Abstract

This chapter surveys and analyzes the substantial literature on optimal remedies for contract breach in a variety of settings. It begins with a standard analysis of the behavioral effects of expectation and reliance damages, then discusses the application of these damage measures in a world where courts are not perfectly informed about the parties’ valuations of the contract. When valuation problems are extreme, courts may turn to alternative remedies such as specific performance, or parties may attempt to solve the problem themselves through liquidated damages clauses. The chapter considers whether these solutions to the valuation problem alleviate or exacerbate opportunistic behavior by the parties. It also highlights the recent contributions that game theory and options theory have made to the understanding of remedial choices.

JEL classification: K12

Keywords: Contract Damages, Remedies

1. Introduction

The principal remedy for breach of contract in Anglo-American law is an award of money damages. The preferred measure of damages is the expectation measure, under which the promisee receives a sum sufficient, in theory, to make him indifferent between the award and the performance. Other damage measures, and other remedies such as specific performance and rescission, are available in special circumstances. This chapter discusses the basic design of the remedial system.

A. The General Problem

2. Sanctions and Incentives

A contract is an exchange of promises or an exchange of a promise for a present performance, and the parties enter into it because each values the thing
received more than the thing foregone. These values are based on expectations about the future because some or all of the contractual performance will occur in the future. When the future diverges from what a party expected, he may conclude that the performance he will receive under the contract is no longer more valuable than the performance he must provide. He has, in the terminology of Goetz and Scott (1980), experienced a ‘regret contingency’ and now would prefer not to perform and not to receive the promised performance from the other party.

Absent a system of contract remedies, a party who regrets entering into a contract will not perform unless he fears that the breach will result in sanctions by the other party (who might have required security for the performance) or by third parties (who might revise their opinion of the breacher and reduce their economic and/or social interactions with him accordingly). The economic function of contract remedies, then, is to alter the incentives facing the party who regrets entering into the contract, which will directly affect the probability of performance and indirectly affect the number and type of contracts people make, the level of detail with which they identify their mutual obligations, the allocation of risks between the parties, the amount they invest in anticipation of performance once a contract is made, the precautions they take against the possibility of breach, and the precautions they take against the possibility of a regret contingency.

An administratively simple system of remedies would aim to reduce the probability of breach to near zero. That could be achieved by the routine (and speedy) grant of injunctions against breach backed by large fines for disobeying the injunction or by imposing a punitively large monetary sanction for breach. This would give promisees a high degree of confidence that the promised performance will occur and induce a high level of investment in anticipation of performance. In the standard parlance, this would be a ‘property’ rule because it would entitle the promisee to the performance except to the extent the promisor is able to negotiate a modification on terms acceptable to the promisee.

### 3. Efficient Contracts and Efficient Nonperformance

Were it possible to enter into complete state-dependent contracts (that is, contracts that identified every possible contingency (state) and specified the required actions of the parties for each), parties would be willing to be bound to contracts even were the sanction for breach punitive. Such contracts would require performance in some states but excuse it in others, in such a way that each party would be willing ex ante to be absolutely bound to perform the required actions in all states. Shavell (1980) defines a ‘Pareto efficient complete contingent contract’ as a complete state-dependent contract to which no mutually beneficial modifications could be made, viewed at the time of
contracting. We will call such contracts ‘efficient’. In doing so, we will assume unless otherwise stated that the parties are risk neutral, each party’s objective is to maximize his wealth, post-contractual renegotiation is prohibitively costly, performance is all or nothing (that is, partial performance is not possible), and the contracts do not create uncompensated gains or losses for third parties.

Under what conditions would an efficient contract excuse performance? Shavell demonstrates that the contract would require performance in all circumstances except those in which nonperformance would result in greater joint wealth. An example will illustrate the point. Imagine that Seller agrees to manufacture and sell to Buyer a machine that Buyer will use in its own manufacturing process. The value of the machine to Buyer is $300; however, Buyer has an opportunity to make certain alterations to his manufacturing plant, at a cost of $50, which will increase the value of the machine to $375. Such investments by a promisee in anticipation of performance are called ‘reliance expenditures’ or ‘reliance investments’ in the literature, and we will use the terms interchangeably. Assume for the moment that the future can be represented as a set of two possible states; Seller’s production cost is $200 in one state and $400 in the other. An efficient contract would require Seller to make the machine in the low-cost state but not in the high-cost state. In the high-cost state, the joint wealth of the parties is greater if Seller does not perform than if it does. This can be seen by comparing the cost of performance to Seller ($400) with the benefit to Buyer ($300 or $375, depending on whether Buyer makes the reliance investment). The contract price is irrelevant as it is transferred from Buyer to Seller and does not affect their joint wealth.

Both parties can be made to prefer this contract to one that requires performance in both states. They can allocate between themselves the extra wealth created by the efficient contract, and there will be some allocations under which each party’s expected gain exceeds the expected gain from the contract that always requires performance. By choosing such an allocation, each party will be better off at the time of contracting and willing to be bound to perform or not perform as required. In the literature, a breach that occurs in circumstances in which an efficient contract would excuse performance is called an ‘efficient breach’.

4. Barriers to Efficient Contracting; Remedies as a Substitute for Efficient Contracts

To reiterate, faced with an efficient contract, courts would have the simple task of requiring strict adherence to its terms. Unfortunately, the writing of efficient contracts is no easy task. It is costly to bargain over remote contingencies and the parties may lack the foresight to deal with all possible states. Moreover, the
parties may not have equal access to the information necessary to tell which state occurs. In the above example, Seller may know whether the cost of manufacturing the machine is $200 or $400, but Buyer may have no way to verify Seller’s assertion.

