5. Contract theory and the economics of contract law

Klaus Ulrich Schmolke

In economics, contract theory deals with questions of how economic actors – as maximizers of their own utility – behave in certain contractual arrangements, and the implications of such behavior in terms of efficiency. Put differently and in more normative terms, contract economics is interested in constructing efficient contract designs given the incentives of the parties. Due to imperfect information and other causes, the parties are sometimes not able to reach the optimal solution, that is, the contract maximizing their joint utility, on their own. Then the question arises whether contract law can help to mitigate the ensuing welfare loss.

I. WHY CONTRACTS?

The core notion of contract economics is – unsurprisingly – the contract. As a legal institution, the contract – defined as a voluntary agreement governing the exchange of goods and services between its parties and being enforced by the courts – is well known to lawyers. As a device that facilitates exchange, the importance of the contract and contract law for a modern market economy is seemingly self-evident. When we recall the microeconomics view on exchange (see Chapter 3), this self-evidence may be doubted. In standard microeconomics, welfare-enhancing exchange is typically thought of as taking place as a series of spot transactions whereby goods, services, and money are simultaneously exchanged at a discrete point in time.

A. Exchange in an Ideal World: The Coase Theorem

When perceiving exchange as a costless spot transaction between two omniscient parties who know all the properties of the intended bargain, allocatively efficient outcomes will be achieved. This is also true when legal entitlements are bargained over. Take the following example from
Ronald Coase’s seminal article on ‘The Problem of Social Cost’: A wants to build a factory emitting smoke with harmful effects right next to the residence of B. Regardless of whether we bestow the right to pollute upon A or the entitlement to prohibit pollution upon B, pollution will turn out to be at the same efficient level. As long as an additional unit of pollution is more valuable to A than the harm caused by it is detrimental to B, the additional pollution will take place. If B has the right to prohibit pollution by A, the latter will pay B for allowing the additional pollution and reap the remaining surplus. If A has the right to pollute, B will stop A from adding a further unit of pollution by paying him as soon as the detriment of this further unit to B is larger than its benefit to A.

This is an illustration of the so-called Coase Theorem, which in its strong form comprises two hypotheses: first, the efficiency hypothesis, which is rather close to the initial statement of this section and in the words of Calabresi goes as follows: ‘If people are rational, bargains are costless, and there are no legal impediments to bargains, transactions will . . . occur to the point where bargains can no longer improve the situation; to the point, in short, of optimal resource allocation.’ The second hypothesis, called the invariance hypothesis, states that this optimal resource allocation is independent of the initial allocation of legal entitlements.

Thus, while the parties (A and B) will cooperate in the described way to maximize their aggregate surplus without regard to the initial allocation of legal entitlements, the final distribution of the maximized surplus between A and B depends on this initial allocation (and on the bargaining power of the actors). However, this distribution is no matter of efficiency.

To do Coase justice, he never believed in the existence of a ‘Coasian World’ of zero transaction costs. He rather wanted to show the importance of legal rules in the real world, that is, a world with transaction costs, for achieving allocative efficiency. We will come back to that later on.

More important for now is the notion that, in this ideal Coasian World, contracts as economic or legal institutions are needed for reasons of efficiency only insofar as they make the parties of the exchange continue to accept the bargain struck as valid (and do not take back by force what they have traded away). But at least with regard to the exchange of goods this result may arguably also be achieved in the absence of contracts by a mere (but enforceable) ban of the rule of force.

---

B. Contracts as Commitment and Coordination Devices

The proper value of enforceable contracts to the parties becomes clear when we add two features with regard to the intended exchange: (1) the passage of time, because one party performs in advance or investments have to be made prior to the then simultaneous exchange, and (2) uncertainty about the counterparty’s intentions and future actions. Examples are easily given: Suppose A needs a loan. Bank B will not grant one to A as long as B cannot enforce its repayment. Because of the advance payment by B (the grant of the loan) and the uncertainty with respect to A’s intentions and future actions (will A repay or not?), B will insist on a credible commitment by A. An enforceable credit contract supplies the necessary commitment device. Such a commitment device may be necessary even if the exchange is performed simultaneously: Imagine A is a tailor specializing in bespoke suits. B wants to buy such a suit, but declines to pay in advance, because A could take the money without manufacturing the suit. Thus, A has to manufacture the suit in advance, but if he does so he has to invest in advance and rely on B to pay the agreed price. A cannot sell the suit to someone else, since it is customized. In other words: B made a specific investment when tailoring the suit. If A has no reason to trust B (i.e., if A is uncertain as to B’s intentions and future actions), a credible commitment device is needed: an enforceable contract.

Contracts are also valuable to the parties as a coordination device. Suppose A and B want to exploit a gold mine. A has the money, B the expertise. Thus, they agree on a business contract promising each other to provide the money and the expertise, respectively. What they further need to get things going is a miner working in the mine. A will find a good enough miner if he looks for one. B knows the ‘mining community’. So he will find a really good miner if he does the search. If both A and B look for a miner independently, they end up with two miners. This would leave the business still profitable, but they would make more profits by hiring only one miner. A and B could coordinate their actions by speaking to each other. B could announce: ‘I will hire a miner.’ But that may be ‘cheap talk’. Assigning the task of hiring a worker to B in the enforceable business contract would therefore be even better. With these economic functions in mind, Shavell defines a contract as ‘a specification of the actions that named parties are supposed to take at various times, as a function of the conditions that then obtain’.

