Contribution to: JEL Classification: K12

The New Palgrave Dictionary

of Economics and the Law

Contracts

Steven Shavell

Abstract: The private and the social functions of contracts and of contract law are examined in this entry. In section 1, on the basic theory of contracts, the topics considered include contract formation, why contract enforcement is valuable, the incompleteness of contracts, the interpretation of contracts, remedies for breach, renegotiation of contracts, and judicial overriding of contracts. In section 2, the economic literature on production contracts is reviewed, and in section 3 other types of contract are discussed.

Contracts

The private and social functions of contracts and of contract law are examined here. In section 1 the basic theory of contracts is considered, in section 2 production contracts are analyzed, and in section 3 several other types of contract are discussed.

1. Basic Theory

1.1 Definitions and framework of analysis. A *contract* is a specification of the *actions* that named parties are supposed to take at various times, as a function of the *conditions* that then obtain. The actions usually comprise delivery of goods, performance of services, and payments of money, and the conditions include uncertain contingencies, past actions of parties, and messages sent by them.

A contract is said to be *complete* if the list of conditions on which the actions are based is exhaustive, that is, if the contract provides explicitly for all possible conditions. Otherwise, a contract will be referred to as *incomplete*. Typically, incomplete contracts do not include conditions which, were they easy to include, would allow both parties to be made better off in an expected sense. It should be noted that an incomplete contract may well not have literal gaps. For example, although a contract stating merely that a specified price is to be paid for a quantity of wheat that is to be delivered is incomplete (it does not mention many contingencies that might affect the buyer or the seller of wheat), the contract has no gaps, for it stipulates what the parties are to do (pay a price, deliver wheat) in all circumstances.

A contract in some relevant class of feasible contracts is called *Pareto efficient* if the contract is impossible to modify (within the class of contracts) so as to raise the expected utility

of both of the parties to it; such a contract will sometimes be referred to simply as *efficient* or as *mutually beneficial*. (Often, below, efficiency will connote the efficient terms in the class of complete contracts, that is, efficiency will refer to a "first-best" contract; this will be clear from context and should cause no confusion.)

Contracts are assumed to be enforced by a *tribunal*, which will usually be interpreted to be a state-authorized court, but it could also be another entity, such as an arbitrator or the decisionmaking body of a trade association or a religious group. (Reputation and other non-legal factors may also serve to enforce contracts but will not be discussed here; on such mechanisms of enforcement, see, for example, Bernstein 1992 and Charny 1990.) Enforcement refers to actions taken by the tribunal when parties to the contract decide to come before it. Tribunals may impose money sanctions — so-called *damages* — for breach of contract or insist on *specific performance* of a contract — require parties to do what a contract specifies (for example, convey land). Tribunals may also fill gaps, settle ambiguities, and override terms in contracts.

1.2 Contract formation. The formation of contracts is of interest in several respects.

Search effort. Parties expend effort in finding contracting partners, and it is apparent that their search effort will not generally be socially optimal. On one hand, they might not search enough: because the surplus gained when one party locates a contract partner would ordinarily be divided between them in bargaining, the private return to search may be less than the social return. On the other hand, parties might search more than is socially desirable because of a negative ("common pool") externality associated with discovery of a contract partner: when one party finds and contracts with a second, other parties are thereby prevented from contracting with that party. Both of these externalities arise in Diamond and Maskin (1979), who examine a

specific model of search and contracting. Although policies to promote or to discourage search might be desirable, one wonders whether social authorities could obtain the information needed to determine the nature of problems with search effort.

Mutual assent and legal recognition of contracts. A basic question that a tribunal must answer is at what stage of interactions between parties does a contract become legally recognized, that is, become enforceable. The general legal rule is that contracts are recognized if and only if both parties give a reasonably clear indication of assent, such as signing their names on a document. This rule obviously allows parties to make enforceable contracts when they so desire. Moreover, because the rule requires mutual assent, it protects parties against becoming legally obligated against their wishes. Thus, it prevents the formation of what might be undesirable contracts, and it means that search for contracting partners will not be chilled due to the risk of unwanted legal obligations.

However, certain legal doctrines sometimes result in parties becoming contractually bound without having given their assent: one party may become contractually bound if the party with whom he is negotiating makes investments in anticipation of contract formation. This legal policy not only may result in undesirable contracts, but also may induce wasteful early investment as a strategy to achieve contract formation. It is true that early investment is sometimes efficient, but a party who wants to make early investment could attempt to advance the time of contract formation or make a preliminary contract about the matter. See Bebchuk and Ben-Shahar (1996), Craswell (1996), and Wils (1993).

Offer and acceptance. Mutual assent sometimes is not simultaneous; one party will make an offer and time will pass before the other agrees. An issue that this raises is how long, and the

circumstances under which, the offeror will want to be held to his offer and whether he should be held to it. If an offeror is held to his terms, offerees will often be led to invest effort in investigating contractual opportunities. Otherwise offerees might be extorted by offerors if the offerees expressed serious interest after investigation. The anticipation of such offeror advantage-taking would reduce offerees' incentive to engage in investigation and thus diminish mutually beneficial contract formation. Hence, it may be in offerors' and society's interests for offered terms to be enforced for some period of time. Yet offerors' circumstances may change, making it privately and socially advantageous for them to alter contract terms. On this and related issues, see Craswell (1996) and Katz (1990, 1996).

Duress and emergency. Even if both parties have given their assent, a contract will not be recognized if it was made when one of the parties was put under undue pressure, as when he is physically or otherwise threatened by another. This legal rule has virtues similar to those of laws against theft; it reduces individuals' incentives to expend effort making threats and to defend against them. In addition, contracts may not be legally recognized if they are made in emergency situations, as when the owner of a ship in distress promises to pay an exorbitant amount for rescue. Nonenforcement in such situations beneficially provides victims with implicit insurance against having to pay high prices, but it also reduces incentives for rescue (yet rescue incentives might tend to be excessive, for the general reasons that there is excessive fishing effort).