Given these barriers to efficient contracting, the law faces a more complex problem than that of compelling adherence to efficient contracts. Instead, it must take incomplete contracts and augment them by damage measures that induce behavior that mimics reasonably closely the behavior that an efficient contract would require. A particular damage measure can be termed ‘efficient’ with respect to a particular decision if it creates an incentive for the relevant party to make the same decision it would under an efficient contract. Because standard damage measures allow a promisor to breach and pay compensatory (rather than punitive) damages, they are called ‘liability’ rules in contrast to property rules as defined above.

5. Other Approaches

An alternative framework for the design of damage measures is offered by Barton (1972). He poses the problem as one of designing damage measures that would induce the parties to make the same decisions regarding performance or breach, and reliance prior to performance or breach, that they would make were the parties divisions of a single, integrated firm and had the sole objective of maximizing the value of the firm. Shavell and Barton each show that the objective of an efficient regime of contract damages is to cause the parties to maximize their joint wealth, although one might prefer Shavell’s conceptual approach on the grounds that Barton’s assumes away the problem by positing a wealth-maximizing firm.

A more recent perspective on contract damages is to consider money damages as an option under which, for example, Seller may purchase Buyer’s entitlement to Seller’s performance. The option expires on the date fixed for performance and its strike price is the damage award (which may from the parties’ perspective be a random variable). The value of the option is reflected in the contract price (see Mahoney, 1995; Ayres and Talley, 1995). This literature derives from the more general use of option theory to analyze decision making under uncertainty (see Dixit and Pindyck, 1994). We will make occasional reference to the options perspective below.
B. The Standard Damage Measures

6. A Taxonomy of Damage Measures

The preferred measure of contract damages is the amount of money that will make the promisee indifferent between performance and damages. It should be noted at the outset that this formulation of the measure of damages is not fully accurate; there are a number of limiting doctrines, discussed in Chapter 4620, that often reduce money damages below the promisee’s subjective valuation of the performance. There is also some evidence that courts award greater damages for breaches that appear opportunistic (see Cohen, 1994). As courts express it, however, the preferred measure of damages is the amount necessary to put the aggrieved party in the same position as if performance had occurred, which is known as the expectation measure.

Fuller and Perdue (1935) provide the standard taxonomy of contract damage measures. They identify three different ‘interests’ of the promisee that are affected by a breach - expectation, reliance, and restitution - and state that the most common damage measures provide compensation for one of the three. The expectation interest is measured by the net benefit the promisee would receive should performance occur, as described above. The restitution interest consists of any benefit the promisee has provided the breaching party. For example, if a seller agrees to make monthly deliveries of a commodity in return for fixed payments due 60 days after each delivery and the buyer repudiates the contract after receiving and retaining two deliveries but making no payments, restitutionary damages would restore to the seller the value of the delivered goods. The reliance interest is measured by the promisee’s wealth in the pre-contractual position. Reliance damages provide compensation both for any benefit conferred on the breaching party and for any other reliance investments made by the promisee in anticipation of performance to the extent such investments cannot be recovered.

In most instances, the restitution measure will provide the lowest recovery and the expectation measure the highest. One complication is how to treat other contractual opportunities that Buyer passed up in order to enter into the contract with Seller. Analytically, these seem similar to reliance investments and are often treated as such. In a competitive market, where Buyer could have entered into another contract at an identical price had she not contracted with Seller, the reliance measure and the expectation measure will converge approximately. ‘Approximately’, because the value of the alternative contract is a function of the probability that it will be performed (see Cooter and Eisenberg, 1985) and of the damage remedy if it is not performed, and thus the problem is somewhat circular. When analyzing the difference between expectation damages and reliance damages below, we will assume that they
differ and that Buyer’s expectation interest exceeds his reliance interest. We will also assume that the reliance interest equals or exceeds the restitution interest, although we will relax that assumption in Section 12 below.

7. Incentives Within an Existing Contract: The Decision to Perform or Breach

The expectation measure leads to efficient decisions to perform or breach an existing contract, given a fixed level of reliance (see Barton, 1972; Shavell, 1980; Kornhauser, 1986). This can be illustrated using the example set out above. Assume that the contract price for the machine is $250 and that Buyer makes an irrevocable decision to invest $50 in reliance, an investment that has no value absent the contract. When production costs are $200, Seller will manufacture the machine and Buyer will pay $250 for it. Buyer then obtains a machine worth $375 to him for a total expenditure (contract price plus reliance expenditure) of $300. The transaction increases Buyer’s wealth by $75. When production costs are $400, Seller will breach. The expectation measure seeks to make Buyer as well off as if Seller had performed. Seller’s breach relieves Buyer from his obligation to pay the contract price. Accordingly, if Seller pays Buyer damages of $125, Buyer will be in the same position as if Seller had performed, having paid out a non-recoverable $50 in reliance and received $125, for a net increase in wealth of $75.

So long as Buyer is awarded $125 in the event of breach, Seller will breach only when the cost of performance exceeds $375, the value of the performance to Buyer. Compare this result to that obtained under the reliance measure. Under the reliance measure, Seller must compensate Buyer for his $50 reliance investment. Assume for a moment that there is a third possible state under which Seller’s cost of production is $350. Performance would be efficient because its value to Buyer exceeds its cost to Seller. Seller will perform given expectation damages, because the damage award of $125 exceeds the net loss from performance ($350 cost minus the $250 contract price). Given reliance damages, however, Seller will breach and pay $50 rather than perform at a loss of $100. More generally, it is obvious that given expectation damages, only when the production cost reaches $376 will Seller become better off by breaching and paying damages than by performing and losing the difference between his production cost and the contract price. The expectation measure, unlike the reliance measure, causes Seller to internalize fully the effect on Buyer’s wealth of Seller’s decision to perform or breach.
8. Incentives within an Existing Contract: The Decision to Rely

