CHAPTER 4

CONTRACT RIGHTS AND REMEDIES

§4.1 The Process of Exchange and the Economic Roles of Contract Law

The last chapter emphasized the importance of voluntary exchanges in moving resources from less to more valuable uses, and it noted various obstacles to value-maximizing exchanges; but the process of exchange, once terms are agreed on, was assumed to operate reliably without legal intervention. This is strictly true, however, only when both parties perform their obligations under the contract simultaneously; and that is rare. When the simultaneity condition does not hold, two dangers to the process of exchange arise—opportunism and unforeseen contingencies—for which the law offers remedies.

A hires B to build a house, payment due on completion. While the house is being built and before any payment has been made, B is at A’s mercy, for he would find it difficult (especially if A owns the land that the house is being built on!) to sell the house to anyone else if A decided not to pay for it. So in the absence of a legally enforceable contract, A could force B to reduce his price after construction was under way. (Since contract law, like all social institutions, does not work perfectly, one is not surprised to find that builders insist on progress payments—and not because their customers are their lowest-cost lenders.) After the house is constructed and A pays B, their roles are reversed. A is now at B’s mercy. For the construction of the house is not really the end of B’s performance but the beginning. A is counting on receiving a stream of services from the house for many years. If B has built the house shoddily, A’s expectation will be disappointed.

Like the disincentive to cultivate land in a world without property rights, the problem of contractual opportunism arises from the sequential character of economic activity. If sowing and reaping were simultaneous, the need for recognition of property rights in land (as distinct from harvested crops) would be less urgent. If contractual exchanges were simultaneous, the need for legal protection of contract rights would be less urgent. Since they are not, the absence of legally enforceable rights would bias investment toward economic activities that could be completed in


§4.1 1. This implies, and we find, that in primitive societies, where contractual exchanges tend to be simultaneous, the law of contracts, in contrast to many other areas of primitive law, is rudimentary. In particular, executory contracts—contracts where neither party has yet begun to perform his contractual undertaking—are not enforced. Richard A. Posner, The Economics of Justice 188-184 (1981).
a short time; and this would reduce efficiency. A wants to sell his cow. There are two bidders, B and C. The cow is worth $50 to B and $100 to C (and only $30 to A), so efficiency requires that the cow be sold to C rather than B. But B has $50 cash in hand while C cannot obtain any cash for a week. C promises to pay $75 to A in a week, and let us assume that this $25 premium would fully compensate A for the costs, in the event of default, of bringing a suit for damages or for return of the cow, discounted by the risk of default — if the law made C’s promise to A enforceable. But if the law does not enforce such promises, A may decide that, since C may fail to raise the money and B in the interim lose interest in the transaction, he is better off selling the cow to B now. If he does, the law’s failure to provide a remedy if C breaks his promise will have induced a misallocation of resources. B might resell the cow to C later, but this would involve an additional transaction cost.

Now suppose that D offers to sell a shirt for $5 and his competitor, E, offers one for $6 which he claims (truthfully) will last three times as long as D’s shirt and is therefore a better value. Because the difference is not apparent on casual inspection or handling, E would be willing to guarantee the superior durability of his shirt; but if his promise is not legally enforceable, consumers may doubt the honesty of his claim and buy D’s shirt instead, again a suboptimum result.

The system of voluntary exchange would not break down completely without a law of contracts. There are contracts in societies that have no formal law-enforcement machinery and, as we’ll see later in this chapter, contracts between nations that recognize no legal constraints on their sovereignty. Someone known not to perform his side of bargains will find it difficult to find people willing to make exchanges with him in the future. Hence there might be more explicit definition, either in writing or by reference to custom, of the undertakings of the parties to an exchange than if contracts were enforceable. Transacting parties would be eager to minimize misunderstandings that might give rise to charges of bad faith, since someone against whom such charges were lodged would find it difficult to get people to make exchanges with him in the future.

But a system of unenforceable contracts would not be efficient. Apart from the costs of credit bureaus and security deposits (especially since return of the deposit could not be compelled), self-protection would often fail. Although someone who was contemplating breaking his contract would consider the costs to him of thereby reducing the willingness of other people to make contracts with him in the future, the benefits from breach might exceed those costs. He might be very old; or (a related point) the particular contract might dwarf all future contracts that he expected to make; or he might not be dependent on making contracts but instead be able to function in the future on a cash-and-carry basis.