Disclosure. The law may impose an obligation to disclose private information at the time of contract formation. Such a legal duty is beneficial in the respect that disclosed information may be desirably employed by the buyer; suppose, for instance, that he learns from the seller that the basement of his new house leaks and thus decides not to store valuables there. However, as

first emphasized by Kronman (1978a), a disclosure obligation discourages parties from investing in acquisition of information. For example, a company might decide against conducting aerial surveys to determine the mineral-bearing potential of land if it would be required to disclose its findings to sellers of land, as sellers would then demand a price reflecting the potential value of the land. The social welfare consequences of the effect of a disclosure obligation on the motive to acquire information, analyzed in Shavell (1994), depend on whether the information is socially valuable or mere foreknowledge, on whether the party acquiring information is the buyer or the seller, and on inferences that would be made from silence. (On inferences from silence, see also Grossman 1981, Milgrom 1981, and Fishman and Hagerty 1990.)

1.3 Why contracts and their enforcement are valuable to parties. At the most general level, parties make contracts when they have a need to make plans. They want contracts enforced to ensure that promised payments are made and to prevent opportunistic behavior that otherwise might occur over the course of the contractual relationship and stymy fulfillment of their plans. There are two basic contexts in which parties make enforceable contracts.

The first is that concerning virtually any kind of financial arrangement. The necessity of contract enforcement here is transparent. For example, because borrowers would not be forced to repay loans in the absence of contract enforcement, loans would be unworkable without enforcement. In financial arrangements, there is always a party who extends credit to another for some time period, and contract enforcement prevents his credit from being appropriated, which would render the arrangements impossible.

The second context in which parties make enforceable contracts involves the supply of custom or specialized goods and services — those which cannot simply be purchased on a spot

market in a simultaneous exchange of money for goods or services. The need for enforcement of agreements for supply of custom goods and services inheres mainly in averting what is often described as the *holdup problem* (discussed further in sections 2.1 and 2.2). To illustrate, consider a buyer who wants a custom desk which would be worth \$1,000 to him and would cost \$700 for a seller to produce. In the absence of contract enforcement, the buyer will not pay the seller in advance (for the seller could walk away with what he receives). The buyer will pay the seller only after he makes the desk. But at that point, the seller's production cost is sunk and he is vulnerable to holdup; the situation is that he has a desk which, being custom-made, has little or no alternative value. The outcome of bargaining between him and the buyer might thus be a price lower than the seller's cost of \$700, say the price is \$500. If so, and the seller anticipates receiving only the \$500 price, he will not produce the desk. This is true even though production and sale at a price between \$700 and \$1,000, such as \$800, would be mutually beneficial for the seller and the buyer. Enforcement of the buyer's promise to pay \$800 for the desk on delivery, or of the seller's promise to produce and deliver the desk (if the buyer paid the price of \$800 in advance), is thus desirable for the parties.

More broadly, enforcement of contracts will stimulate all manner of investments which, like the seller's expenditure on production, have specific value in a contractual relationship. Enforcement will lead buyers to train workers to use new contracted-for equipment, sellers to engage in investigation to reduce production costs, and so forth. In the absence of contract enforcement, there would be too little investment in these things, for, at the final stage of negotiation for performance and for payment, each side would be subject to holdup by the other, so would tend to obtain only a part of the surplus created by its investment.

The foregoing idea of contract enforcement as a cure for holdup-related underinvestment was initially stressed in the economic literature by Klein, Crawford, and Alchian (1978), Grout (1984), and Williamson (1975). However, the general notion that contract enforcement is privately and socially desirable because it fosters production and trade is made (usually with little articulation) by most writers on contract (see, for example, Farnsworth 1982: 16-17 and Pound 1959: 133-134) and one supposes has always been appreciated.

1.4 Incomplete nature of contracts and their less-than-rigorous enforcement.

Although enforceable contracts are desirable, they are observed to be substantially imperfect.

They are significantly incomplete, leaving out all manner of variables and contingencies that are of potential relevance to contracting parties, and they also often fail to employ included variables in a mutually beneficial manner. Moreover, contracts are not enforced rigorously, despite the seeming strength of the reasons for contract enforcement; penalties for violation of contractual obligations are often modest, and breach is not an uncommon event.

There are two main reasons for the incompleteness of contracts. The first is the *cost of writing more complete contracts*. Parties may not include variables in a contract, or not in a detailed, efficient way, due to the cost of evaluating, agreeing upon, and writing terms. (In particular, parties will tend not to specify terms for low probability events, because the expected loss from this exclusion will be minimal, whereas the cost of including the terms is borne with certainty.) Also, parties may not include some variables (effort levels, technical production difficulties) because they cannot be verified by tribunals. Of course, many such variables can be made verifiable (effort could be made verifiable through videotaping), but that would involve expense.

The second reason for the incompleteness of contracts is that the *expected consequences* of incompleteness may not be very harmful to contracting parties. Incompleteness may not be harmful simply because a tribunal might interpret an imperfect contract in a desirable manner. In addition, the prospect of having to pay damages for breach of contract may serve as an implicit substitute for more detailed terms. Furthermore, the opportunity to renegotiate a contract often furnishes a way for parties to alter terms in the light of circumstances for which contractual provisions had not been made. Finally, in some settings parties' concern for their reputation may induce them to refrain from opportunistic behavior.

That contracts are less than rigorously enforced is intimately related to their incompleteness. For incomplete contracts to function to the parties' advantage, tribunals must be able to reinterpret or override imperfect contractual terms rather than always enforce these terms as written. Also, for damage measures to be employed beneficially by parties, notably for parties to be able to escape from contractual obligations when performance and renegotiation are difficult, damages for breach must not be excessive. Additionally, for parties to avoid bearing high risks in the form of payments they would be induced to make when renegotiating imperfect contractual terms, the damages for breach must again not be severe. These points will be expanded in the discussion below of contract interpretation, damage measures for breach, and renegotiation.

1.5 Interpretation of contracts. Contractual interpretation, which includes a tribunal's filling gaps, resolving ambiguities, and overriding literal language, can benefit parties by easing their drafting burdens. For example, if it is efficient to excuse a seller from having to perform if his factory burns down, the parties need not incur the cost of specifying this exception in their

contract, assuming that they can trust the tribunal to interpret their contract as if the exception were specified.