---

II. MARKET IMPERFECTIONS AND THE CASE FOR CONTRACT LAW

Given the aforesaid and focusing exclusively on efficiency considerations, one may still wonder why there is more than just one contract law rule stating: ‘All contracts entered into by mutual assent of the parties are recognized by law and enforced by the courts.’ In the real world, however, there are factors and circumstances at play which often hamper the parties of a contract to strike the optimal bargain (market or bargaining failure, see Chapter 3, section V on market failure in general). Thus, even though both parties seek to maximize their joint utility by contracting, the content of the contract finally agreed upon may fall short of this aim. Worse still, the bargaining situation at hand may be affected by parameters that thwart the conclusion of a contract, so that there will be no contract at all. To illustrate the latter case, think of a consumer C who wants to buy a Blu-ray movie. The electronics retailer R – as all other retailers – only sells her products under the condition that the consumer agrees to her fine print. For C, being a very busy person, it is too onerous and costly to invest his precious time in reading the lengthy fine print for the purchase of a mere Blu-ray disc. But if C skips the reading and concludes the purchase anyway he takes the risk that R has ‘hidden’ a clause in the fine print that heavily disadvantages C. If C perceives the costs of taking this risk as higher than the benefit of the Blu-ray movie, he will abstain from purchasing the desired Blu-ray movie.

In such cases contract law may provide efficient means to lessen (the impact of) market imperfections, most notably imperfect information and incentive problems.\(^4\) Even if the law cannot bring about welfare-maximizing contracts (first-best solution), it may at least be able to effect a welfare gain (second-best solution). But the aforesaid only holds true under one crucial condition that those calling for the ‘legal cure’, that is, the intervention of the courts or the legislator, tend to forget: Intervention by (contract) law or otherwise is only indicated if the costs of intervention do not wholly consume the welfare gain accomplished by the intervention, or – even worse – exceed this gain (see also Chapter 6, section I on government failure). When this insight is neglected by the courts or the legislator, contract law may – in terms of efficiency or welfare maximization (on state intervention by means of contract law in the name of distributive justice,

\(^4\) Cf. with regard to the regulation of ‘contractual terms which have not been individually negotiated’ (‘fine print’), e.g., the European Directive 93/13/EEC on unfair terms in consumer contracts, 1993 O.J. (L 95) 29.
see below, section VI) – in itself become part of the problem instead of being part of the solution. In other words, merely leaving things be may be the best solution achievable!

Against the background of the above, we can easily see that contract theory and the economic analysis of contract law go hand in hand. Contract theory provides the insights to evaluate the legal environment of contractual arrangements: Are the legal constraints of the behavior or decision at hand merely an additional cost factor or a means to overcome market failure, and thus to reap additional welfare gains? As for the causes of market or bargaining failure, economists traditionally distinguish certain categories. Among these are (1) externalities, (2) imperfect or asymmetrically distributed information (information asymmetries), (3) market power, (4) imperfect rationality (limits of cognition), and (5) public goods (see Chapter 3, section V, and on imperfect rationality, Chapter 8). When looking for a ‘common denominator’ of all or most of these phenomena, Nobel laureate Oliver E. Williamson and like-minded economists found it in the notion of transaction costs (see Chapter 3, section V.C), thus establishing the so-called field of transaction cost economics. Others, especially those contract theorists inspired by game theory (see Chapter 4), tend to disagree: They see imperfect information of the parties as the main source of their non-optimal contracting. However, this dissension should not be overestimated, since the burden of overcoming imperfect information may be translated into transaction costs.

III. IMPERFECT INFORMATION – A CLOSER LOOK

Therefore, not only those contract theorists focusing on the economics of information, but also most transaction cost economists agree that imperfect information (most notably information asymmetries) is practically the most significant hindrance for efficient contracting. The impact of imperfect information on the efficiency of a contractual bargain becomes clearer if we retrace the following consideration: Pareto optimality of a contractual agreement requires the parties to be able exactly to price each and every parameter constituting the subject matter of the contract,

---

5 Oliver E. Williamson, The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting 17 (The Free Press: New York 1985): ‘This book advances the proposition that the economic institutions of capitalism have the main purpose and effect of economizing on transaction costs.’
including the expected value (see Chapter 2, section I.C.1) of the risks and contingencies attributed to one or the other party by the contract terms. This, in turn, makes it necessary for all parties to the contract to share every piece of information that is relevant to the pricing of these parameters. In reality, however, one party often has private information on the contractual subject matter not (initially) shared by the counterparty. The ignorance of this private information may lead to false prices and, in consequence, to inefficient outcomes.

A. The Problem of Adverse Selection

To show the detrimental impact of information asymmetries, George A. Akerlof, in his seminal article on the market for ‘lemons’, referred to the automobile market by way of example (he himself called it a ‘finger exercise’), which shall be explained in the following: Assume that two types of used car are on the market: good cars, which are of good quality, and bad cars (known in the United States as ‘lemons’), which are of bad quality. The seller who drove the car for a while knows whether his car is a good or a bad one. The potential buyers do not. What they do know is that it is either a good or a bad car. Without further information potential buyers will attach a probability of 0.5 to the car at hand being a good car \((p_g)\) and – correspondingly – a probability of 0.5 to the car being a bad one \((p_b = 1 - p_g)\). Now suppose that the bad car is worth €2,000, while a good car is worth €4,000. In this case, the expected value of the car sold is €3,000 \((= 0.5 \times €2,000 + 0.5 \times €4,000)\) if the potential buyer is risk neutral (see Chapter 2, section I.C.1 on risk preferences).

What will happen under these circumstances? The potential buyers will not be willing to pay more than €3,000 for the used car. The sellers of good cars, however, will not sell below €4,000. Thus, they will be driven out of the market. Only the sellers of bad cars will remain. As a result, even though there is demand for used cars of good quality, their sellers will leave the market. This is already bad, but things get even worse: Suppose the remaining bad cars are not of a single bad quality, but their quality is equally distributed within the range of corresponding values from €1,000 (worst cars) to €3,000 (best of the bad cars). In other words, the expected value of the bad type is €2,000. After the cars of a quality worth more than €3,000 drop out of the market, the potential buyers will observe this decline in overall quality and hence adapt their value estimations to an expected value of the individual car of €2,000. As a consequence, now even the remaining cars of a quality higher than €2,000 leave the market. This game will repeat itself until there are only cars worth €1,000 left in the market (assumed to be of the lowest quality). This ‘race to the bottom’ is
known as the problem of *adverse selection*. To sum up, due to the hidden characteristic of the used car’s quality, the welfare-maximizing sale of used cars of good quality will not be concluded. State intervention (by law) may be warranted if the market does not come up with a solution of its own.

