Cadillac Contracts and Up-Front Payments: Efficient Investment Under Expectation Damages

Aaron S. Edlin
Cadillac Contracts and Up-front Payments: Efficient Investment Under Expectation Damages

Aaron S. Edlin
University of California, Berkeley and National Bureau of Economic Research

This article shows that up-front payments can eliminate the overinvestment effect identified by Shavell (1980), by controlling which party breaches a contract. At the same time, "Cadillac" contracts (contracts for a very high quality or quantity) can protect against underinvestment due to Williamsonian holdups. This combination provides efficient investment incentives when courts use expectation damages as a remedy for breach. The expectation damages remedy is therefore well-suited to multidimensional but one-sided investment problems, in contrast to specific performance, which is well-suited to two-sided but unidimensional investment problems.

1. Introduction

The overinvestment caused by a default legal remedy of expectation damages has become a textbook result in law and economics. Polinsky (1989: 37), for example, writes:

The expectation remedy generally leads to too much reliance because it gives the relying party the value that would have been created by the reliance investment if the contract had been performed.

This insight can be traced to Shavell (1980) and later Rogerson (1984). Their essential idea is that the expectation damages remedy promises full compensation to the victim of breach. Fully insured, the potential victim undertakes risky actions that offer low social returns but high damages. This leads Cooter and Eisenberg (1985: 1467) to propose that damages be limited to what they would be under efficient investment.

This essay expands upon earlier research (Edlin, 1993: Section 3.3). I thank Tai-Yeong Chung for his extensive comments and also thank two referees, Eric Emch, Mario Epelbaum, Benjamin Hermalin, Stefan Reichelstein, William Rogerson, Alan Schwartz, Joseph Stiglitz, Robert Wilson, and participants at the 1994 meetings of the American Law and Economics Association, the winter 1995 meetings of the Econometrica Society, and Northwestern University's 1994 Micro Week Conference.

© 1996 by Oxford University Press. All rights reserved. 8756-6222/96/$5.00
This article argues that restricting expectation damages is unnecessary if the parties choose an appropriate contract and if the law is otherwise accommodating. The overinvestment effect arises if we assume that the investing party will be the victim of breach, but we argue that the identity of the breach victim should not be treated as exogenous. An "up-front payment" before performance can control which party chooses to breach and thus eliminate the overinvestment effect.

The parties still may face an underinvestment problem, due to Williamsonian holdups (Williamson, 1975), but we show that this too can be eliminated if the parties sign what we call a "Cadillac contract," a contract for a very high quantity or quality. The combination of up-front payments and Cadillac contracts can induce efficient investment in a variety of contracting situations, as long as only one party must invest and the courts have enough information to form an unbiased estimate of damages.\(^1\)

The intuition behind combining up-front payments and expectation damages is simple. The party who makes a sufficiently large up-front payment will not want to breach, since finishing performance requires only a commensurately small subsequent payment.\(^2\) Therefore, the other party will commit any breach.

Controlling who breaches is critical under an expectation damages remedy, because the breacher gets the residual that remains after paying the victim compensatory damages. If the breacher is also the investor, he will invest efficiently, maximizing this residual.

As an example, consider the not-so-fictional story of an economics department near Chicago. The department's building is undergoing substantial remodeling, remodeling that will definitely be completed in a year and that might be completed earlier. In the meantime, the department rents space in a commercial district on a one-year lease. Although the landlord sets the rent high, he provides an up-front payment as an inducement to rent the space: the first two months are free.

Given the high rent, the landlord has no desire to terminate the tenancy (after the first two months) since no other tenant will offer to pay higher rent. However, the department may desire to breach the contract and move out of the office space if the remodeling nears completion before the lease terminates.

---

1. Williamson (1983) previously argued that there was a strong connection between up-front payments and relationship-specific investments. Whenever investments are relationship-specific, the investor risks a "holdup"—that is, he risks losing the returns to his investments in negotiations conducted after the investments are sunk. Adequate protection must be provided somehow, and Williamson (1983) argued that up-front payments will protect investment returns in situations where it is impractical to enforce contracts in courts. (Courts may be an impractical option either because litigation is prohibitively expensive or because one party is unreliable and may flee from the court's jurisdiction or become judgment-proof due to insolvency.) What we find here is that such payments are useful even in the polar case when courts are costless and when there is no fear of a judgment-proof defendant.

