in many firms. Obviously this theory is based on certain assumptions about people’s motivation in working and their affinity to shirk.

As a result, shirking is more likely to be a problem for larger “teams” where shirking is less likely to be detected and where members are less motivated by a sense of loyalty to their entity.

Corporate governance, in this context, should try to minimise the scope for shirking by all participants and conserve the gains from using a firm. In theory, this should be easy to implement as the team members are likely to gain more from less shirking by other team members than they lose from their resulting inability to shirk.
Responses to shirking are the reduction or removal of the metering problem and monitoring in the form of management. The former may be possible if we think about many modern trends in organisation such as outsourcing, franchising and the replacement of employees with self-employed contractors. In these entire organisation forms the traditional team and the problems related to it are reduced.

If shirking undermines the gains from team production, Alchian and Demsetz argue, the efficient solution is monitoring with the help of a hierarchical management structure. This includes a central manager who co-ordinates the individual contributions or inputs. Being the recipient of the team’s residual income, she takes all the surplus profits and bears any losses. Under these conditions the manager has the optimum incentive to maximise the output of the team.

To explain management hierarchies with the help of team production we assume that the monitors themselves are opportunistic and therefore exposed to shirking. Consequently the monitors need monitoring themselves. To avoid problems with the implication that the monitors at the top of the hierarchy have no monitors, incentive systems need to be established. This is one of the tasks of a good corporate governance system.

For company law implications the theory of team production is useful in highlighting problems in the hierarchical structure of companies. If the different monitors do not keep in mind what is best for the company but misuse monitoring to gain private benefits, shareholders should enjoy some rights. They should be able to replace individuals or whole boards that are not monitoring on their behalf.

7. Outlook and conclusion

The market and the firm are two alternative forms of allocating scarce resources. The firm exists because in many cases it is more efficient to organise production in a non-price environment. Coase’s transaction cost theory was one of the first neo-classical attempts to define the firm theoretically in relation to the market. This theory was later extended by Williamson through a deeper analysis of different forms of contracts. He developed governance structures which can be seen as a spectrum of contracts between the extremes of markets and firms. Alchian and Demsetz’s analysis of team production is another extension of the earlier work by Coase. According to this theory the firm is preferred to the market because of the benefits of team production.
By looking at the common elements of all of these theories it is obvious that they provide negative reasons for the existence of the firm. All theories revolve around the existence of firms to overcome inefficiencies. None of these offer a positive theory of the firm, with functions like generating value. But Demsetz himself in his 1988 article considered the positive function of the firm as an efficient device for accumulating, storing and using information.

Whereas most theories of the firm are primarily concerned with predicting the behaviour of firms in external markets, the last two decades brought forward a knowledge-based theory of the firm as an alternative view. Knowledge is seen as the most important resource of a firm. This includes all the intellectual abilities and knowledge possessed by employees, as well as their capacity to learn and acquire more knowledge. Thus, knowledge resources include what employees have mastered, as well as their potential for adapting and acquiring new information. These resources are seen as being extremely important for sustaining competitive advantage in today’s environment. Participants and “input-providers” are viewed as assets rather than sources of inefficiency.

Based on the idea that the firm is a nexus of contracts, the incomplete contract literature helps us to understand the need for corporate governance. If it would be possible to fix the relationships between the owners and the managers of a company contractually consistently, there would be no divergences between the two parties. But as the costs of writing such contracts are prohibitively high, the abstract structure of corporate governance has to be introduced.

Complete contracts, which do not exist in the real world, would also provide a solution to the agency theory. If one party (the agent) has discretion which she is supposed to exercise for the benefit of another (the principal), she may instead exercise it to maximise her own utility. The shareholder-manager relationship is the classical example for such an agency relationship. Agency theory results from information asymmetries between parties, where especially the concepts of adverse selection and moral hazard are of relevance in a corporate governance framework.

Corporate governance as a research area has been affected by all of these theories of the firm. Arguably the agency theory formed the starting point of an area ranging from a number of disciplines including law, economics, accounting, management, and organizational behaviour. Nevertheless, these theories are just the starting point for companies increasingly realising that they have to take care not only of their shareholders but other constituencies in a society
in which they are established. This endeavour of including other stakeholders of the company within its analysis became well known as corporate social responsibility (CSR).

8. Bibliography