B. Signaling

What the Akerlof model of the adverse selection dynamic does not take into account is that at least some sellers and buyers may have incentives to provide or acquire more information about the quality of the respective car to close the information gap.

1. The concept of signaling

In our example of a market for used cars, the sellers of good cars are supposedly eager to convey their private information of the cars’ quality to the potential buyers. This can be accomplished by *signaling* the quality of their car by way of the terms of their offer. Since all sellers are likely to advertise their cars as ‘good’ ones, this signal has to be credible in such a way that the initially uninformed buyers are capable of distinguishing the quality of the car on the basis of the signal. The sellers of good cars will thus try to choose a signal that is too costly for sellers of bad cars.

In our used car example, the issuance of a warranty by the seller would be a signal with such properties. Suppose the seller promises to repair the car free of charge within a certain time period after the conclusion of the contract, in case a defect should occur. Such a warranty is evidently less costly for the seller of a good car, since the probability of a defect occurring is relatively low, while it is relatively high for bad cars. Thus, a seller of good cars can afford this signal, whereas the signaling costs of such a warranty may be too high for sellers of bad cars. In this case, they will only offer their cars without such a warranty.

2. Signaling that is too costly

However, a signal will not be sent if it does not pay off because the costs of signaling are too high, that is, the benefits gained by signaling are lower than the signaling costs themselves. Joseph E. Stiglitz illustrated this mechanism by referring to an employee being assigned to an assembly line. The employer is not able to determine by herself the productivity of the individual worker. Thus, every worker earns the same wages, whereas

---

the employer would be willing to pay more to workers with high productivity. The worker could signal his high productivity by working really hard, which comes with a cost \( c \). This would result in a pay raise of \( b \). As long as \( c \geq b \), the signal is not worth the effort. Hence, no signaling will take place.

3. Signaling and state intervention

Signaling is a market mechanism to overcome the problem of adverse selection. One may also think of alternative state intervention to cope with the welfare loss associated with this problem, especially when signaling may be too costly. For example, in the US, several states passed so-called ‘Used Car Lemon Laws’ which require the dealers of used cars to give consumers a written warranty that depends on the mileage of the car. Sellers of really bad cars then drop out of the market, since for them this warranty obligation is too costly. The main purpose of such laws, however, is not to overcome asymmetric information, but to protect consumers from potentially dangerous cars.

Such laws typically only provide minimum standards as to the quality of a product or the abilities of a service provider. Therefore, informed sellers or service providers may still have incentives to send distinguishing signals showing the superiority of their products or services. For example, Articles 3 and 5 of the European Directive on certain aspects of the sale of consumer goods\(^7\) provide for a mandatory liability of sellers in case the sold goods are not in conformity with the contract and the lack of conformity becomes apparent within two years as from delivery of the goods. A seller who is convinced of the superior quality of her goods may therefore signal this information by voluntarily prolonging this liability beyond the mandatory two-year period.

C. Screening

1. Screening as a means to gain information about potential counterparties

While Michael Spence contributed the theory of signaling to the analyses of markets affected by information asymmetries,\(^8\) Joseph Stiglitz was among the first to propose a theory on screening.\(^9\) Screening describes the

\(^7\) Directive 1999/44/EC, 1999 O.J. (L 171) 12.

\(^8\) Michael Spence, Job Market Signaling, 87 Q. J. Econ. 355 (1973).

inverse scenario to signaling: Here it is the uninformed party who adopts costly means to get more information about the abilities of the potential counterparty or the quality of her goods, respectively. The uninformed party ‘screens’ the potential counterparties.

Rothschild and Stiglitz illustrate this mechanism by referring to the insurance market where insurance companies sell insurance contracts to individuals. These individuals, being risk-averse, purchase insurance to smooth out their expected income losses by accidents occurring with probability \( p \). Since their marginal utility of income declines (see Chapter 2, section I.C.1), they are willing to pay an insurance premium that is slightly higher than the expected income loss due to accidents. Now assume there are two types of individuals: those with a low probability of accidents \( p_l \) (‘good type’) and those with a high probability of accidents \( p_h \) (‘bad type’).

From the insurance companies’ point of view, the question now arising is which contracts to offer, and to whom, in order to maximize their expected profits.

However, while individuals know to which type they belong, the insurance companies do not know – by assumption – the accident probabilities of the individuals and are therefore not able to discriminate between good and bad types without further information. Thus, they will search for screening devices in order to get this information and, in turn, be able to adjust the contract terms to the individual’s propensity to incur accidents.

2. Self-selection by contract
One technique of screening is to force the counterparties, that is, the insurance customers, to reveal private information by the terms of a pricing scheme included in the contract offer. This mechanism is called self-selection. In our insurance market example, such a self-selection device may look as follows: The insurance companies offer two different kinds of contract. The first kind offers full coverage and comes at a relatively high price (insurance premium). The second kind provides for retention of a certain amount in case an accident occurs (partial coverage), which is reflected in a lower insurance premium. For customers of the good type, it may be more attractive to choose the contract that only partially covers their losses, since their probability of incurring an accident is low. At the

Economic Sciences was awarded to Akerlof, Spence, and Stiglitz, ‘for their analyses of markets with asymmetric information’.

same time, the bad-type customers may be better off with full coverage due to their higher propensity to be involved in accidents.