2. Implicit in this comment is the assumption that the breaching party cannot sue. This assumption is central, and this essay provides examples of situations where it holds and where it does not. Where it does not hold, policy conclusions follow.
In fact, the department will move out at the very time that moving is efficient, assuming that it must pay the landlord damages calculated to compensate him for the breach (expectation damages). Moreover, the department will make efficient (multifaceted) ex ante investment decisions: it will invest optimally in hastening the remodeling, and will also “settle into” the downtown building to just the right extent, making appropriate expenditures to decorate, move books, and buy new business cards and stationery; expenditures that take into account the likely length of stay.

The combination of the expectation damages remedy, the up-front payment, the high rent, and the length of the lease is critical to the efficiency of this contract. The up-front payment induces the department to accept the high rent. The high rent, in turn, ensures that the alternative-rental value of the office space stays below the rent. Otherwise, the department might not unilaterally move out early even if moving were efficient; and since the lease gives the department the right to stay for the full year, the landlord would then have to bribe the department to move. The more “settled” it is, the larger the bribe would need to be. This means the department would have a (bribe-seeking) incentive to “over-rely” on the lease, that is, to overinvest in the specific investment of settling in. The overinvestment problem would worsen if the landlord could unilaterally breach, terminating the tenancy and paying the tenant expectation damages.

The one-year lease in the example above is what we call a Cadillac contract, because the remodeling is certain to be completed in a year. If the remodeling could take longer, the department might underinvest, for fear that investment returns would be lost in future negotiations to extend the lease—that is, in a Williamsonian holdup.

The example illustrates how the parties can control their own fate and avoid both under- and overinvestment by choosing an appropriate contract. This furthers the theme of Edlin and Reichelstein (1995), who study both expectation damages and specific performance, and the theme of Chung (1991) and Aghion, Dewatripont, and Rey (1994), who study specific performance. The contractual solution presented here differs significantly from the solution in those three articles, however. They all assume the parties can eliminate the full insurance problem by promising to trade an intermediate quantity, so that an investor sometimes receives more and sometimes less than the marginal social return to investment. Unlike such balancing contracts, the combination of Cadillac contracts and up-front payments ensures that the investor always gets the full marginal social return. This will be advantageous if investment decisions are multidimensional, such as when the economics department above needs to decide its expenditures to speed the remodeling as well as its expenditures to settle into the downtown offices.

This article spells out explicitly how two parties can use Cadillac contracts and up-front payments to induce one of them to invest efficiently. A surprising number of legal doctrines turn out to become involved in this result. In particular, how the courts view the duty to mitigate damages, what they think constitutes unjust enrichment, and under what circumstances an up-front payment is an illegal penalty can all affect the outcome. This article emphasizes
how the courts can choose legal doctrines to facilitate the parties’ solving their contracting problem.

Section 2 presents the model. Section 3 illustrates the approach in the simple models of Shavell (1980) and Rogerson (1984), and Section 4 extends it to settings where both parties’ valuations are uncertain and where one party makes multidimensional choices about investment and trade. The analysis in these sections assumes that the court is able to assess the damages to the victim of breach. However, Section 5 shows that what is really needed for the analysis to apply is that the court observe a signal of damages and make an unbiased assessment. Section 6 concludes and discusses situations where the results presented here do and do not apply.

2. The Model

We model trade between a buyer and a seller, both of whom are risk neutral. At date 1, they have the opportunity to sign a contract. At date 2, the seller must make an investment decision. The parties’ essential contracting problem is that this investment is not contractible and must be made before date 3, when some uncertainty about the value of trade is resolved.\(^3\) Trade, or perhaps breach of the contract, occurs at date 4, after which any outstanding claims are settled or go to court. The parties are fully and symmetrically informed at this stage about the buyer’s valuation and the seller’s cost.

This timeline is illustrated in Figure 1, and a more detailed description of each date follows.

2.1 Date 1: The Contract

We consider a fixed-price contract to trade \(q\) at a price \(p\) to be paid at the time of trade. The buyer also pays the seller an amount \(T\) up front, which may be viewed as compensation for the seller’s investment.

---

3. Problems might actually occur even if investment were contractible. Suppose, for instance, that the parties contracted for the efficient investment. Would this induce efficient investment? That depends. It would if the damages for breaching on the investment were extremely high. Nevertheless, since courts usually enforce only limited damages, a serious inquiry is required. Finally, it is unclear how a court would think about the damages to one party from another’s failure to make appropriate investments.
Definition 1. An up-front payment is anything valuable delivered at the time a contract is signed (other than the promise of payment for the goods). It might represent money paid, another profitable trade, or the signing of a separate profitable agreement.