It may be worthwhile elaborating somewhat by viewing contract interpretation more formally, as a function that transforms the contract individuals write into the *effective* contract that the tribunal will enforce. Given a method of interpretation, parties will choose contracts in a constrained-efficient way. Notably, if an aspect of their contract would not be interpreted as they want, the parties would either bear the cost of writing a more explicit term that would be respected by the tribunal, or else they would not bear the cost of writing the more explicit term and accept the expected loss from having a less than efficient term. The best method of contract interpretation will take this reaction of contracting parties into account and can be regarded as implicitly minimizing the sum of the costs the parties bear in writing contracts and the losses resulting from inefficient enforcement.

1.6 Damage measures for breach of contract. When parties breach a contract, they often have to pay damages in consequence. The damage measure, the formula governing what they should pay, can be determined by the tribunal or it can be stipulated in advance by the parties to the contract (in which case it is often called *liquidated damages*). One would expect parties to specify their own damage measure when it would better serve their purposes than the measure the tribunal would employ, and otherwise to allow the tribunal to select the damage measure. In either case, we now examine the functioning and utility of damage measures to contracting parties.

Clearly, the prospect of payment of damages is an *incentive to perform contractual obligations*, and thus generally promotes enforcement of contracts and the goals of the parties, as

discussed in section 1.3. As emphasized in section 1.4, however, damages for breach in fact are not chosen to be so high that they virtually guarantee performance of contracts as written. Under the commonly employed *expectation measure*, damages equal the amount that compensates the victim of breach for his losses; these damages are often quite willingly paid by a party who commits a breach.

Why are damages not chosen to be so high as to guarantee performance? An important explanation is that parties do not always desire performance of the less-than-complete contracts that they write; it may not be Pareto efficient for there to be performance in circumstances for which the parties did not make provisions in their incomplete contract. For example, suppose that a contract is very incomplete: it states only that "The seller will produce a custom desk for the buyer and the buyer will pay \$800 in advance." The buyer and the seller do not really want the desk always to be produced. It is readily shown that, had they made a complete contract, they would have specified performance if and only if the production cost were less than the \$1,000 value of the desk to the buyer. (For instance, they would have jointly decided against a contractual term specifying performance when the production cost is \$2,000, for the seller would have been willing to reduce the contract price sufficiently to induce the buyer to strike that term.) Now if the incomplete contract calling for the desk always to be produced is enforced by the expectation measure of damages of \$1,000, the seller will behave exactly as he would have under the complete contract, that is, he will perform if and only if the production cost is less than \$1,000. Higher damages than the expectation measure might induce performance when it is inefficient, and lower damages might induce breach when that is inefficient. Indeed, for this reason, the parties would often agree to choose the expectation measure over another measure of

damages.

This understanding of damage measures as a device to induce the behavior that the parties would have specified in more complete contracts sheds light on the notion, held by many legal commentators, that contract breach is immoral, as it constitutes the breaking of a promise. That belief is often incorrect, it is submitted, and might fairly be considered to be the opposite of the truth. The view that a contract breach is the breaking of a promise overlooks the point that the contract that is breached is generally an incomplete contract, and that the breach is what the parties *want* and would have specified in a complete contract. In the example of the simple incomplete contract calling for a desk to be produced, the seller who finds that his production cost would be \$2,000 will commit breach under the expectation measure. But in so doing, he will be acting precisely as would have been set out in a complete contract, and it is that contract which is best regarded as the promise between the parties that ought to be kept.

The point that a moderate damage measure, and in particular the expectation measure, is desirable because it induces performance if and only if the cost of performance is relatively low is apparently first clearly stated (informally) in Posner (1972), who emphasized the social efficiency of the measure, and (formally) in Shavell (1980), who stressed the mutual desirability of the measure for contracting parties and also its role as a substitute for more complete contracts. (Two earlier writers, Birmingham 1970 and Barton 1972, adumbrate these points, although the meaning of their articles is at points obscure; see also Diamond and Maskin 1979, who consider damage measures in analyzing search behavior.)

Several more comments should be made about damage measures and incentives. First, damage measures influence the motive of contracting parties to make *reliance investments* (so

called because the investments are made relying on contract performance). Reliance investments are illustrated by the earlier-noted instance of a buyer training workers to use a contracted-for machine or by advertising the contracted-for appearance of an entertainer. Under the expectation measure, there is a tendency for reliance investment to exceed the Pareto efficient level: the buyer will treat an investment like advertising as one with a sure payoff — either he will receive performance or receive expectation damages, a form of insurance — whereas the actual return to investment is uncertain, due to the possibility of breach (advertising will be a waste if the entertainer does not appear). This tendency toward overreliance due to the receipt of contract damages was initially noted in Shavell (1980), and stands in contrast to the problem of inadequate reliance investment associated with lack of contract enforcement. The issue of reliance investments has been elaborately analyzed in the economic literature, as will be described in section 2.2.

A second comment is that the value of damage measures as an incentive toward efficient performance would not exist if renegotiation of contracts in problematic contingencies would always result in efficient performance. But, as will be discussed below, the assumption that renegotiation would necessarily lead to efficiency is implausible.

An important function of damage measures which is quite distinct from their incentive role concerns *risk-spreading and compensation*. Notably, because the expectation measure compensates the victim of a breach, the measure might be mutually desirable as a form of insurance if the victim is risk averse. However, the prospect of having to pay damages also constitutes a risk for a party who might commit breach (such as a seller whose costs suddenly rise), and he might not want to bear this risk due to his risk aversion. The latter consideration

may lead parties to want to lower damages (see Polinsky 1983) or to avoid use of damages as an incentive device, by writing more detailed contracts (for instance, the parties could go to the expense of specifying in the contract that a seller can be excused from performance when his costs are high.) A full consideration of damage measures and efficient risk-allocation would also take into account whether the risk a party bears is detrimental or beneficial (if a party wants to breach because he has a superior opportunity, payment of damages would not be likely matter as much on risk-bearing grounds), and whether the risk is monetary or non-monetary (if the victim's loss is non-monetary, such as the loss due to failure of musicians to appear at a wedding, he may have little desire for receipt of damages as a form of insurance).

1.7 Specific performance as a remedy for breach. As observed at the outset, an alternative to use of a damage measure for breach of contract is specific performance: requiring a party to satisfy his contractual obligation. (Some economists have employed the term "specific performance" in an unconventional and broader sense, to refer to enforcement of all provisions in a contract, *including* any damage measure named in it. Thus, they would say that a contract is specifically performed when the parties name expectation damages in their contract and parties might well commit breach and pay these damages.) Specific performance can be accomplished with a sufficiently high threat or by exercise of the state's police powers, such as by a sheriff removing a person from the land that he promised to convey. (Note that if a monetary penalty can be employed to induce performance, then specific performance is equivalent to a damage measure with a high level of damages.)