The numerical example in Table 5.1 illustrates the aforesaid: The bad type has a probability $p_b$ of 20 percent to incur an accident at a cost of 100, which translates into an expected loss of 20. Thus, full coverage at a premium of 21 leaves her with a certain loss of 21, that is, the premium payment $(0.2 \times 100 - 0.2 \times 100 - 21)$, whereas a partial coverage of 80 percent at a premium of 18 amounts to an expected loss of 22 $(0.8 \times 0.8 \times 100 - 0.8 \times 100 - 18)$. She therefore prefers full coverage. On the other hand, the good type having a probability $p_g$ of 10 percent to incur an accident at a cost of 100 has an expected loss of 20 when choosing only partial coverage $(0.8 \times 0.1 \times 100 - 0.1 \times 100 - 18)$, while full coverage at a premium of 21 results in a certain loss of 21 $(0.1 \times 100 - 0.1 \times 100 - 21)$. Thus, she prefers the contract with partial coverage as long as the utility gain achieved by excluding the residual risk is $< 1$.

Even though both types incur pecuniary losses, they nevertheless conclude an insurance contract as long as they have a utility gain due to their supposed risk aversion. However, compared to a state of perfect information, there is a utility loss to the good type, who prefers to obtain full coverage insurance due to his risk aversion!

It goes without saying that screening by offering self-selecting contracts only works under the assumption that the uninformed party (the insurance company) comes up with a pricing scheme that separates the different customer types because all types are better off with the contract aimed at their respective type (separating equilibrium).\textsuperscript{11} In contrast, a pricing scheme does fail as a screening device, when different types of

\begin{table}[h!]
\centering
\caption{Self-selection in the insurance market (separating equilibrium)}
\begin{tabular}{|l|c|c|c|c|c|c|}
\hline
Type & Probability of accident & Costs of accident & Premium for full coverage & Premium for 80\% coverage & Certain loss in the case of full coverage & Expected loss in the case of 80\% coverage \\
\hline
Good & 10\% & 100 & 21 & 18 & 21 & 20 \\
Bad & 20\% & 100 & 21 & 18 & 21 & 22 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{11} Michael Rothschild and Joseph E. Stiglitz, \textit{Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information}, 90 Q. J. Econ. 629, 633 (1976) define: ‘Equilibrium in a competitive insurance market is a set of contracts such that, when customers choose contracts to maximize expected utility, (i) no contract in the equilibrium set makes negative expected profits; and (ii) there is no contract outside the equilibrium set that, if offered, will make a nonnegative profit.’
counterparties choose the same kind of contract, because it is the utility-maximizing choice for both (or all) types (pooling equilibrium). In consequence, no private information is revealed to the uninformed offeror. To illustrate this, take the numerical example shown in Table 5.2.

The numerical example in Table 5.2 differs from the one given in Table 5.1 only insofar as the premium for partial coverage is 19 instead of 18: For the bad type, nothing changes. The certain loss of 21 when choosing the contract with full coverage is now even more attractive in comparison to the contract with partial coverage, which for her results in an expected loss of 23 ($0.8 \times 0.2 \times 100 - 0.2 \times 100 - 19$). For the good type, however, the situation differs from the former scenario: Choosing the partial coverage now results in an expected loss of 21 ($0.8 \times 0.1 \times 100 - 0.1 \times 100 - 19$), while full coverage at a premium of 21 results in a certain loss of 21. Thus, under these conditions, the good type also prefers the contract with full coverage, since he prefers – by assumption – to eliminate the residual risk that is associated with the partial coverage.

When such a pooling equilibrium occurs, the good types pay a rather high price for the respective service or good by which they typically cross-subsidize the bad types who pay too low a price given their risk characteristics. Welfare losses may occur because the pooling contract provides suboptimal incentives for the bad types to avoid the respective risk and, furthermore, because for some good types insurance is too expensive. As a consequence, the latter abstain from buying insurance even though they would prefer to have it.

3. Reduction of screening costs by state intervention

The costs of screening may be (substantially) reduced by state intervention. For example, section 21(1) of the Australian Insurance Contracts Act 1984 (IC Act) states that:

[A]n insured has a duty to disclose to the insurer, before the relevant contract of insurance is entered into, every matter that is known to the insured, being a matter that . . . the insured knows to be a matter relevant to the decision of the

<table>
<thead>
<tr>
<th>Type</th>
<th>Probability of accident</th>
<th>Costs of accident</th>
<th>Premium for full coverage</th>
<th>Premium for 80% coverage</th>
<th>Certain loss in the case of full coverage</th>
<th>Expected loss in the case of 80% coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>10%</td>
<td>100</td>
<td>21</td>
<td>19</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Bad</td>
<td>20%</td>
<td>100</td>
<td>21</td>
<td>19</td>
<td>21</td>
<td>23</td>
</tr>
</tbody>
</table>
insurer whether to accept the risk and, if so, on what terms; or . . . a reason-
able person in the circumstances could be expected to know to be a matter so
relevant.

If the insured fails to comply with this duty of disclosure or makes a
misrepresentation in a non-fraudulent manner, the claim of the insured
against the insurer is, according to section 28 of the IC Act, reduced ‘to
the amount that would place the insurer in a position in which the insurer
would have been if the failure had not occurred or the misrepresentation
had not been made’.

D. Market Power and Asymmetric Information

Monopoly or market power is known as a classic cause of market failure.
As has already been shown, a monopolist reduces the supply of his goods
to a level below the optimal to maximize his producer surplus, causing a
welfare-decreasing ‘dead weight loss’ (see Chapter 3, section V.A). This
concept has been transferred to the realm of contracts in order to explain
inefficient contract terms. In a paper from the 1940s pioneering the idea,
the German emigré Friedrich Kessler wrote that standard contracts were
typically used by enterprises with strong bargaining power. They offered
these kinds of contracts to the ‘weaker’ counterparties who were only left
to accept the standard terms or abstain from the bargain. Such standard-
ized contracts were therefore ‘contracts of adhesion; they are à prendre ou
à laisser’.\footnote{12 Friedrich Kessler, \textit{Contracts of Adhesion: Some Thoughts about the Freedom of Contracts}, 43 COLUM. L. REV. 629, 632 (1943).} Hence, the argument continues, the law has to step in.\footnote{13 Cf., for a real-life example, the decision \textit{Macaulay v. Schroeder Publishing Co Ltd}, [1974] 1 W.L.R. 1308 (H.L.); on this decision, see Michael Trebilcock, \textit{The Doctrine of Inequality of Bargaining Power}, 26 U. TORONTO L.J. 359 (1976).}