Sections 3 and 4 show that under an appropriate legal regime involving an expectation damages remedy, the parties can sign a fixed-price contract that provides optimal investment incentives. The parties have an incentive to sign such a contract since it maximizes joint surplus and since this surplus can be divided arbitrarily with the up-front payment.

We consider a setting where the seller may supply many goods and services to the buyer in various quantities or qualities. Let \( q \) represent a list of the goods and services delivered, chosen from some set \( Q \), and let \( v(q, \theta) \) represent the value to the buyer of \( q \) in contingency \( \theta \).

Assume \( Q \) has a maximal element \( \bar{q} \) with nonnegative value, that is,

\[
\exists \bar{q} \text{ s.t. } \forall \theta, q, \quad v(q, \theta) \leq v(\bar{q}, \theta), \quad \text{and } v(\bar{q}, \theta) \geq 0. \tag{1}
\]

Definition 2. A Cadillac contract is a contract to trade a maximal element \( \bar{q} \).

We will show that a Cadillac contract is efficient if it is combined with a suitable up-front payment. Cadillacs are large, luxurious cars, traditionally at the top of GM's line. Accordingly, a Cadillac contract is a top-of-the-line contract for as large a quantity or quality as is generally efficient to trade. In fact, we have defined an extreme form of Cadillac contract involving the most valuable trade possible. If the Department of Defense orders 50 jet fighters of extraordinary capabilities, this may be a Cadillac contract. It will function as one if any renegotiations will be for jets with fewer capabilities and/or for fewer jets. Similarly, when a moving company promises to deliver all your belongings to your new residence intact, this is as valuable as their performance could be. (Those who have moved will know that their performance occasionally falls below this high standard.)

2.2 Date 2: Investment

At date 2, the seller can buy \( m \) assets, denoted by \( S \in \mathbb{R}_+^m \), which may affect the cost of performing \( q \). (Each of the \( m \) assets may vary in scale or quality since each is associated with a real number.) The assets cost \( I(S) \), where \( I: \mathbb{R}_+^m \rightarrow \mathbb{R} \). Let the ex post cost of producing \( q \) be \( c(S, q, \theta) \). This cost represents the most economical method of producing \( q \) with \( S \) given \( \theta \). In some contingencies, this may involve selling as salvage some of the assets embodied in \( S \); in others, all of \( S \) may be fully used in production. It is worthwhile keeping in mind that the function \( c \) is a reduced form, which may capture sophisticated deployment decisions.\(^4\)

\(^4\) As an example, let the set \( A \subset \{1, \ldots, m\} \) represent the indices of the assets put to alternative use. Then the assets in alternative use can be represented as \( \sum_{i \in A} S_i e_i \), where \( S_i \) is the \( i \)th element of \( S \) and \( e_i \) is the \( i \)th unit vector \((0, \ldots, 1, \ldots, 0)\). Suppose, the value derived from the assets put to alternative use is \( \bar{a} (\sum_{i \in A} S_i e_i, \theta) \). Suppose further that the costs of production of \( q \) using the
2.3 Date 3: Uncertainty

The contingency \( \theta \) may affect both the seller's cost \( c \) and the buyer's value \( v \). The literature on investment trade models generally assumes that some important uncertainty remains unresolved until after the time at which investment choices must be made.\(^5\) This timing is the essence of the contracting problem, because otherwise the parties could contract after all uncertainty is resolved but before investment and could promise to trade whatever turns out to be efficient. Such a contract would induce efficient investment.\(^6\)

2.4 Date 4: Trade, Breach, and Beyond

Any performance, breach, or renegotiation occurs at date 4. The legal regime, together with the date 1 contract, provides discipline and ultimately determines the parties' payoffs. In this article, we study the standard expectation-damages remedy for breach.

**Definition 3.** *Expectation damages* are the amount that makes the victim of a breach exactly as well off as she would have been if the other party had performed.

The expectation formula compensates the victim of breach for economic damages as measured ex post. (Nonlawyers beware: expectation damages have little to do with an expectation operator.)\(^7\) We express damages algebraically in the following sections. In Sections 3 and 4, we assume the court observes whatever is necessary of cost and valuation to accurately calculate damages. Section 5 considers the case where the court observes only a signal of damages.

It is necessary now to point out a number of ambiguities regarding breach and damages that the previous literature has not squarely confronted, and which are not settled by simply saying the legal regime involves an expectation damages remedy.