While the expectation measure produces efficient decisions to breach given reliance, it does not produce efficient levels of reliance. In general, expectation damages result in excessive reliance expenditures, because they cause Buyer to act as if performance were always forthcoming. In the example, Buyer will always spend $50 to increase the value of performance from $300 to $375, because either (1) the performance will be forthcoming or (2) Buyer will be compensated for the lost $375 in value. In the high-cost state, however, the parties’ joint wealth would be greater if Buyer refrained from investing. Seller would be liable for damages of $300 less the $250 contract price, or $50. By contrast, if Buyer relies, he receives $125 in damages as shown above and increases his wealth by $75 net of the reliance expenditure. Unlike the no-reliance case, where Buyer gains $50 and Seller loses $50, here Buyer gains $75 and Seller loses $125. The difference reflects the fact that the $50 expenditure is wasteful in the high-cost state. Expectation damages, then, do not cause Buyer to internalize fully the effect on Seller’s wealth of Buyer’s decision to make a reliance investment.

The reliance measure is subject to the same objection. Under the reliance measure, Buyer will recover $50 if it invests that amount in reliance. Once again, Buyer’s investment decision will be made as if the investment is not risky, even though it is (because performance is inefficient in some states). Indeed, reliance damages create a perverse incentive for Buyer in some circumstances. Assume for a moment that Seller’s production cost is $310. Under the reliance measure, Seller will pay damages of $50 rather than perform and suffer a loss of $60 ($310 minus the $250 contract price). Breach deprives Buyer of a $75 gain (showing again that any measure of damages less than the expectation measure induces inefficient breach decisions). Buyer may be able to avoid breach, however, by making an additional (and we will assume wasteful) reliance expenditure of $11. Now reliance damages amount to $61, and Seller performs. Thus the excessive breach problem can be cured in part, but at the cost of excessive reliance. In general, as Shavell (1980) demonstrates, the reliance measure will result in greater (inefficient) reliance expenditures than the expectation measure. There is no measure of damages that results both in efficient decisions to perform or breach and efficient decisions to make or not make reliance expenditures. However, expectation damages do better than reliance damages at inducing efficient breach decisions, and do no worse than reliance damages at inducing efficient reliance decisions. Accordingly, given the various assumptions outlined above, the expectation measure is preferable on efficiency grounds.

The analysis to this point has assumed risk neutrality. A risk-averse Buyer would have additional cause to prefer the expectation measure, because it eliminates variability from Buyer’s outcome. At the same time, the expectation measure introduces greater variability into Seller’s outcome than does the
reliance measure. It is accordingly possible that where both parties are risk averse, they may find that a sum of damages greater than the reliance measure but less than the expectation measure offers the highest joint utility level. The precise formulation of the damage amount would depend on the parties’ comparative levels of risk aversion (see Polinsky, 1983). It seems plausible that courts have not tried to alter damage measures to accommodate risk aversion (except to the extent specific performance can do so, as discussed in Section 10 below) because of the administrative and error costs that would result.

The analysis has also assumed that renegotiation at the time of breach is prohibitively costly. Were negotiation costless, the damage rule would be irrelevant, as the parties would in all cases bargain to Pareto efficient breach/performance and reliance decisions, as per the Coase Theorem (Coase, 1960). In the more plausible situation where renegotiation is costly but not always prohibitively so, the choice of remedy is necessarily more difficult. Some critics have therefore argued that much of the literature on damage remedies is beside the point, as the choice of remedies should be informed principally by an analysis of transaction costs (see Friedmann, 1989; MacNeil, 1982). Friedmann analyzes potential transaction costs in a variety of contractual settings and argues that over compensatory remedies (remedies that provide compensation to the promisee in excess of the expectation interest) will generally be efficient.

9. Incentives at the Stage of Contract Formation

Friedmann and MacNeil are surely correct to argue that a better understanding of the costs of postcontractual renegotiation is necessary for making efficient remedial choices. It is also, however, worth paying attention to the effect of remedies on precontractual negotiations.

The price Seller will require to enter into a contract is increasing in the damage measure. Returning to our example, when Buyer makes a $50 reliance expenditure and Seller breaches, again assuming no opportunity costs, Buyer’s wealth decreases by $50. Buyer can be no worse off from entering into the contract so long as the remedy for breach is at least $50. Will Buyer be willing to pay more for the more generous expectation measure, and will Buyer and Seller prefer the resulting contract to one that provides for reliance damages only? As Friedman (1989) notes, the difference in remedies affects the contract price, the quantity contracted, and the quantity actually consumed, with effects that vary with market structure and utility functions. In general, however, the range of contract prices for which the contract increases both parties’ wealth will be greater under a reliance measure than an expectation measure. That is, the reliance measure will create a greater bargaining range, which might increase the number of contracts entered into.
The choice of remedies where precontractual as well as postcontractual incentives are analyzed remains an underdeveloped area. Friedman (1989) provides a formal analysis of expectation and reliance for two contexts in which those measures diverge. The first is the case of a breaching buyer who has contracted to purchase from a monopolist selling at a single price. The second is the case of a breaching buyer in a competitive market where the seller does not know its production cost in advance but the buyer does. Friedman demonstrates that neither damage remedy dominates the other under those conditions. Friedman’s analysis is limited, however, by his assumption that reliance is fixed and exogenous. The situations he analyzes, moreover, have the desired formal characteristics (expectation and reliance measures diverge) but are probably not very common.

A possible alternative would be to start by assuming that the expectation and reliance measures diverge without specifying market structure in detail. A model could then be developed in which the choice between expectation and reliance damages affects the structure of the contract, the decision to breach, and the decision to rely. The equilibrium and comparative statics of such a model might shed light on the type of market conditions under which expectation or reliance damages would be more nearly optimal. It would also be valuable to consider carefully whether there are plausible conditions under which the cost of negotiating around an inefficient damages measure at the time of contracting is greater or less than the cost of renegotiating at the time of performance.