Economists will only be bothered by bargaining power as far as it causes
a welfare loss. In case it does, the lawyers will ask whether and how the
legal environment may reduce this welfare loss. We can only answer this
question when we remind ourselves that monopoly power alone is not
the cause of the welfare loss, described elsewhere as ‘deadweight loss’.
If the monopolist were able to discriminate price perfectly, she would reap
the maximum surplus without any welfare loss. Thus, the welfare loss
occurring in cases of monopoly power is ultimately caused by the imper-
fect information of the monopolist about the individual reservation prices
of her counterparties (see Chapter 3, section V.B).

Drawing on this insight, Richard Craswell rightly pointed out that a
monopolist may have the incentive to charge a high price to reap a high profit. But she will not have an incentive to do this indirectly by demanding inefficient contract terms, when this can be done straightforwardly by raising the monetary price of the good or service offered. This is so, because an inefficient contract term by definition does not yield as much profit for the monopolist as it causes costs for the counterparties. If we therefore observe inefficient terms in contracts offered by a monopolist, these terms may be used to discriminate between different groups of counterparties, that is, as a screening device. To intervene legally by declaring such terms null and void would only cause the monopolist to demand the same price from all potential counterparties. However, the welfare consequences of such a strategy change are far from clear. Thus, Craswell concludes, ‘the focus on monopoly power is really a red herring where contract terms are concerned. If courts and laypeople tend to associate inefficient terms with monopolies, it’s probably because monopoly is the only form of market failure that courts and laypeople are familiar with.’

IV. COGNITIVE LIMITS AND COGNITIVE FAILURES

A. The Limits of Cognition as a Source of Imperfect Information

Like the Coase Theorem, contract theory traditionally assumes rationality of the contracting parties in line with the model of the homo economicus (see Chapter 2, section I.C). The assumption of rationality is crucial to the case for freedom of contract and against state intervention. The parties can only be relied upon to maximize their joint welfare if they are capable of correctly calculating the (subjective) expected utility of their options (see above, section III).

However, there is overwhelming empirical evidence that this assumption of (formal) rationality does not hold. To begin with, it cannot be denied that human decision-makers only have a limited capacity to search for, absorb, and compute information. This bounded rationality of human actors is the starting point for a more realistic model of decision-making, which its inventor Herbert Simon dubbed ‘satisficing’. At its core

---

lies the idea that the human actor does not aim for the optimal, that is, utility-maximizing choice, since this is too costly due to his limited capacities to gather and assess information. Thus, he contents himself with choosing a satisfactory option, thereby saving ‘choosing costs’. This is, on balance, the rational course of action for a boundedly rational actor. Even worse, human decision-makers are not only boundedly rational, but also prone to making systematic errors, especially when deciding under uncertainty (see Chapter 8, sections III.B and C for further details on bounded rationality and systematic decision-making errors of human actors). Melvin A. Eisenberg was supposedly the first scholar to perform an indepth analysis on what these findings imply for the freedom of contract and its limitation by contract law. His main argument is that contracting parties misperceive or miscalculate contractually relevant risks as well as the utility of contractual terms allocating these risks. As a consequence, the conclusion of the contract and its terms are incorrectly priced by one or both parties. Therefore, they do not opt for the contract that maximizes their joint utility. The ensuing welfare loss may, in turn, justify state intervention by contract law.

With regard to bounded rationality, contract theorists soon incorporated the insights of Herbert Simon into their analyses without abandoning the traditional rational choice model. They did this by recognizing ‘costs of reading’ or ‘costs of understanding’ as a further type of transaction costs. From this perspective, bounded rationality and, more generally, the limits of cognition of one or both parties to a contract may be perceived as a problem of imperfect information. Most economists nowadays, however, especially the proponents of behavioral economics (see Chapter 8 for details) classify cognitive limitations and errors in decision-making caused by biases and the application of heuristics as an independent category of market or bargaining failure (see above, section II).

B. State Intervention by Paternalist Contract Law

The discovery by psychologists and experimental economists of ever new ‘anomalies’, that is, deviations of human decision-making behavior from the rational choice model, inspired legal scholars to propose a policy of legal paternalism: The erring and rationally deficient human actor shall be protected by the law from his own false decisions. These issues will be discussed in general and at length in Chapter 8. For the purposes of this

---

chapter, it suffices to give a concluding example of (paternalist) contract law addressing the limits of cognition of the (prospective) parties.

Under Article 18 of the European Directive on mortgage credits, the creditor is not only obliged to assess the creditworthiness of the consumer (debtor). He also has the duty to make the credit available to the consumer only under the condition that ‘the result of the creditworthiness assessment indicates that the obligations resulting from the credit agreement are likely to be met in the manner required under that agreement’. Given the typical risk structure of a credit contract and the incentives of the parties to such a contract, this provision seems odd. If the creditor is ready to take the default risk, why should she be obliged to refrain from concluding the credit contract which the consumer obviously wants as well?

For an answer to this question, one has to take into account that the European legislator adopted this provision (among others) ‘to prevent household over-indebtedness’. Thus, the European legislator is skeptical when it comes to the ability of consumers to assess correctly their risk of defaulting on a mortgage credit contract. Consumers therefore shall be kept from concluding credit contracts which – unbeknown to themselves – overburden them.