It is apparent from what has been said about incomplete contracts and damage measures that parties should not want specific performance of many contracts that they write, for they do

not wish their incomplete contracts always to be performed. It is therefore not surprising that, in fact, specific performance is not used as the remedy for breach for most contracts for production of goods and for provision of services. Additionally, it may be observed that specific performance might be peculiarly difficult to enforce in these contexts because of problems in monitoring and controlling parties' effort levels and the quality of production.

However, specific performance does have advantages for parties in certain contexts, such as in contracts for the transfer of things that already exist, like land, and specific performance is the usual legal remedy for breach of contracts for sale of land. This point is discussed briefly below, in section 3. On specific performance and its general comparison to damage remedies, see Bishop (1985), Kronman (1978b), Schwartz (1979), Shavell (1984), and Ulen (1984). (Specific performance also is examined in many of the articles on production contracts cited in section 2.2.)

1.8 Renegotiation of contracts. Parties often have the opportunity to renegotiate their contracts when problems arise. Indeed, the assumption that they will do this has appeal because, having made an initial contract, the parties know of each other's existence and of many particulars of the contractual situation. For this reason, much of the economic literature on contracts assumes that renegotiation always occurs when inefficiency would otherwise result; see, for example, Hart (1987) and Rogerson (1984).

Nevertheless, in many circumstances contracts will not be renegotiated because parties are not in contact with each other when difficulties are experienced and one party would benefit from acting quickly. A problem may occur during the course of production and the producer may have to decide on the spot whether to abort the process or proceed at greater cost. Or a party

may hear a new bid and have to answer it immediately. Furthermore, even if the parties are in contact with one another, asymmetric information may lead to breakdowns in renegotiation.

In any event, let us assume that successful renegotiation tends to occur when difficulties arise and consider how it affects the welfare of contracting parties. Plainly, renegotiation often allows parties to prevent Pareto inefficient outcomes. For example, if damages exceeding the expectation measure or specific performance were the remedy for breach, a seller might be led to perform when production cost exceeds the value of performance to the buyer. To avoid this inefficient outcome, the seller might pay the buyer to release him from his obligation to perform. That renegotiation may result in performance if and only if it is efficient means, as noted, that damage measures for breach are not necessary to accomplish this, and also helps to explain why contracts lack detail.

But even if renegotiation occurs, it may represent only a partial substitute for explicit contractual terms or for appropriate damage measures for breach. One reason (see section 2.3) for this assertion is that renegotiation cannot affect precautions which were taken before the time of renegotiation and which influence the likelihood of nonperformance; renegotiation can only affect future decisions about breach. Another reason involves the allocation of risk-bearing. Consider, for instance, the substantial risks borne by a producer who may have to purchase a release from an obligation to perform when his production costs would be extremely high. Such risks could be mitigated by use of a clause excusing him from performance or by a damage measure such as expectation.

The prospect of renegotiation also influences the incentives of parties to invest in the contractual relationship. Renegotiation may not lead to efficient levels of reliance investment if

it results in the extraction of part of the surplus that a party's investment creates. Yet renegotiation is affected by, among other elements, the damage measure that applies for breach; and if the damage measure is appropriately chosen, the damage measure together with renegotiation may spur desirable reliance investment; see section 2.2.

One presumes that the ability to renegotiate is usually desirable for contracting parties, because it allows them to improve their situation when difficulties arise and to write simpler contracts than otherwise. Thus, we would expect that parties will want their renegotiated contracts enforced, and the law generally does enforce renegotiated contracts. However, for somewhat subtle reasons, the ability to renegotiate can also work to the detriment of parties. See Jolls (1997) and the literature cited therein, especially Fudenberg and Tirole (1990). That renegotiation might be undesirable is not surprising from a formal perspective, for the ability to renegotiate might prevent parties from committing themselves to particular outcomes in their initial contract. Nevertheless, the law usually prevents parties from binding themselves not to renegotiate, even though that could in principle be done. (If courts were to refuse to enforce renegotiated contracts, such contracts would not be made, even secretly, as long as one of the parties to the original contract would prefer that it be enforced over any renegotiated contract.)

1.9 Judicial overriding of contracts. A basic rationale for judicial overriding of contracts is the existence of externalities. Contracts that are likely to harm third parties are often not enforced, for example, agreements to commit crimes, price-fixing compacts, liability insurance policies against fines, and certain simple sales contracts (such as for machine guns). Also, individual provisions of contracts may not be enforced because of externalities (consider clauses in sales contracts that would bar subsequent purchase by blacks).

Another general rationale for nonenforcement of contracts is to prevent a loss in welfare to one or both of the parties to contracts. This concern may motivate nonenforcement when a party is incompetent, lacks relevant information, or is in an emergency situation (see section 1.2). The rationale also applies in the context of contract interpretation by tribunals; as discussed (see section 1.5), contract interpretation may amount to overriding terms of contracts, and this may promote the welfare of contracting parties by allowing them to write simpler contracts. Also, at least in theory, nonenforcement of contracts might also be beneficial to parties where they would be led to include terms constituting wasteful signals of unobservable characteristics; see Aghion and Hermalin (1990).

Additionally, contracts sometimes are not enforced because they involve sale of things said to be inalienable, such as human organs, babies, and voting rights. In many of these cases, however, the inalienability justification for lack of enforcement can be recognized as involving externalities or the welfare of the contracting parties. See generally Rose-Ackerman (1985); and see also Trebilcock (1993: 23-77).

2. Production Contracts

In this section the economic literature on production contracts is discussed. The first case considered is that where symmetrically informed, risk-neutral parties contract, and the only variables of concern are the value of performance and production cost. Then the case where parties make reliance investments to raise the value of the contract during the contract period is examined. Finally, several other issues, including risk-bearing and asymmetric information, are reviewed. Throughout, when damages for breach are discussed, one can imagine them either to be chosen by the parties or by the courts.