C. Alternative Damage Measures

10. Specific Performance

Disappointed promisees are not in all cases limited to an award of damages; under appropriate circumstances, they may seek the equitable remedy of specific performance. A decree of specific performance requires the breaching party to perform according to the contract. The principal criterion for awarding specific performance is a demonstration that money damages are insufficient to compensate the promisee for the lost performance. Traditionally, this was most often found when the breaching party was a seller who had agreed to sell a ‘unique’ good. Real estate has long been presumed in many jurisdictions to be unique, while other goods such as artworks and heirlooms are often found to be unique.

Specific performance is analogous to a punitive sanction that seeks to deter breach absolutely. In order for it to have that effect, we must assume that renegotiation is costly. It would then seem clear that expectation damages are preferable to specific performance, because the latter would sometimes result
in performance even though nonperformance would result in greater joint wealth. On the other hand, it should be clear that the assumption that courts can adequately calculate a sum of money sufficient to make the promisee indifferent between damages and breach is not always accurate, particularly where cover is not possible. In such circumstances we must rely on a lost surplus measure of damages, and the calculation of Buyer’s consumer surplus is necessarily subjective. This is not a fatal objection if we believe that courts will guess correctly on average, but if they systematically underestimate Buyer’s surplus, the monetary remedy will result in too much breach, just as specific performance results in too much performance.

Kronman (1978) started the law and economics debate on specific performance by employing a framework similar to that of the prior paragraph. He notes that specific performance is a property rule in the sense defined in Section 2 above; it effectively assigns the promisee an absolute entitlement to the goods from the moment the contract is made. This does not make sense in most instances because renegotiation (meaning a transfer of the property right back to the promisor) is costly and the result will be an inefficiently high level of performance. The danger of undercompensation, which would result in an inefficiently low level of performance, is normally lower because there is often a substitute price available. When, however, there is no substitute price available (the case of ‘unique’ goods), the danger of undercompensation likely outweighs the cost of renegotiation. Accordingly, the legal rules, in a rough manner, promote efficiency.

Schwartz (1979) argues that undercompensation is not merely an isolated problem limited principally to goods for which there is no obvious substitute, but is built into the structure of money damages. The reluctance of courts to award damages that are uncertain, difficult to measure, or unforeseeable (see Chapter 4620), or to provide compensation for emotional harm resulting from a breach, makes money damages systematically under compensatory. Schwartz argues that the resulting inefficiencies are likely greater than those resulting from renegotiation costs, and accordingly that specific performance, rather than money damages, should be the default remedy.

Bishop (1985) adopts a similar analytic approach but argues that both Kronman and Schwartz have overgeneralized their arguments. He breaks down contract breaches into a number of categories depending on the identity of the breaching party (buyer or seller), the type of contract, and the alternative transactions available to buyer and seller. He also identifies another cost of awarding specific performance. Because the value of a specific performance award (including the amount the promisor will pay to be released from performance) will in some cases exceed the value of performance to the promisee, the promisee will be tempted to behave opportunistically in hopes of causing a breach and satisfying the conditions for specific performance. Bishop
argues that in some categories the problem of excessive breach resulting from undercompensation will dominate, and in others the problem of excessive performance resulting from renegotiation costs and opportunism will dominate. The relative magnitudes of the inefficiencies generated by costly renegotiation and undercompensation are ultimately empirical questions and to date the literature does not provide data from which we could confidently identify the preferred remedy. Accordingly, Mahoney (1995) takes a different approach to the problem, using the option methodology outlined above. The methodology is first employed to confirm the argument made by Craswell (1988) that were renegotiation costless and money damages perfectly compensatory, risk-averse contracting parties would always prefer money damages to specific performance. The intuition is that entering into a contract with a money damages remedy is analogous to holding a hedged position in a commodity, whereas the identical contract with a specific performance remedy is analogous to holding an unhedged position. The variance of possible outcomes is greater for both parties with the unhedged contract and they will accordingly prefer money damages. In the face of costly renegotiation and undercompensation, we can still make some sense of the case law using the option heuristic. Many contracts involving ‘unique’ goods are prompted by the desire of the buyer to speculate on the future value of the land, artwork, and so on, and speculation involves holding an unhedged position. Thus buyer and seller would likely prefer specific performance. Other cases in which specific performance has been consistently awarded (long-term contracts to supply a fuel input to a public utility or other regulated entity) can be explained by noting that the buyer is likely more risk averse with respect to price fluctuations than is the seller, and specific performance better accommodates that distinction.

11. Liquidated Damages

We began the analysis of damages by arguing that court-awarded damages function as a substitute for complete state-dependent contracts. The court’s application of an efficient damages rule creates appropriate incentives to perform or not perform, rely or not rely, and so on, and thereby saves the parties the trouble of drafting their contract to provide for all contingencies. Some parties, however, choose to create a tailor-made incentive structure by specifying the amount of damages payable in the event of breach. Courts have adopted a skeptical attitude toward these so-called liquidated damages clauses. In general, courts will enforce a liquidated damages clause only if (a) at the time of contracting, the damage that the promisee will suffer in the event of breach (that is, the promisee’s expectation) is uncertain, and (b) the amount of liquidated damages is both a reasonable estimate of (the mean of) those
damages and not disproportionate to the actual (ex post) damages. A larger amount is called a ‘penalty’ and is unenforceable.

This attitude is puzzling to most law and economics scholars. Absent some reason to believe that one or both parties misunderstood the terms of the agreement, we would normally assume that they went to the trouble to specify liquidated damages because the resulting contract is Pareto superior to a contract that calls for the ordinary court-awarded damages. The courts’ approach might increase the net wealth of the parties, but only if the existence of a penalty is strong evidence that the contract does not represent the parties’ actual intent, perhaps because one defrauded the other. The focus of scholarship in the area, therefore, has been to ask under what circumstances rational, well-informed parties would agree to a penalty clause.