The basic aim of contract law (as recognized since Hobbes’s day) is to deter people from behaving opportunistically toward their contracting parties, in order to encourage the optimal timing of economic activity and (the same point) obviate costly self-protective measures. But it is not always obvious when a party is behaving opportunistically. Suppose A hires B to paint his portrait “to A’s satisfaction.” B paints a portrait that connoisseurs of portraiture admire, although not enough to buy it themselves at the contract price. A rejects the portrait and refuses to give any reason for the rejection. If the rejection is not made in good faith, A will be held to have broken the contract. Good-faith performance — which means in this context refraining from taking advantage of the vulnerabilities created by the sequential

2. Thomas Hobbes, Leviathan 70-71 (1914 [1651]).
character of contractual performance—is an implied term of every contract. No one would voluntarily place himself at the mercy of the other party, so it is reasonable to assume that had the parties thought about the possibility of bad faith they would have forbidden it expressly.

Should the law go further, and read into the contract an implied duty of reasonableness on A’s part? It should not (and does not). The parties probably meant A to be the sole judge of the adequacy of B’s performance. The language of the contract suggests, though not conclusively, and the suggestion is reinforced by reflecting on the incompetence of a judge or jury to determine whether A, though in fact (by assumption that he is acting in good faith) dissatisfied with the portrait, ought to have liked it. In contrast, if the contract, although containing the same language, were for the painting of the outside walls of a factory, the court might decide that the parties had not intended to make the buyer’s whimsy the measure of the seller’s compliance, as a judge or jury could determine without great difficulty whether the paint job was adequate to its workaday purpose.3

Let us change the example. A, a manufacturer, gives B, a dealer, an exclusive dealership for some area. That is, A agrees not to sell his product to anyone else in that area for the period of the contract. In a famous opinion by Judge Benjamin Cardozo, the court held that a contract for an exclusive dealership contains an implied condition that the dealer shall use his best efforts to sell the supplier’s product.4 Without such a condition, B could make the contract worthless to A simply by not selling A’s product, substituting the products of other manufacturers instead. The contract would be wholly one-sided; this presumably was not intended.

This example shows that another name for opportunism is—monopoly. The contract between the manufacturer and the dealer gave the latter a monopoly by depriving the manufacturer of the right to sell to competing dealers within the area specified in the contract. The law assumes that the parties did not contemplate that the dealer would be free to take advantage of the monopoly, so it interpolates a best-efforts condition.

Yet it could be argued that had the manufacturer wanted such protection, he would have negotiated for it, and that if he did not it may be because the parties preferred to avoid the possibility of litigation over the meaning of “best efforts.” They may have preferred to rely instead on the dealer’s interest in landing future contracts, or the shortness of the contract term, or the inclusion of a clause permitting either party to terminate the contract on short notice, to protect the manufacturer against the dealer’s exploiting the monopoly conferred by the contract. In other words, some gaps in contractual protection may be deliberate—the product of a tradeoff between the danger of opportunism on the one hand and the direct and indirect costs (including risk of error) of litigation, on the other.5

So opportunism is one problem with which the law of contracts must cope and another is the limitations of foresight. The longer the performance of a contract will take—and remember that contract “performance” includes the entire stream of future services that the transaction contemplates—the harder it will be for the parties to foresee the various contingencies that might affect performance. Persons contemplating a transaction can reduce the potential error costs arising from

5. This issue is discussed further in §21.16 infra. See also Douglas G. Baird, Self-Interest and Cooperation in Long-Term Contracts, 19 J. Legal Stud. 583 (1990); Alan Schwartz, Relational Contracts in the Courts: An Analysis of Incomplete Agreements and Judicial Strategies, 21 J. Legal Stud. 37 (1992).
imperfect foresight by shortening the duration of their contract (consider employment at will, or spot markets), since the near future is more predictable than the distant future. Or, what is similar, they can agree on just a few things and leave the rest for a future negotiation ("agreeing to agree"). Another alternative is the substitution of vertical integration for contracting—a producer might choose for example to make rather than buy an input, though really this would just be substituting employment contracts, plus contracts to buy whatever physical materials would be used in manufacturing the input in-house, for a contract to buy the input. A fourth alternative, however—and the one that deserves particular emphasis in a book on economic analysis of law—is for a court or an arbitrator to fill any gap in the contract when and if the gap emerges and precipitates a dispute.