2.1 Value of performance and production cost. Assume that a risk-neutral buyer and a risk-neutral seller have met; the seller faces uncertain production cost c, which he will learn before he decides whether to produce; v is the certain value of performance to the buyer; and the parties are symmetrically informed. The Pareto efficient outcome is for the seller to produce if and only if c < v. (That is, in a contract with terms for all contingencies, performance would be required if and only if c < v; a change in the contract price would compensate a party for agreeing to alter a term from any initially considered contract under which performance does not occur if and only if c < v.)

In the absence of contract enforcement, then (amplifying on section 1.3) there would be too little production because the buyer would only pay the seller for actual delivery of the good and cannot guarantee the price. In particular, supposing that the seller would obtain a fraction α of the surplus from a transaction (α reflects bargaining strength), he would obtain a price of αv . (After the seller produces the good, the surplus from the transaction would be v, presuming for simplicity that the custom good has no alternative value for the seller.) Thus, the seller would decide to produce only when $c < \alpha v$, rather than whenever c < v.

Suppose now that there is contract enforcement and that the parties are not able to renegotiate before the seller decides whether to produce (an assumption that is relaxed below). If c is verifiable by the tribunal, the parties could write a complete contract specifying efficient performance if and only if c < v. The parties would want a damage measure d for breach of this contract to be sufficiently high definitely to induce performance when c < v, and thus any d exceeding c would work. If c is not verifiable, the parties are able to write an incomplete contract specifying "The buyer shall deliver the good to the seller at price p, paid at the outset,"

accompanied by damages d for breach. (The assumption that the price p is paid at the outset rather than at performance is only simplifying; it does not affect the conclusions.) Under such a contract, there will be performance when c < d and breach otherwise. If the expectation measure is employed, that is, d = v, the seller will perform when c < v, so that performance will be efficient. If damages d exceed v, there will be excessive performance, as there will be if there is specific performance. If d is less than v, there will be too little performance. The points of this paragraph were, as noted, emphasized in Posner (1972) and Shavell (1980).

If, instead, it is assumed that the buyer and the seller can renegotiate their contract after c becomes known but before the seller decides whether to produce, then, given symmetric information, it is natural to suppose that there will always be Pareto efficient performance, regardless of d.

Let us also note that if the buyer's value v is uncertain as well as the seller's production $\cos c$, the major difference in conclusion is that v must be verifiable for the expectation measure d = v to be applied by the tribunal (c still need not be verifiable).

2.2 Reliance investment during the contract period. Now assume that parties can make investments during the period of the contract that affect its value v or the production cost c. Such investments are, as noted, sometimes called reliance investments, since they are made in anticipation of contractual performance. We will begin with the case in which just one party invests before discussing the case where both sides invest.

Buyer makes reliance investment and seller's costs are uncertain. Suppose that one party to the contract invests, for concreteness the buyer, and that the other party faces uncertainty. Specifically, let r be the buyer's reliance investment (training of workers to use a contracted-for

machine) and let v(r) be the value of performance given r, where v is increasing in r. The buyer chooses r before the seller learns c and decides about producing. The Pareto efficient decision of the seller is to produce when c < v(r), and the efficient decision of the buyer is therefore to choose r to maximize

$$v(r)$$

$$\int (v(r) - c)g(c)dc - r,$$
0

where g is the density of c. Thus, the optimal r, denoted r^* , is determined by v'(r)G(v(r)) = 1, where G is the cumulative distribution of c. The point to note here is that the marginal return to reliance investment is only an *expected* return, for the investment pays off only when c < v(r) — when production turns out to be efficient.

In the absence of contract enforcement, there will be too little production, as before; it will occur only when $c < \alpha v(r)$. But now, in addition, the buyer will choose an incorrect value of r because he will only obtain a fraction 1 - α of the value created by the investment r.

Assume next that there is contract enforcement and that *the parties do not renegotiate* before the seller's production decision. This is the setting analyzed in Shavell (1980), who first studied reliance investment. If c and r are verifiable by the tribunal, the parties can write a contract specifying efficient performance (when c < v) and also specifying r^* ; again, they would want the contract enforced by a damage measure high enough to ensure performance, and any such measure of damages would serve their purposes.

Now assume that c and r are not verifiable, that the parties write a simple contract specifying "The buyer will pay price p and the seller will deliver the good to him," and consider what occurs under different damage measures. If the expectation measure is employed, that is, d

= v(r), the seller will perform when c < v(r), so that performance will be efficient. However, as the buyer will always receive v(r) (either he obtains performance, worth v(r), or damages of that amount), he will choose r to maximize v(r) - r. Consequently, the buyer will select an inefficiently high r; the problem is that the buyer does not take into account that investment does not have any value when performance does not occur. Under a sophisticated expectation measure based on *efficient* investment, namely $d = v(r^*)$, however, investment as well as performance may be verified to be efficient (as was first noted by Cooter 1985).

Another damage measure that has been examined is the *reliance measure*, according to which the buyer would receive his price p plus reliance investment r if the seller breaches. Under this measure, if there is a breach, the buyer will be placed in the position he would have enjoyed had he not invested and made the contract. It can be shown that, under the reliance measure, investment would be even more excessive than under the expectation measure, and there would be too little performance. (Note, however, that to apply the reliance measure, courts must be able to verify investment r, and that if this is so, r^* could be achieved simply by the parties naming it in their contract.) Finally, under specific performance, there is excessive performance, but r is chosen optimally given that level of performance (because performance always occurs).

Next assume that *the parties do renegotiate* after reliance investment is made and before the seller decides about production, so that, assuming symmetric information, the supposition is that there will always be efficient performance. This version of the model of production contracts was originally studied by Rogerson (1984). Here, damage remedies may influence investment through their effect on the outcome of renegotiation. To illustrate, consider what would occur under specific performance. Under this remedy, as suggested earlier, there will be

renegotiation in which the seller pays the buyer to be allowed not to perform whenever c > v(r), since then performance would be inefficient. In particular, the assumption is that the seller would pay the buyer $v(r) + (1 - \alpha)(c - v(r))$ to be allowed not to perform; for v(r) is needed to compensate the buyer for not receiving performance, $1 - \alpha$ is the buyer's share of the surplus from renegotiation, and c - v(r) is that surplus. Anticipating this, the buyer can be shown to choose an r exceeding the efficient level. The nature of the results about reliance investment in the case with renegotiation are very close to those where there is no renegotiation. Indeed, they are identical under the expectation measure, essentially because there is no renegotiation under the expectation measure; thus, with d = v(r), investment will be excessive because the buyer will always be compensated for his investment. Furthermore, under the sophisticated expectation measure based on efficient investment, $d = v(r^*)$, investment will be efficient (Spier and Whinston 1995).