Goetz and Scott (1977) note, in terms similar to those of the later specific performance debate, that damages measures do not adequately compensate for the subjective value the promisee attaches to performance, particularly when close substitutes for the performance are unavailable or the timing of the performance is critical. In those circumstances, normal damage measures will lead to excessive breach.

The promisee could purchase third-party insurance that would pay off in the event of breach in an amount sufficient to make him whole. In some circumstances, however, the probability of breach is determined not just by exogenous variables beyond the parties’ control, but also the level of care taken by the promisor. For example, a delivery service can affect the probability of timely delivery by the level of care it takes with the package. In such circumstances, the promisor can insure more cheaply than a third-party because the promisor can alter its level of care in response to the insurance clause. In other circumstances, the promisor may be better informed about the probability of breach than a third-party insurer. The delivery service, for example, may be better informed than the promisee or any third-party insurer about the breakdown rate of its trucks. In such cases, the promisor can use its willingness to agree to a penalty clause as a means of credibly signaling a low probability of breach.

It is accordingly not true that the mere existence of a penalty clause is a strong indicator of fraud or mistake; there are plausible conditions under which a penalty clause would make both parties better off. Schwartz (1990) supplements this analysis by considering the effects of a penalty clause on pre-contractual incentives. He notes that the price that the promisor will charge is increasing in the damage measure. To the extent promisees insist on inefficiently large penalties, therefore, they will either pay too much or enter into too few contracts. The promisee accordingly has an incentive to demand a penalty clause only when it would increase the joint wealth of the parties.
Other commentators have argued that penalty clauses can create externalities. Aghion and Bolton (1987), for example, demonstrate that a supply contract containing a penalty clause for buyer’s breach can restrict entry by competing sellers. We might at first conclude that buyer would have no incentive to agree to a contract that limited competition from other sellers. A well-designed penalty, however, will merely redistribute wealth from the new entrant to the contractual buyer and seller. To see why, assume a contract between Buyer and Seller with a contract price of $200, and Seller’s cost of performance is $150. At a later time, Entrant appears, who can provide the good for $100. Seller’s expectation damages will be $50, so absent a penalty Buyer can profitably breach if Entrant offers a price of $149 or less. Assuming that Entrant does not face competition, Entrant will demand a price of $149. Now imagine that Buyer must pay a penalty of $80 upon breach. Now it is not profitable for Buyer to breach unless Entrant offers a price of $119 or less. Entrant can still profitably sell at that price, and will do so. Thus the penalty transfers wealth from Entrant to Seller (which Seller can agree \textit{ex ante} to share with Buyer).

Aghion and Bolton’s analysis works only if Entrant has market power. The externality arguments for the most part are not sufficiently general to provide a compelling explanation for the judicial hostility toward penalty clauses. Moreover, while they can provide support for part of the judicial approach (disallowing liquidated damages that are not a fair \textit{ex ante} estimate of actual damages), they cannot explain the failure to enforce liquidated damages that are excessive \textit{ex post}.

12. Rescission/Restitution

Courts divide contract breaches into ‘partial’ and ‘total’ breach. A partial breach gives the promisee the right to seek a remedy but not to refuse his own performance. The classic example is when a builder constructs a house that contains a minor deviation from the agreed architectural plan. The builder must compensate the owner for the difference in value (in theory, the difference in subjective value to buyer, but it will usually be difficult to convince a court that this differs substantially from the difference in market value). The owner may not, however, refuse to accept delivery of the house and to pay the agreed price. A total breach, by contrast, permits the promisee to refuse to render his own performance. In effect, a total breach permits the promisee to rescind the contract.

As courts express it, a promisee can respond to a total breach by seeking expectation damages or by rescinding and seeking recovery of any value he has provided to the breaching promisor. The latter alternative is equivalent to the restitution measure of damages (although in some circumstances the promisee
may seek return of the performance in specie rather than its monetary equivalent). Restitution is also a remedy in quasi-contractual situations, such as when parties partly performed a contract that is voidable for mutual mistake, but the following discussion will be limited to restitution damages as a remedy for breach.

In the typical case, expectation damages will exceed restitutionary damages and the promisee will seek the former. There are two instances, however, in which we would expect the promisee to seek the latter. The first is when the promisee is risk averse and prefers the certainty of the return of money or property that he has given the promisor to the uncertainties of a jury’s assessment of his expectation and the additional litigation costs that would be incurred in the attempt. The second is when the contract was a losing deal for the promisee, so that his expectation is negative. Where the promisee has provided something of value to the promisor that cannot be easily returned, but can be valued in a judicial proceeding, the promisee may be better off receiving that value in cash than receiving the promised performance.

This might be thought a remote possibility, but it occurs in a number of reported cases. The textbook example is one in which a builder agrees to build a house for an owner and the builder’s costs turn out to be greater than expected, making the contract a losing one for the builder. The owner, however, later decides it does not want the house and repudiates the contract when the house is partly completed. The builder’s expectation is negative because of the unexpectedly high costs of construction, so the builder seeks restitution. Restitution in this instance is measured by the value the builder has conferred on the owner, or the market value of the nearly-completed house. By hypothesis, this exceeds the contract price.