Adjudicative gap-filling is a particularly economical method of dealing with contingencies that, even if foreseeable in the strong sense that both parties are fully aware that they may materialize, are so unlikely to do so that the costs of careful drafting to deal with them exceed the benefits when those benefits are discounted by the (low) probability that they will ever be realized. It may be cheaper from a social standpoint for the court to "draft" the contractual term necessary to deal with the contingency if and when the contingency materializes. (One adjudication may substitute for 1,000 drafting sessions.) Judicial or arbitral gap filling is similar to the use of form contracts (discussed further in §4.9 infra) to economize on drafting costs. The forms contain standard clauses designed to resolve contingencies that may arise in the course of performance. The clauses, like the gap-filling doctrines, are guesses as to what the parties probably would have specified had they negotiated rather than used the form.

The analogy to form contracts can help us to see that judicial gap-filling can be efficient even when a contingency is readily foreseen, as in the case of the exclusive dealership. If a court is rightly confident that most parties to such contracts would agree to impose on the dealer a best-efforts obligation, then reading such an obligation into every contract that does not mention best efforts avoids costs of negotiation and drafting. The parties can always opt out of it, but if they are less likely to opt out than they would be to opt in, the judicial interpolation of the provision is more efficient. The more remote the contingency, the more favorable the balance between negotiating over it in every contract and judicial interpolation in the rare case in which the contingency materializes inclines toward the latter approach.

The costs of judicial gap-filling can, however, be prohibitive, and then the court refuses to fill the gap; as in the common law's refusal to enforce a contract that lacks a price or quantity term. The alternative of interpolating a "reasonable price" or "reasonable quantity" clause is rejected because it would be too burdensome for a court to figure out what price or quantity the parties would have negotiated. Consistent with the reason behind the rule, however, the Uniform Commercial Code (in force in all U.S. states) permits the parties to omit price if there is a published spot price. Not only is the ascertainment of price mechanical in these circumstances, but specifying the price in the contract would defeat the contract's purpose, which is to transact at an ever-changing market price rather than at a fixed price.

It might seem that courts would never have good information for deciding what gap-filling rules would be optimal. But the norm of economic efficiency can provide some guidance. Each party wants to maximize his gain from the transaction, and that is usually best done by agreeing to terms that maximize the surplus created by

6. Arbitration as an alternative to adjudication is discussed in §§19.1, 21.16 infra.
the transaction — the excess of benefits over costs, the excess being divided between the parties. Of course, each party will be concerned not with the total surplus as such but only with the absolute size of his share of it. But he will be more likely to maximize his share if there is enough surplus for the other party to do well also. Hence gap-filling rules based on notions of efficiency will tend to mimic the terms that the parties would have incorporated into their contract explicitly had they foreseen the gap and been unwilling to rely on the courts to fill it sensibly. (Do you see an analogy to the Coase Theorem and to the merger solution to the problem of airplane noise, both discussed in the preceding chapter?) Since, moreover, the judicial gap-filling contract rules are only gap fillers, the parties can negate such a rule by expressly rejecting it in their contract. This point enables empirical answers to be given to such questions as whether contracting parties would really prefer to do without a legally enforceable best-efforts clause in order to reduce the risk of litigation. If that is their preference, their contracts will contain provisions disclaiming any duty to use best efforts; best-efforts litigation will dry up; and eventually the default rule will be changed. (None of this has happened.)

Filling potential gaps in contracts should be distinguished from disambiguating specific terms, which is the heart of the problem of contract interpretation. A contract might contain an explicit best-efforts clause, yet the wording of the clause might leave a doubt as to what exactly it required of the dealer. Gap filling and disambiguating are both, however, "interpretive" in the sense that they are efforts to determine how the parties would have resolved the issue that has arisen had they foreseen it when they negotiated their contract. And in both types of judicial intervention in the contractual process, economics can play a key role. The difference, however, is in the case of ambiguity the court cannot just lift a ready-made clause off the shelf and plug it into the case to decide the interpretive question, reasonably confident that if the rule didn't fit, the parties would have negated it in their contract.

1. What if it the parties' intentions, as gleaned from the language of the contract or from testimony, are at variance with the court's notion of what would be the efficient term to interpolate into the contract? If the law is to take its cues from economics, should efficiency or intentions govern? Paradoxically, the latter. Parties to a transaction are more trustworthy judges of their self-interest than a judge (or jury), who has neither a personal stake in nor firsthand acquaintance with the parties' venture. Yet discrepancies between (apparent) agreement and efficiency can be important clues to the existence of mistake, incapacity, or other grounds for believing that a literal interpretation of the contract would not really promote the parties' joint ends.