Both parties make reliance investments and both the value of performance and production costs are uncertain. Here let $v = v(r, \theta)$ and $c = c(t, \theta)$, where t is reliance investment of the seller and θ is the state of nature; t lowers t given t. In this more general situation, what occurs can be understood in many respects by analogy to the case just discussed. For example, under the expectation measure, investment will tend to be excessive for both parties, but performance will be efficient.

Much recent literature, beginning with Hart and Moore (1988), has focused on this general situation, assuming that parties can renegotiate after reliance investments are made and θ is revealed, and that they will always then agree on efficient production decisions because information is symmetric. The literature in question furthermore usually supposes that none of

the variables (costs, values of performance, reliance investments) are verifiable by the tribunal. Thus a contract can depend only on what is recorded in it, certain subsequent communications between the parties, whether there has been performance, and, if not, who committed breach.

Of note are a number of results establishing the existence of contracts that will produce efficient outcomes, that is, in both parties choosing efficient levels of reliance investment (performance will always be efficient). Aghion, Dewatripont, and Rey (1994) and Chung (1991) demonstrate the efficiency result using a contract in which one party is effectively given the right to make a single take-it-or-leave-it offer to the other in renegotiation. It is evident that this party will invest efficiently, as he can extract in bargaining the full marginal return from his investment. For instance, if the buyer has the right to make an offer and is paying the seller to perform, he will pay only the minimum needed to induce the seller to do so, and will pocket any increase in value $v(r,\theta)$ due to his having chosen a higher r. Less apparent is how the other party is given an incentive to invest efficiently; that is accomplished by properly choosing the quantity of the good or the probability of delivery. (For instance, if the named quantity of the good is chosen to be higher than is likely to be efficient, the buyer will usually pay the seller to agree to lower the quantity. The amount the buyer will pay must compensate the seller for the profits he would have made at that higher contracted-for quantity. But the profits the seller would have made will depend on his investment in lowering production costs — thereby giving the seller an incentive to invest in lowering his production costs, and an incentive that is greater the higher the contracted-for quantity.) Also, Nöldeke and Schmidt (1995) establish that a simple option contract will induce efficient investments for reasons that are closely related to those just reviewed. Additionally, Edlin and Reichelstein (1996) and Hermalin and Katz (1993) adduce

contracts leading to efficiency under somewhat different conditions; and Rogerson (1992) shows that efficiency can be achieved under wide circumstances, but assuming that parties can commit not to renegotiate their contracts.

Cooperative reliance investments. It has been assumed above that a reliance investment benefits directly only the party who makes it. Another possibility is that a reliance investment benefits the other party to the contract; importantly, suppose that a seller's investment raises product quality and value for the buyer. Such cooperative reliance investment is studied in Che and Chung (1995). As they emphasize, when cooperative investment cannot be verified by courts, then under the expectation measure, there will be too little investment (in contrast to the usual case under the expectation measure, where investment is excessive). Indeed, there will be no investment under that measure because the seller who makes a cooperative investment will not benefit directly or in damages that he receives in the event of breach. Moreover, there is no contract that will result in efficient cooperative investment (again in contrast to the usual case); this point is stressed in Che and Hausch (1996), who also demonstrate that contracting offers no advantage over no contracting in wide circumstances when cooperative investment is not verifiable.

2.3 Further considerations. *Risk-bearing.* We have not discussed in this section the allocation of risk among possibly risk-averse contracting parties, about which several comments should be made. First, if all variables are verifiable by a tribunal, the presence of risk-averse parties clearly does not affect when it is Pareto efficient to perform; it continues to be efficient to perform if and only if c < v. However, efficiency requires that the resulting risk be allocated appropriately; for instance, if the seller is risk averse and the buyer risk neutral, the seller would

be insured against fluctuations in c by the buyer's paying him c plus a constant. The level of efficient reliance investment also will generally be affected by considerations of risk-bearing.

Second, when variables of relevance are not verifiable, then damage measures and other mechanisms that may be employed to induce efficient behavior when parties are risk-neutral have to be reconsidered. For instance, the expectation measure imposes risk on the party who might breach and pay these damages; so if that party is risk-averse, the expectation measure would become less attractive relative to lower measures of damages. Furthermore, as earlier noted, renegotiation does not generally lead to efficient risk-bearing, even though it may lead to efficient performance.

Asymmetric information. Another factor about production contracts that has not been examined is asymmetric information between the parties. When parties are asymmetrically informed, renegotiation of contracts might not be successful, so that it becomes more important that the initial contract induces efficiency; Hermalin and Katz (1993) show that efficiency can be achieved under certain types of asymmetry of information using a relatively complicated mechanism in the contract. Asymmetry of information also affects bargaining about initial contracts; on various aspects of this issue, see Ayres and Gertner (1989), Bebchuk and Shavell (1991), and Stole (1992).

New entrants. Mention has not been made of the possibility that new buyers would appear and bid for the seller's good (a similar possibility is that new sellers would appear and make offers to the buyer). In this regard, it should be noted that it is Pareto efficient for the initial contracting parties that a sale be made to a new buyer if and only if his bid exceeds the contract buyer's valuation. Moreover, the contracting parties will want to maximize the amount

that they can extract from a new buyer if he purchases the good. This observation raises the possibility that the buyer and the seller may wish to set damages for seller breach at a high level in order to induce a new party to bid more (which he would have to do to make it in the seller's interest to commit breach). Such an incentive of contracting parties to set damages at high levels can, though, result in too little breach and sale to new parties; thus, at least in principle, the incentive in question is a ground for tribunals not to enforce the high damage level specified by the contracting parties. This point was first made in Diamond and Maskin (1979) and has been refined in a number of articles; see Aghion and Bolton (1987) and Chung (1992). However, Spier and Whinston (1995) observe that three-way renegotiation would seem to vitiate the advantage to the contracting parties of setting high damages; yet they emphasize another reason (concerning induced reliance investment) that the parties will, after all, benefit from setting high damages.