When promisees have attempted to recover reliance damages for a losing contract, courts have concluded that the expectation measure puts an upper bound on the recovery (see *L. Albert & Son v. Armstrong Rubber Co.*). By contrast, some courts have permitted a promisee to recover restitution damages in excess of expectation (see *Boomer v. Muir*). This seeming inconsistency has been largely ignored in the law and economics literature. The most useful discussions appear in a symposium issue of the *Southern California Law Review* in 1994. In it, Kull (1994) provides an analysis of restitution that is similar in many respects to Bishop’s analysis of specific performance. Money damages are not always an adequate substitute for performance and the damage calculation is in any event uncertain. Thus where the promisee has provided something of value to the promisor that can easily be returned, the promisee may prefer to rescind the transaction, putting both parties back in the pre-contractual position. For example, the promisee may have paid in advance for a good or service that the promisor fails to provide. Taking litigation costs into account, the promisee may prefer to rescind the transaction and retrieve the
One example of a situation in which it seems likely that rescission and restitution will minimize the costs associated with breach is where a seller delivers goods that do not conform to the contractual specifications. The perfect tender rule, recognized under the common law and the Uniform Commercial Code, permits a buyer to reject nonconforming goods even if the variation is minor. As noted by Priest (1978), the administrative costs involved in calculating the difference in value between the goods as delivered and as promised will likely exceed the cost of returning the goods to Seller and money to Buyer. The costs associated with salvaging the nonconforming goods might also be minimized by the perfect tender rule, as in many instances it will be cheaper for Seller to find another purchaser for the goods than it will be for Buyer to adapt the goods to Buyer’s own use.

On the other hand, where the contract is a losing one for the promisee and the promisee has conferred a benefit on the promisor that cannot easily be returned, the remedy of rescission and restitution is potentially over compensatory. Kull argues that the threat of opportunistic behavior (that is, socially wasteful efforts to exploit an inefficient remedy to obtain an unbargained-for benefit) will be substantial for such contracts. The promisee can turn a loss into a gain by inducing breach by the promisor (or convincing a court that mutual uncooperativeness constituted or resulted from such a breach). By contrast, the perfect tender rule permits a buyer to behave opportunistically by unreasonably claiming that goods are defective when their market value has declined, but because the goods can be returned to Seller, the parties are spared the additional cost of a court proceeding to determine their value.

D. Calculation of Expectation and Reliance Damages

13. A Categorization of Approaches to Calculating Damages

There is consensus that the expectation measure is in most circumstances superior to reliance or restitution damages. A separate but no less important question is how expectation and reliance are to be defined and measured in typical contractual settings. Parties’ valuations are often unknown to one another and to the court, and promisees have an incentive to overstate their valuations, making the calculation of expectation damages difficult in some settings. Cooter and Eisenberg (1985) present a very helpful categorization and analysis of alternative calculation methods. They identify five broad categories and note that the calculation of money damages in reported cases usually falls into one of these categories. They are:
Contract Remedies: General

(i) Substitute Price. Often there is a spot market for the contractual performance at the time and place that performance was due, most obviously if the performance consists of the delivery of a marketable commodity. In such an event, Buyer can respond to Seller’s breach by cover, or the purchase of the commodity on the spot market. (Seller can respond to a breach by Buyer by selling on the spot market.) The difference between the contract price and the price at which cover occurred or could have occurred is then a measure of the cost of making Buyer (or Seller) indifferent between the contract and the substitute performance. We should note, however, that the substitute price measure can be overcompensatory when a promisee chooses not to cover but instead to sue for the difference between the contract price and the spot price. That choice itself suggests that the promisee may value the commodity at less than its market price.

(ii) Lost Surplus. When cover is unavailable, Buyer’s expectation can be thought of as the lost consumer surplus from the contract. In our ongoing example, if Buyer cannot cover, he loses the difference between his valuation of the machine ($300 or $375, depending on reliance) and the $250 contract price. The analysis of Seller’s lost producer surplus from Buyer’s breach is analogous. The lost surplus measure is feasible only when a court can obtain credible evidence of the promisee’s valuation.

(iii) Opportunity Cost. If a market exists for the performance, Buyer could have entered into a contract to buy the machine from any one of a number of competing sellers. The value to Buyer of the best alternative contract available at the time of the contract with Seller is an important component of his reliance. This value cannot be measured objectively because Buyer did not enter into this hypothetical contract and we do not know whether the hypothetical contractual party would have performed. Assuming that the probability of performance of the alternative contract is high, however, then the difference between the spot price at the time and place of breach and the price of the foregone contract is a good measure of reliance (augmented by any out-of-pocket expenditures in reliance on the contract with Seller). In a competitive market, the next-best price and the contract price should be the same, and the opportunity cost measure will equal the substitute price measure (a conclusion consistent with Fuller and Perdue’s conclusion that expectation and reliance damages are equal in a competitive market).

(iv) Out-of-Pocket Cost. This is the amount of reliance investment, less any salvage value of that investment. Out-of-pocket cost is the most common measure of reliance damages; a more complete measure of reliance damages is out-of-pocket cost plus opportunity cost.

(v) Diminished Value. So far we have ignored partial performance. In the real world, however, performance is often rendered but is defective or incomplete. In such cases an appropriate measure of Buyer’s lost expectation is the
difference between Buyer’s valuation of the promised performance and his valuation of the actual performance.

As these alternative methods of calculation should make clear, the measure of damages is usually straightforward and uncontroversial where cover is possible. The accepted measure of damages in such cases is the difference between the cover price and the contract price, which is easy to apply and provides appropriate incentives regarding the decision to perform or breach. The difficult questions arise when there is no perfect substitute for the performance (or there is room for debate about whether the substitute is adequate) or where the manner or timing of the breach causes harm that cannot be remedied by cover. We will provide two examples of cases that arise frequently and that have been the much discussed in the literature, in which there is debate over the appropriate means of measuring the non-breaching party’s expectation.