V. INCENTIVE PROBLEMS AND IMPERFECT INFORMATION AFTER CONTRACT CONCLUSION

A. Moral Hazard

1. The phenomenon of moral hazard

Information asymmetries not only occur at the time the contract is concluded, but also thereafter when it is executed. Such information asymmetries may lead to welfare-decreasing incentive problems, the most famous of which is known as moral hazard. The term purportedly has its origins in the insurance industry and can be easily depicted in this very context by referring and extending the example given above (see above, section III.C.2 with Table 5.1): Suppose the customer who is more prone to having accidents (bad type) buys insurance with full coverage from an insurance company. Due to his probability of 20 percent of being involved in an accident with costs of 100 ex ante, the bad type has to pay an insurance premium of 20 plus administration costs and entrepreneurial profit.

of the insurance company. The problem now is that, due to the full insurance, the bad type has less than optimal incentives to take precautions for avoiding the risk. In consequence, the probability of an accident actually occurring rises.

The insurer would want to take care of this problem by imposing certain duties of conduct on the insured to restrain the scope of his actions. An insurance company insuring car accidents could, for instance, contractually oblige the insured to take care of the lights and brakes of her car. But this mechanism will not help as far as the actions of the insured cannot be observed by the insurance company (hidden action). For example, the insurance company typically does not know whether the insured drives aggressively and inattentively or safely and foresightedly.

2. Agency contracts: The principal–agent conflict

The incentive problems just described may also occur with regard to agency contracts, such as employment or service contracts, or in similar arrangements like the appointment as a corporate director: The agent acts on behalf of his principal, that is, the employer or the corporation (and the shareholders as its owners), and is therefore obliged by contract or appointment to act in the best interest of the principal. The principal, in turn, provides ‘payoff rules’ or a ‘fee schedule’ that lay out the pay the agent receives depending on her actions. Therefore, the agent’s actions affect both the welfare of the principal as well as the welfare of the agent.

By assumption, though, the agent is as much a maximizer of her own utility as the principal. Thus, incentive problems arise under the condition of asymmetric information, that is, uncertainty of the principal about the actions of the agent. Arrow distinguished two main categories of such unequal information distribution.\(^{18}\) Firstly, information asymmetries occur where the actions of the agent\(^{19}\) cannot be observed and cannot be precisely inferred from the outcome either (hidden action). Secondly, the agent has private information because she made an observation the principal has not made; thus, the principal cannot observe whether the agent uses this information in the best interest of the principal or not (hidden information).

To illustrate the problem, think of a managing director of a company doing business on behalf of the shareholders of the company. The managing director of the company is the agent of the shareholders (principal) and

---


\(^{19}\) Id., at 3: ‘The most typical hidden action is the effort of the agent.’
therefore obliged to act in the best interest of the company and its owners. Now suppose the managing director runs a company that manufactures and sells beverages (the L-Company, henceforth L). L buys large amounts of Cola syrup from the Coca-Cola Company which is sold at L’s stores, but L is investigating a change of supplier due to the high costs of the syrup. The managing director now acquires the information that the National Pepsi-Cola Company, which possesses a secret formula and trademark, is bankrupt. Hence, formula and trademark are offered for sale to the director in his capacity as an agent of L. Instead of buying the formula and trademark for L and thereby seizing the corporate opportunity for his principal, the director captures it for himself without the knowledge of the board of directors of L, sets up a new company wholly owned by him and further on sells the Pepsi syrup to L with a profit.20

3. Solution strategies: Monitoring and alignment of interests by incentive pay

What to do about these incentive problems? Economic theory has put forward mainly two different strategies to address these issues, which may be combined. The first strategy is monitoring: The principal may control the actions of the agent by requesting reports, paying her unheralded visits, or assigning a different kind of agent with the task of controlling and supervising the agent. Thus, by monitoring, the principal intends to uncover the formerly hidden action or hidden information. The information gap is thus narrowed. The range of occasions on which the agent may deviate from the best interests of the principal is reduced, as is the range of actions deviating from the principal’s best interests, should such an occasion occur.

For example, a company’s board of directors regularly consists not only of managing directors who run the company, but also of independent directors supervising and controlling the actions of the former. In German stock corporation law, there are actually two boards: A managing board (Vorstand), which is controlled and monitored by a supervisory board (Aufsichtsrat).

The second strategy to address the incentive problems of an agent acting on behalf of a principal is the alignment of interests. Here, the agent is not ‘forced’ to comply with the interests of the principal by uncovering (potential) hidden actions and thereby reducing the information gap between principal and agent. Instead, the agent is ‘induced’ to act in the best

20 Cf. the Delaware case Guth v. Loft, Inc., 5 A.2d 503 (Del. 1939), the facts of which have been simplified here.
interest of her principal by making it in her own best interest to do so. The most common measure to achieve this goal is incentive pay. Think again of a company’s managing director. Typically, his remuneration not only consists of a fixed salary, but also of a proportion of variable pay. This variable pay is typically tied to the company’s performance. At first, stock options were deemed to be the ideal tool to align the interests of executives with the best interests of their firm. Nowadays, too large a proportion of variable pay and especially the provision of stock options are considered harmful by many since this induces risk-seeking behavior. In the aftermath of the 2008 financial crisis, a lively debate has been going on as to how bankers and managers should be properly incentivized in order to induce them to act in the long-term interest of their firms.

However, implementing these strategies comes at a cost. Such costs are called *agency costs*, which also comprise the original welfare loss caused by using suboptimally incentivized agents.

4. **State intervention by law**

The state tries to lower agency costs through legal intervention. Company law, for example, provides monitoring structures, which experience has shown to be useful in restraining the management from acting to the detriment of the principal. One such device is the mandatory supervisory board that German stock corporation law stipulates. A means of interest alignment would be to hold the agent liable for not acting in the best interest of her principal. Recall the managing director of the L-Company (see section V.A.2): US corporation law recognizes the so-called ‘corporate opportunities doctrine’, whereby directors and officers of a company are held accountable for misappropriating the business opportunities of their company. The idea is that due to these liability rules it is no longer worthwhile for a director or officer to diverge from the company’s, that is, the shareholders’, best interest *ex ante*. Finally, it seems noteworthy that after the financial crisis of 2008 many company laws stipulated rules for executive remuneration to prevent a compensation scheme from setting perverse incentives to the detriment of the company and its shareholders.