Precautions and probabilistic breach. It has been supposed throughout that breach occurs when a party decides not to perform, but often breach does not occur in this way: rather a party chooses a level of precaution which affects the likelihood of performance, and a random factor then determines whether breach or performance results. For example, a shipper's care in packing dishes affects the likelihood that they will arrive unbroken, and a chance event (a jolt) determines whether they arrive broken or unbroken. In this setting, the conclusions reached about damage measures in the absence of renegotiation continue to apply: the expectation measure results in efficient precautions, the buyer's reliance investment is excessive, and so forth. However, the very issue of renegotiation is made moot because the precautions are chosen before breach might occur (if the dishes arrive broken, it is too late for renegotiation). See

Cooter (1985), Craswell (1988), and Kornhauser (1983).

3 Other Types of Contract

3.1 Contracts for transfer of possession. A different contractual context from production is where something that already exists is to be conveyed to a buyer. Examples include contracts for transfer of real estate, goods in inventory, used durable goods. Here a major uncertainty of interest concerns bids by new parties. With regard to these bids, the points just discussed about new entrants apply; the parties would like for there to be a sale to a new buyer when he will pay more than the contract buyer's valuation, and so forth.

It is of interest to explore why contracting parties often adopt specific performance as the remedy for breach of contracts for transfer of possession even though damage measures are commonly employed for other types of contract. Initially suppose that the contract buyer and the contract seller have equal access to bids from new parties. Then the buyer's always receiving the good does not result in any loss of opportunity to sell to a new party willing to bid a high amount. Moreover, specific performance offers an advantage over damage measures. Namely, because under specific performance it will always be the buyer who will be bargaining with a new party, the good will never be sold to a new party bidding less than the buyer's valuation. In contrast, such an inefficient sale could occur if the seller might pay damages, commit breach, and bargain with a new party (suppose that bargaining is not three-way, involving the contract buyer as well). And an inefficient sale by a breaching seller to a new party would make the original contracting parties collectively worse off than under specific performance: either the buyer would not obtain the good even though his valuation exceeds what the seller received from the new party; or the buyer would obtain the good through repurchase from the new party, but in general at a higher

party. (Although the contracting parties would be worse off if the buyer repurchases at a higher price, society would not be worse off as the good would still be allocated to the user who places the highest value on it.) See Shavell (1984) and Bishop (1985).

The foregoing advantage of specific performance in preventing inefficient sales to new parties is clearly reduced if the buyer does not have equal access to bids from new parties (suppose that the seller is a dealer and the buyer is not). Also, use of specific performance might increase transaction costs, if the new party turns out to purchase only after delivery of the good to the buyer.

Notice too that some of the disadvantages of specific performance in the production context are less significant in the present context of transfer of possession. In production contracts, specific performance imposes a possibly large risk of loss on sellers whose production costs might be very high; here, specific performance only reallocates a beneficial risk (of a sale at a high price) from seller to buyer. In addition, enforcement of specific performance in the context of contracts for transfer of possession is often easier than in the production context, where enforcement might involve policing the quality of production or services.

3.2 Donative contracts. An important category of contractual arrangement is donative, concerning gifts. Assuming that the motivation for gifts is altruism, a basic question is why a donor would want to defer his gift rather than make it immediately. The answers include the possibilities that the donor may face liquidity constraints and that he may wish to wait for resolution of uncertainties concerning, among other factors, his own needs and future income, and the donee's needs, future income, and character.

Given that a donor does desire to defer making a gift, would he want to make a contract that would in some way bind him? The disadvantage of so doing is that it may not be feasible for him to limit as he wishes the conditions under which he makes the gift (due to the costs of specifying these conditions and to the problems the courts would have in verifying them). The principal advantage of a contract to the donor is that it may induce the donee to engage in reliance activities that will increase the value of the gift to the donee (a high school student might study more if he anticipates a gift that will finance his college education). Such reliance activities will in turn inure to the benefit of the donor because of his altruism. However, if the donee knows about the altruism of the donor, a contract may not be necessary to induce donee reliance activity; if so, a contract would be disadvantageous for the donor. On these issues, see Goetz and Scott (1980) and Shavell (1991); and see also Posner (1977) and Posner (1997).

3.3 Additional types of contract. In this entry, mention has not been made of many additional types of contract, including principal-agent contracts, even though they have been studied, often intensively, in the economic literature. The omission of such contracts is explained in part by convention (by what is and is not considered to be a law and economics topic) and in part by the economic literature's relative inattention to contract enforcement.

References

- Aghion, P. and Bolton, P. 1987. Contracts as a barrier to entry. *American Economic Review* 77: 388-401.
- Aghion, P., Dewatripont, M. and Rey, P. 1994. Renegotiation design with unverifiable information. *Econometrica*. 62: 257-282.
- Aghion, P. and Hermalin, B. 1990. Legal restrictions on private contracts can enhance efficiency. *Journal of Law, Economics*,

 & Organization. 6: 381-409.
- Ayres, I. and Gertner, R. 1989. Filling gaps in incomplete contracts: an economic theory of default rules. *Yale Law Journal*. 99: 87-130.
- Barton, J. 1972. The economic basis of damages for breach of contract. *Journal of Legal Studies*. 1: 277-304.
- Bebchuk, L.A. and Shavell, S. 1991. Information and the scope of liability for breach of contract: the rule of *Hadley v. Baxendale. Journal of Law, Economics, & Organization* 7: 284-312.
- Bebchuk, L.A. and Ben-Shahar, O. 1996. Pre-contractual reliance. Discussion Paper No. 192, John M. Olin Center for Law, Economics, and Business, Harvard Law School.
- Bernstein, L. 1992. Opting out of the legal system: extralegal contractual relations in the diamond industry. *Journal of Legal Studies*. 21: 115-157.
- Birmingham, R.L. 1970. Breach of contract, damage measures, and economic efficiency. *Rutgers Law Review* 24: 273-292.
- Bishop, William. 1985. The choice of remedy for breach of contract. *Journal of Legal Studies*. 14: 299-320.
- Charny, D. 1990. Nonlegal sanctions in commercial relationships. *Harvard Law Review*. 104: 373-467.
- Che, Y-K., and Chung, T-Y. 1995. Contract damages and cooperative investments. Department of Economics, University of Wisconsin-Madison.
- Che, Y-K., and Hausch, D.B. 1996. Cooperative investments and the value of contracting. Department of Economics, University of Wisconsin-Madison, forthcoming, *American Economic Review*.