14. Example 1: Anticipatory Repudiation

Common law judges and scholars initially found anticipatory repudiation - a definitive statement by a promisor, made prior to the time for performance, that he intended to breach - extraordinarily vexing. Some concluded that any such statement must be without legal effect; the performance was due on a particular date and breach could therefore only occur on that date (Williston, 1901). Courts eventually came to the view that the promisee could treat the repudiation as a breach (Hochster v. De La Tour), but found it more difficult to decide how damages should be measured. The most famous early case, Missouri Furnace v. Cochrane, held that the appropriate measure was the difference between the contract price and the spot price at the time specified for performance. The Uniform Commercial Code, by contrast, encourages prompt cover, presumably in the futures market. As noted by Jackson (1978), the legal literature on anticipatory repudiation from the early part of this century is voluminous.

Jackson argued that in applying the Uniform Commercial Code’s provisions on cover to anticipatory repudiation, courts should fix damages at the difference between the contract price and the futures price at the time of repudiation. He noted that the Missouri Furnace method is systematically overcompensatory. Imagine, for example, that Seller breaches a contract to supply a commodity in the future and that the spot and futures prices at the time of repudiation are higher than the contract price. Over a large number of contract breaches, however, the spot price at the time of performance will sometimes be higher, and sometimes lower, than the contract price (in present value terms). Whenever it is lower, Buyer will not bring a damages action because he has been made better off by the breach. He is under no obligation to share this gain with Seller. When the spot price is higher than the contract price, Buyer will
recover the difference between the two. Averaged over a large number of contracts, buyers in the aggregate receive more than would be required to make them as well off as they were under the contract. Awarding the difference between the contract price and the futures price, by contrast, puts each buyer in the position he occupied prior to the repudiation and at a lower average cost to sellers.

We might simplify Jackson’s argument by noting that the Missouri Furnace rule replaces a forward contract by an option with a strike price equal to the forward price. Because the value (prior to expiration) of an option with a strike price of $X$ is always greater than the value of a forward contract with a contract price of $X$, the Missouri Furnace damage measure is overcompensatory.

15. Example 2: The Lost-Volume Seller

Sellers in a competitive market have often argued that the Substitute Price measure of damages, which awards them the difference between the contract price and the spot price, is undercompensatory. In many instances, there is little or no difference between the contract price and the spot price, and accordingly the damage award is trivial. Sellers contend, however, that they are not ‘made whole’ by selling in the spot market; the seller had the capacity to sell to both the substitute buyer and the original buyer at the market price, and the breach reduced their sales volume by one unit. Thus in place of two sales and two profits, they have received only one sale and one profit. Courts have often awarded the so-called ‘lost-volume seller’ an amount of damages equal to its ordinary profit on one sale. In the well-known case of Neri v. Retail Marine Corp., Retail Marine, a dealer in boats, agreed to sell a boat to Neri at a fixed price. Retail Marine ordered the boat from the manufacturer but Neri repudiated the contract. Retail Marine sold the boat to another customer for the same price and successfully sued Neri for the profit it would have made on the sale to him. The court concluded that Retail Marine, as a dealer, had an ‘inexhaustible’ supply of boats, and Neri’s breach deprived it of a profitable sale.

There is a substantial law and economics literature on the lost-volume seller. An early contribution appeared in an anonymous student-written comment (Anonymous, 1973). The comment noted that in a perfectly competitive market, each seller would choose output by equating marginal cost with demand and the demand curve would be presumed horizontal. At the chosen output, the firm’s marginal cost would be rising and therefore any additional sale would be at a cost in excess of the price. Because the seller could not, in fact, satisfy additional buyers at the market price, the breach and resale would create no ‘lost volume’ in a perfectly competitive market. A seller with market power (that is, one facing a downward-sloping demand curve) might be able to make additional sales at a profit. However, by hypothesis, such a seller
could eliminate the ‘lost volume’ by reducing its price and making an additional sale. Thus the standard contract price minus cover price measure would fully compensate such a seller.

Goetz and Scott (1979) provide an additional argument against awarding lost profits to the retailer who has market power. They note that the breach removes the breaching buyer as a competing seller. The buyer presumably breaches because it no longer wants the good at the contract price. In lieu of breaching, however, the buyer could complete the purchase and then resell the good. This resale, if made in the same market in which the retailer operates, shifts the demand curve facing the retailer to the left by one unit. Once again, if we compare the retailer’s position after the breach to its position assuming no breach but resale by buyer, there is no lost volume.

Goldberg (1984) disputes Goetz and Scott’s analysis. He first argues that the observation that a non-breaching buyer could sell in competition with the retailer is unrealistic. In fact, he argues, the buyer, lacking expertise, would have to engage the services of a retailer. The retailer’s usual markup is a reasonable estimate of the fee the retailer would charge for his services. Accordingly, the award of lost profit to the retailer approximates the result that would obtain if the buyer purchased and resold.

Goldberg also argues that it is inaccurate to say that the retailer ‘saves’ the marginal cost of a sale when the original buyer breaches and then incurs that marginal cost when the substitute buyer appears. He contends that the retailer’s cost of servicing an additional buyer consists principally of the cost of ‘fishing’ for a buyer, or convincing the marginal buyer to purchase (represented, perhaps, by costs of advertising, wages paid to salespeople, and so on). That cost is irretrievably lost once a contract is concluded with the original buyer and must be incurred again in order to induce another buyer to purchase. More recently, Scott (1990) argues that Goldberg’s equation of marginal cost with the cost of ‘fishing’ is inaccurate; for some goods, the cost of delivery and preparation for delivery are significant, and those costs are not incurred twice when a buyer defaults. Cooter and Eisenberg (1985) provide an analysis similar to Goldberg’s, but focus on the seller with market power. They argue that many sellers hold price at a constant level reflecting expected demand and marginal cost over some period, rather than constantly adjusting price to reflect realized demand. Such sellers can lose volume in a particular period.