B. **Long-Term Contracts, Opportunism, and the Cost Trade-Off of the Parties**

1. **The limits – transaction costs and bounded rationality**

The aforementioned problems are typically exacerbated when the parties enter long-term contracts, sometimes also called ‘relational contracts’. As their name already indicates, they are intended to govern the conduct of the parties for a longer, sometimes indeterminate duration. Such contracts
Economic methods for lawyers

are concluded when their subjects require considerable specific investments and consecutive transactions. The subject matter covered by such contracts is typically complex. This complexity, as well as the need to project the contractual relationship into the contingent future, leads to a high level of uncertainty. In short, relational contracts are concluded ‘where transactions (1) are recurrent, (2) entail idiosyncratic investment, and (3) are executed under greater uncertainty’.21

Due to the bounded rationality of the parties, that is, their bounded capacity to absorb and compute information (see above, section IV.A.), those contracts are necessarily incomplete to a considerable extent: Even though the parties recognize the need to modify and specify their contractual relationship in the future, they do not negotiate the respective terms ex ante due to the prohibitive costs of such negotiations. As a consequence, even important aspects of the contractual arrangement may not be governed by a precise and substantiated program of duties.

Against this background of a necessarily incomplete contract, the parties expect to renegotiate and adapt the original allocation of risks when future events so demand. The termination of the contract is typically no feasible alternative path of action since it would devalue the specific investments already made by the parties. This is true at least as long as those investments have not been redeemed. Hence, the parties are ‘locked in’ the contractual relationship.

2. The threat – opportunism ex post

These features of a long-term contractual relationship would be less of a problem if the parties agreed upon a rule whereby they commit themselves to filling the gaps in their contract sequentially in a way that maximizes their joint utility.22 However, this efficient path of mutual cooperation is threatened by the parties’ opportunism. Williamson defines such opportunism as ‘self-interest seeking with guile’.23 It may cause the parties to act in ways that further their own advantage, but frustrate outcomes


22 Cf. Oliver E. Williamson, The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting 48 (New York: The Free Press 1985): ‘[P]roblems during contract execution could be avoided by ex ante insistence upon a general clause of the following kind: I agree candidly to disclose all relevant information and thereafter to propose and cooperate in joint profit-maximizing courses of action during the contract execution interval, the benefits of which gains will be divided without dispute according to the sharing ration herein provided.’

23 Id., at 47.
that are superior from a welfare perspective. Since cooperation during the contract execution period is generally for the mutual benefit of both parties, opportunistic behavior only emerges where the expected short-term gains of defection are higher than the expected long-term gains of cooperation. The occurrence of such ‘high-value opportunism’ obviously depends on the discount rate the parties apply to future gains.

3. The trade-off – transaction costs versus governance costs
Given the aforesaid, the parties of a long-term contract have to accomplish a twofold task: On the one hand, they aim to minimize the transaction costs ex ante, not least due to their bounded rationality. On the other hand, the parties strive to implement safeguards against ex post opportunism. Both goals have to be traded off, however, since the reduction of transaction costs ex ante typically increases the hazard of opportunism or, alternatively, the governance costs necessary to hold opportunism at bay ex post.

4. The law – part of the solution or part of the problem?
What part does the law have to play in this long-term contract scenario? At its best, the law can be an effective means to lower the sum of the costs of opportunism and the costs to avoid opportunism, which comprise the transaction costs ex ante as well as the governance costs ex post. Legal default rules may help to further this aim (at least) in three different ways: Firstly, they inform the parties about the issues and aspects which are of significance to their contractual arrangement. Secondly, they relieve the parties of the burden to negotiate each and every aspect of their contractual relationship, thus allowing them to focus on the important issues. Thirdly, default rules fill the gaps the parties left after concluding their negotiations.

Donald J. Smythe distinguishes (at least) five beneficial effects that contract law, as a means of reducing the likelihood of ex post opportunism, may have on long-term contractual relations: Such a law ‘will (1) increase the longevity of relational contracts, (2) improve the cooperativeness of relational contracts, (3) increase the size of investments under relational contracts, (4) decrease expenditures on special arbitration procedures, and (5) decrease the volume of the transactions conducted under less efficient governance structures . . . ’25

So far, legal intervention in long-term contractual arrangements sounds like a success story. But the law can also become part of the problem: For example, it is well known that legal rules and arrangements that are intended to protect one party from opportunistic behavior of the other party may under unfortunate circumstances be misused by the protected party for a different kind of opportunism. This can be perceived as a manifestation of moral hazard: The protective legal rule may work like an insurance insulating the protected party from the adverse consequences of defection. This, in turn, may stifle his incentives to uphold the cooperative equilibrium of the contractual arrangement. Furthermore, when contractual partners take their quarrels to court, there is always the possibility that the judge – being an outsider to the parties’ relationship – may misinterpret a certain course of action. What he thinks of as opportunistic behavior may actually be an appropriate means of retaliation that pushes the defecting counterparty back on the path of cooperation and vice versa.

Because of these dangers of dysfunctional or misapplied law, some legal scholars argue for restraining the role of law (in favor of social norms) as a means of stabilizing a welfare-enhancing cooperative equilibrium in long-term contractual relations. The law should – so the argument goes – only be employed for a ‘large strike’, that is to impose strict sanctions in the case of ‘large scale opportunism’.