- Chung, T-Y. 1991. Incomplete contracts, specific investments, and risk-sharing. *Review of Economic Studies*. 58: 1031-1042.
- Chung, T-Y. 1992. On the social optimality of liquidated damage clauses: an economic analysis. *Journal of Law, Economics, & Organization.* 8: 280-305.
- Cooter, R. 1985. Unity in tort, contract, and property: the model of precaution. *California Law Review*. 73: 1-51.
- Craswell, R. 1988. Contract remedies, renegotiation, and the theory of efficient breach. *Southern California Law Review*. 61: 629-670.
- Craswell, R. 1996. Offer, acceptance, and efficient reliance. Stanford Law Review. 48: 481-553.
- Diamond, P.A. and Maskin, E. 1979. An equilibrium analysis of search and breach of contract, I: steady states. *Bell Journal of Economics*. 10: 282-316.
- Edlin, A.S. and Reichelstein, S. 1996. Holdups, standard breach remedies, and optimal investment. *American Economic Review*. 86: 478-501.
- Farnsworth, E.A. 1982. Contracts. Boston: Little Brown.
- Fishman, M.J. and Hagerty, K.M. 1990. The optimal amount of discretion to allow in disclosure. *Quarterly Journal of Economics*. 105: 427-444.
- Fudenberg, D. and Tirole, J. 1990. Moral hazard and renegotiation in agency contracts. *Econometrica*. 58: 1279-1319.
- Goetz, C.J. and. Scott, R.E. 1980. Enforcing promises: an examination of the basis of contract. *Yale Law Journal*. 89: 1261-1322.
- Grossman, S.J. 1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics*. 24: 461-483.
- Grout, P.A. 1984. Investment and wages in the absence of binding contracts: a Nash bargaining approach. *Econometrica*. 52: 449-460.
- Hart, O. 1987. Incomplete contracts. 752-759, in *The New Palgrave*, J. Eatwell, M. Milgate, and P. Newman (eds.), New York: MacMillan Press.
- Hart, O. and Moore, J. 1988. Incomplete contracts and renegotiation. *Econometrica*. 56: 755-785.

- Hermalin, B.E. and Katz, M.L. 1993. Judicial modification of contracts between sophisticated parties: a more complete view of incomplete contracts and their breach. *Journal of Law, Economics, & Organization.* 9: 230-255.
- Jolls, C. 1997. Contracts as bilateral commitments: a new perspective on contract modification. *Journal of Legal Studies*. 26: 203-237.
- Katz, A. 1990. The strategic structure of offer and acceptance: game theory and the law of contract formation. *Michigan Law Review*. 89: 215-295.
- Katz, A. 1996. The economics of promissory estoppel in preliminary negotiations. *Yale Law Journal*. 105: 1249-1309.
- Klein, B., Crawford, R.G., and Alchian, A.A. 1978. Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics*. 21: 297-326.
- Kornhauser, L.A. 1983. Reliance, reputation, and breach of contract. *Journal of Law and Economics*. 26: 691-706.
- Kronman, A.T. 1978a. Mistake, disclosure, information, and the law of contracts. *Journal of Legal Studies*. 7: 1-34.
- Kronman, A.T. 1978b. Specific performance. *University of Chicago Law Review*. 45: 351-382.
- Milgrom, P.R. 1981. Good news and bad news: representation theorems and applications. *Bell Journal of Economics*. 12: 380-391.
- Nöldeke, G. and Schmidt, K.M. 1995. Option contracts and renegotiation: a solution to the holdup problem. *Rand Journal of Economics*. 26: 163-179.
- Polinsky, A.M. 1983. Risk sharing through breach of contract remedies. *Journal of Legal Studies*. 12: 427-444.
- Posner, E.A. 1997. Altruism, status, and trust in the law of gifts and gratuitous promises. *Wisconsin Law Review*. 1997: 567-609.
- Posner, R.A. 1972 (1st ed.). Economic analysis of law. Boston: Little-Brown.
- Posner, R.A. 1977. Gratuitous promises in economics and law. *Journal of Legal Studies*. 6: 411-426.
- Pound, R. 1959. An introduction to the philosophy of law. New Haven: Yale University Press.

- Rogerson, W. 1984. Efficient reliance and damage measures for breach of contract. *Rand Journal of Economics*. 15: 39-53.
- Rogerson, W.P. 1992. Contractual solutions to the hold-up problem. *Review of Economic Studies*. 59: 777-794.
- Rose-Ackerman, S. 1985. Inalienability and the theory of property rights. *Columbia Law Review*. 85: 931-969.
- Schwartz, A. 1979. The case for specific performance. Yale Law Journal. 89: 271-306.
- Shavell, S. 1980. Damage measures for breach of contract. *Bell Journal of Economics*. 11: 466-490.
- Shavell, S. 1984. The design of contracts and remedies for breach. *Quarterly Journal of Economics*. 97: 121-148.
- Shavell, S. 1991. An economic analysis of altruism and deferred gifts. *Journal of Legal Studies*. 20: 401-421.
- Shavell, S. 1994. Acquisition and disclosure of information prior to sale. *Rand Journal of Economics*. 25: 20-36.
- Spier, K.E. and Whinston, M.D. 1995. On the efficiency of privately stipulated damages for breach of contract: entry barriers, reliance, and renegotiation. *Rand Journal of Economics*. 26: 180-202.
- Stole, L.A. 1992. The economics of liquidated damages clauses in contractual environments with private information. *Journal of Law, Economics, & Organization*. 8: 582-606.
- Trebilcock, M.J. 1993. *The limits of freedom of contract*. Cambridge: Harvard University Press.
- Ulen, T.S. 1984. The efficiency of specific performance: toward a unified theory of contract remedies. *Michigan Law Review*. 83: 341-403.
- Williamson, O.E. 1975. *Markets and hierarchies: analysis and antitrust implications*. New York: Free Press.
- Wils, W.P.J. 1993. Who should bear the costs of failed negotiations? A functional inquiry into precontractual liability. *Journal des Economistes et des Etudes Humaines*. 4: 93-134.