Suppose A buys goods from B, with delivery to take place in a month, and during the month B's warehouse burns down and the goods are destroyed. The contract is silent on the allocation of the risk of loss before delivery. But since B can prevent (or insure against) a fire in his own warehouse at lower cost than A, the parties, had they thought about the matter, would probably have assigned the risk to B, even though he no longer "owns" the goods; and that is the assignment the court should make in the absence of contrary evidence of the parties' intentions.

2. Generalizing, we expect the seller to warrant those dimensions of performance that are within his control rather than the buyer's. Thus he is held to warrant that the goods he sells are fit for their intended use — but not that they will last indefinitely, for their durability may depend on how the buyer uses them. The State of Wisconsin once hired a man named Bentley to build wings on the state capitol under the direction of the state's architect. Bentley followed the architect's plans
faithfully, but they were no good, and the wings collapsed shortly after being completed. The state sued Bentley, contending that he had guaranteed his work against such a calamity. The contract said nothing germane to the subject; obviously neither party had thought it likely that the wings would collapse because the architect's plans were bad. The state lost its suit. This is the right economic result. The state could have prevented the calamity at lower cost than Bentley, by more careful selection or supervision of the architect. Even so, might not one purpose of the contract have been to insure the state against the collapse of the wings from any cause? Insurance is one way of dealing with unforeseen contingencies, and contracts are often a method of insurance (see §4.5 infra). But it is unlikely that Bentley was a better insurer than the state. Bentley would have had to go out and buy an insurance policy; the state could self-insure against the particular risk.

We have thus far hewn to the notion of contract law as a handmaiden of exchange, but this takes too narrow a view of the subject, as some examples will illustrate:

(1) A wealthy man in an expansive moment promises to pay my way through college. I give up my part-time job, but he then breaks his promise, and I am unable to get a new job.

(2) A promises to deliver goods to B "on the twelfth." B thinks he means the twelfth of this month, but actually A means the twelfth of next month—he could not possibly deliver as soon as B (unbeknownst to him) expects.

(3) A steel company agrees to deliver steel to a bridge-building company within 60 days, but the company is shut down by a wildcat strike and cannot make delivery within that time.

In none of these cases is the issue whether a party to an exchange has refused to carry out his end of the bargain. There is no exchange in the first case; giving up my part-time job confers no advantage on the wealthy promisor. He may not even have known that I gave it up. In the second case there is no exchange; the parties intended different transactions. In the third case performance was prevented by circumstances beyond the promisor's control. Yet in all three cases there is an economic argument for imposing sanctions on the party who failed to perform.

(1) The wealthy man's idle promise induced reliance that cost the promisee heavily when it was broken, and such a cost can be avoided for the future by holding such a promisor liable for the promisee's cost of having relied. (2) In the case in which the buyer and seller confuse the date, suppose the custom in the industry is that a delivery date without specification of the month refers to the current month. A is new to the industry and ignorant of the custom. Nevertheless, to hold him to the promise understood by B will have the salutary effect of inducing newcomers to master the language of the trade promptly—although to be confident that this would be the optimal result we would have to consider whether existing firms might not be the cheapest source of newcomers' information about the custom, and the possible anticompetitive effects of placing the burden of acquiring this information on new entrants. (3) In the third case, the steel company is probably in a better position than the buyer to anticipate and take appropriate safeguards.

against an interruption of production due to a wildcat strike. If so, placing the risk of such an interruption on the steel company, by making it liable for damages to the purchaser from delay, may be the cheapest way of minimizing the costs of such delays in the future.

The question whether to treat a failure to carry through an undertaking as a breach of contract is similar to the question whether to treat an interference with a neighbor's land use as an invasion of the neighbor's property rights. We ask: Will imposing liability create incentives for value-maximizing conduct in the future? The difference is that less is at stake in a contract case. The setting is one of low transaction costs and therefore a judicial failure to discover the efficient solution can be rectified for the future by a drafting change. This point suggests, incidentally, that contract law cannot readily be used to achieve goals other than efficiency, as a ruling that fails to interpolate the efficient term will be reversed by the parties in their subsequent dealings.

To summarize (and anticipate), contract law has five distinct economic functions: (1) to prevent opportunism, (2) to interpolate efficient terms either on a wholesale or a retail basis (gap-filling versus ad hoc interpretation), (3) to punish avoidable mistakes in the contracting process, (4) to allocate risk to the superior risk bearer, and (5) to reduce the costs of resolving contract disputes.