Goldberg also notes that consumer demand is decreasing in the damage measure. Accordingly, were the legal rule to shift suddenly from a substitute price damages measure to one awarding lost profits, the demand curve facing the retailer would shift downward, offsetting the benefit of the higher damage awards. Whether consumers and producers would prefer the resulting contract to one that provides only substitute price damages again depends on comparative levels of risk aversion.
It appears that the literature on the lost-volume seller is at an impasse. The identification of the best damages rule turns on complex and contestable claims about market structure. A better avenue of inquiry would be to pay attention to actual contractual practice. Many sellers of custom goods require non-refundable deposits, which in effect contracts for a lost-profits measure. Other sellers (such as many computer retailers) permit a buyer to return an item for a full refund for some period after delivery, which in effect contracts for an even more lenient approach than the substitute price measure. It seems likely that greater ground will be gained by empirical analysis of the characteristics of markets in which varying cancellation/return policies are used than by further refinements of the theoretical arguments.

E. Conclusions

16. The Puzzle of Overcompensatory Remedies and Some Suggestions for Further Research

Most of the prior analysis could be summed up as follows: when courts and contracting parties are well-informed about each party’s valuation of the contract, money damages measured by the promisee’s valuation (or expectation) provide reasonably good incentives for efficient pre- and post-contractual behavior. Problems arise, however, when there are significant informational asymmetries between the parties and/or between each party and the court. Such asymmetries raise two pervasive issues in contract law. The first is subjective value. The existence of potentially overcompensatory remedies such as specific performance, liquidated damages and restitution can be attributed to judicial recognition that money damages measured by the promisee’s expectation will sometimes undercompensate, because courts use objective indicators of value that may diverge from the promisee’s subjective valuation. Only a few brief attempts have been made, however, to explore subjective value as a unifying theme in contract remedies (see De Alessi and Staaf, 1989; Muris, 1983).

The second issue is opportunism. The possibility that a remedy, although designed to be perfectly compensatory, will in fact undercompensate (overcompensate) may encourage the breaching party (non-breaching party) to use the defect in the remedy to gain bargaining leverage over the other party. The risk of opportunism is the likely reason why courts have not responded to the problem of subjective valuation by instituting overcompensatory remedies across the board.

A worthwhile avenue for additional work would be a careful comparison of the ways in which courts have or have not managed to reduce the risk of opportunism across a range of remedial choices. A promising approach to this
question appears in the liability rule versus property rule literature. When neither party knows the other’s true valuation of the contract, each has an incentive to over- or understate his valuation in an attempt to capture as much as possible of the gains from contract modification or cancellation. The result is to make agreement more costly. The costs imposed by asymmetric information, which we will call ‘bargaining costs’, are a subset of the cost of reaching a deal. The key question is whether the choice of remedy affects bargaining costs.

A specific application to liquidated damages is offered by Talley (1994). He uses the mechanism design branch of game theory to analyze the effects of different enforcement rules on bargaining costs, concluding that enforcement of liquidated damages that exceed actual damages \textit{ex post} creates significant bargaining costs. By refusing to enforce penalty clauses, courts may make it more likely that the parties will bargain to an efficient outcome. The argument is unique in offering a plausible economic justification of the \textit{ex post} component of the liquidated damages rule.

Ayres and Talley (1995), employ game theory to argue that bargaining costs are generally lower under liability rules than under property rules. The intuition is as follows. Going back to our contract between Seller and Buyer, imagine that Seller wishes to breach, and believes Buyer’s valuation of the contract to be uniformly distributed on the interval [\$300, \$400]. Consider a rule that provides for damages of \$500 in response to Seller’s breach. Buyer’s offer to rescind the contract for a payment of \$400 would provide Seller with no new information - Seller already knows that Buyer’s valuation is no greater than \$400. Now consider a rule providing for damages of \$350. Buyer might now conceivably offer to cancel the contract in return for a payment from Seller (if Buyer’s valuation is less than \$350), or it might offer Seller a payment to forego breach (if Buyer’s valuation is more than \$350). Thus the type of offer that Buyer makes conveys information about its valuation and ameliorates the bargaining costs resulting from asymmetric information. Johnston (1995) offers an analogous argument to show that bargaining costs can be lower under a ‘standard’, in which an entitlement is dependent on a discretionary judicial determination, than under a ‘rule’, in which the entitlement is more precisely defined.

Kaplow and Shavell (1995) criticize Ayres and Talley’s analysis on the grounds that it is not a marginal analysis. They argue that in most contexts in which bargaining is impossible or prohibitively costly, liability rules will dominate property rules for the reasons outlined in our discussion of expectation damages above. Thus for liability rules to dominate property rules where bargaining is possible does not prove that they generate lower bargaining costs; the latter point would be proved conclusively only if liability rules dominate property rules to an even greater extent where bargaining is possible than where it is impossible.
It is perhaps unfortunate that the game-theoretic analysis of bargaining costs has been used principally to analyze the relative efficiency of liability and property rules. The more general question is the design of remedies that will create optimal incentives for the parties to reveal their actual valuations or other private information about the state of the world. A liability rule (that is, a rule under which the damage award may be greater or less than the promisee’s valuation) might create superior incentives compared to a property rule (that is, one under which the damage award is at or beyond the endpoint of the promisee’s valuation), but we should be able to make similar comparisons between different liability rules. The discussion to date has covered liability and property rules generally, whether located within the law of property, torts and contract. There is accordingly room for a more focused look at bargaining costs in contractual settings, with an additional emphasis on \textit{ex ante} mechanisms other than judicially-crafted damage rules that might help to reduce \textit{ex post} bargaining costs.

Acknowledgements

Paul G. Mahoney thanks Eric Talley and two anonymous referees for helpful comments.

Bibliography on Contract Remedies: General (4600)


**Cases**