5. Application – minority protection in the close corporation
A prominent field where legal rules are applied to abate ex post opportunism in long-term contractual arrangements is the protection of minority shareholders of close corporations. These rules address the following problem: close corporations own the features of long-term contracts, as described above. To smooth the decision-making process with regard to the conduct of the underlying business, the majority principle is generally the legal default rule. This may become dangerous for a shareholder who finds himself in the minority faction later on. The majority might be tempted to abandon the course of cooperation that maximizes the joint profits of the shareholders in favor of a more opportunistic course of action that disproportionately benefits the majority faction. Since there is no liquid market for memberships in a close corporation, the exit by sale of shares is no realistic option for the aggrieved minority. Hence, the minority shareholder is ‘locked in’. Company law provides certain protective measures to help the minority in distress: In UK company law, for example, according to section 994 of the Companies Act 2006, a member of a company may apply to the court by petition for an order on the grounds ‘that the company’s affairs are being or have been conducted in a manner that is unfairly prejudicial to the interests . . . of some part
of its members (including at least himself). To take another example, in German close corporation law a shareholder is entitled to leave the company ‘for good cause’, which also captures cases of a severe disruption of the members’ relationship. In this case, the departing shareholder is entitled to ‘full compensation’ for the loss of his membership. More problematically, this right to ‘full compensation’ is also granted to departing members who have been expelled by the company, that is, the majority of the shareholders, for good cause. However, this statutory default provision may backfire, since it insures the shareholder even when he himself is defecting (moral hazard). Therefore, the parties often provide for a deduction from the full compensation in the articles of association to incentivize cooperative behavior.

VI. THE UNEASY CASE FOR ‘DISTRIBUTIVE JUSTICE’ IN CONTRACT LAW

The economic approach to contracts and contract law aims at maximizing social welfare, which in the context of contracts often means maximizing the welfare of the (prospective) parties to the contract. Hence, contract theorists are typically unconcerned about how the welfare gains reaped by the conclusion and execution of a contract are distributed among the parties. From a legal perspective, this seems odd at first glance, since the notion of ‘distributive justice’ plays a rather prominent role in contract law debates among legal scholars, especially when the law of consumer contracts is concerned. This raises the question: Is common ignorance with regard to distributive issues a serious flaw in the economic approach to contract design and contract law, or is there a convincing justification for focusing on issues of welfare maximization?

While economists are not oblivious to the necessity of redistributing resources by state intervention (see Chapter 1, section IV.D.1 with regard to the constitutional boundaries of exclusively pursuing welfare maximization in terms of efficiency gains), most of them agree that private law is a comparatively clumsy, unreliable, and inefficient means to redistribute resources. With regard to contract law in particular, economists further some noteworthy arguments for this viewpoint, which should at least be known by a contract lawyer, even if she dissents from their conclusion: The party burdened by the law with the costs of disadvantageous contract terms, for example, the entrepreneur or professional vis-à-vis consumers, regularly reacts to such intervention by raising the price for her product or service. As a result, the consumer who is supposed to benefit from the legal intervention bears its costs in the end. The distributive effect
between professional and consumer parties is naught. To illustrate this mechanism, let us draw again on the European Directive on the sale of consumer goods, which provides for a mandatory liability of sellers in the case where the sold goods are not in conformity with the contract and the lack of conformity becomes apparent within two years as from delivery of the goods (see above, section III.B.3). This liability comes at a cost for the seller, who therefore calculates the expected costs for each good sold and raises the price of the good accordingly. The consumer is now insured, which she appreciates under the assumption of risk aversion, but pays the price for this insurance.

Now let us assume that the party burdened by mandatory contract terms cannot pass on the whole costs of such terms, but only part of the costs. In this case, the intentionally burdened party has indeed to bear the costs, or part of the costs, associated with the legal rule. But does it benefit the counterparty? It has rightly been remarked that the unwillingness of the counterparty to bear the whole costs of such mandatory contract terms indicates that he does not appreciate the terms that much, at least not enough to pay for them. Craswell generalized this insight with regard to consumer contracts: ‘Paradoxical as it may seem, the rules whose costs are most heavily passed on are also the rules that will benefit consumers the most.’

Some scholars believe that the unwillingness of contractual parties to pay for certain terms is the very reason why the social planner (the legislator) makes them compulsory. However, those who try to justify interventions that cause such effects (solely) by reference to the notion of ‘distributive justice’ have a hard time. The crucial point seems to be that the social planner assumes that the unwillingness to pay for the term is based on a flawed decision-making process, that is, because it is caused by limits of cognition (see above, section IV): If the consumer correctly understood the impact of the mandatory term, he would pay for it.

Mandatory contract terms may have distributive effects of a different kind: Suppose the preferences of the demand side (consumers) are heterogeneous with regard to the mandatory contract term. As a consequence, some consumers profit from the term in question, while others incur losses. In this scenario, the redistribution does not take place between the consumers and their professional counterparties, but among different subgroups of consumers. To illustrate this, let us draw again on the mandatory liability of sellers provided for in the European Directive on

---

the sale of consumer goods (see above, section III.B.3). This mandatory liability functions as an insurance against the risk that the good does not conform to the contract, the specifications of which are uniform and establish a minimum standard that cannot be waived. The seller is thus compelled to offer a pooling contract with regard to this liability. This, in turn, may lead to a cross-subsidization of high-risk consumers by low-risk consumers (redistribution). Some low-risk consumers may even drop out of the market and thereby cause a welfare loss (see above, section III.C.2).

VII. CONCLUSION

Contract theory provides an analytic apparatus to identify the welfare implications of various bargaining scenarios and contractual arrangements. It has been especially successful in detecting sources and conditions of market failure where the parties fail to conclude the optimal, that is, welfare-maximizing, contract. The most important source of such market failures is imperfect information of the parties and consequent incentive problems. Whenever a market failure occurs, the law, most notably contract law, may be used to lessen the ensuing welfare loss (second-best solution). However, legal intervention is only warranted if the costs of intervention are lower than the welfare gain accomplished by the intervention.

The concepts of contract theory can be applied to very different bargaining settings. The examples given in this chapter covered the sale of goods, insurance contracts, labor contracts, and the governance of companies. But the scope of contract theory is much greater. Thus, its insights may, for example, also help to conceptualize and solve issues of public choice (see Chapter 6, section III.C) or (public) international law.27

FURTHER READING