In particular, we know that a breach victim will be entitled to sue for damages. But what if the victim benefits from the breach so that damages are actually

\begin{align*}
\text{assets not in } A & = \bar{c}((\sum_{i \in A} S_i e_i, q, \theta)). \text{ Then, we would have} \\
\mathcal{C}(S, q, \theta) & = -\max_A \left[ \bar{c}((\sum_{i \in A} S_i e_i, \theta)) - \bar{c}((\sum_{i \in A} S_i e_i, q, \theta)) \right].
\end{align*}

6. This proposition is not quite as straightforward as most presume, however. Although the result is immediate if the contract could have a sufficiently high nonrenegotiable damage term so that performance is ensured, some argument would still be required if damages were limited and/or if renegotiation were possible after investment.
7. If Buyer Bob needs wheat and contracts with Seller Sally on November 1 for the delivery of 100 bushels of wheat on December 1 for $100/bushel, and Sally fails to deliver, damages will depend upon the price Bob must pay on December 1 for wheat. If Bob must pay $120/bushel, his damages are $(120/bushel - 100/bushel) \times 100 \text{ bushels} = 2,000$. 


negative? The breach victim won’t want to sue in such a case. The important question is whether the breacher can sue for damages.

Another question is: What happens if both parties try to breach the contract? The doctrinal answer begins by pointing out that a contractual duty must exist in order to be breached, and one party’s material breach may discharge the other’s duty, so that the other party has no further obligation. This implies that the court’s analysis would turn on a factual investigation of timing, asking whether party A failed to deliver before or after party B canceled the order. This determination could produce difficult questions, but as this analysis points out, such nettlesome questions can be avoided if the first question is answered appropriately.

Lastly, we should ask how the law treats lump-sum payments made at the contracting stage. Will they affect the answer to the first question? That is, if the contract breacher made such a payment, will he be entitled to a refund? The answer to all these questions is: It depends. Like most interesting legal questions, answers vary with the facts of the case and the jurisdiction. This essay does not provide an exhaustive description of the law as it is. Instead, it suggests answers to these questions that allow the parties to solve their contracting problems. The answers are not far-fetched, however, and generally accord with the outcomes in a significant subset of cases.

The difficulty with purely positive analysis that describes the economic consequences of the law as it is, is that the law does not exist as a singular entity, and even if we ignore this fact or focus on some well-defined area of the law, results can hinge upon so many particulars as to preclude economic analysis. On the other hand, purely normative analysis runs the risk of becoming divorced from the institutional constraints the law poses. Here, we walk a fine line, exploring what the law should be but tethered by a sense of realism that prevents us from running too far from what it is.

3. Discrete Trade

This section restricts the model to the more standard one of discrete trade and one-dimensional investment. Thus, \( Q = \{0, 1\} \) and represents no-trade and trade. The investment expenditure is simply \( S \in \mathbb{R} \), and \( I(S) \equiv S \). Only the buyer’s value of trade is uncertain.\(^8\)

For the purpose of the restricted model, we simplify notation as follows:

- \( v(\theta) \): buyer’s value of trade. The value of \( q = 0 \) is normalized to 0.
- \( c(S) \): cost of production given investment expenditure \( S \).
- \( a(S) \): alternative-use or salvage value of asset if there is no production (denoted as \(-c(S, 0, \theta)\) in the general notation of Section 2).

---

\(^8\) In Shavell (1980) and Rogerson (1984), the buyer invests and the seller has uncertain costs. This difference is unimportant here. However, when trade is not discrete, as in Section 4, it becomes important. There we explicitly present a solution method that involves a broad duty to mitigate damages. On the other hand, if the buyer invests, a parallel analysis would involve the contract being “divisible,” a concept we elaborate later.
Production and trade are efficient ex post if and only if the value to the buyer is sufficiently high that

\[ v(\theta) - c(S) \geq a(S). \]  
(2)

Otherwise, it is better to sell the investment as salvage for \( a(S) \).

We assume that some optimal investment level or levels exist. Since both buyer and seller are risk neutral, any first-best level of investment satisfies

\[ S^* \in \arg \max_{S \geq 0} \int \max[v(\theta) - c(S), a(S)] dF - S, \]  
(3)

where \( F(\theta) \) is the cumulative distribution for contingencies \( \theta \).

This section shows that a Cadillac contract together with a suitable up-front payment will solve the parties' contracting problem. Note that in this restricted setting, a contract to trade the good (\( q = 1 \)) is a Cadillac contract, since \( Q \equiv [0, 1] \).

We assume that if either the seller or the buyer breaches the contract, the other party may choose to sue to get an expectation damages remedy imposed. In contrast, we assume the breaching party cannot sue for breach of contract; after all, the other party did not breach the contract. Furthermore, we assume that the breaching party cannot sue in what is called “quasi-contract,” and that the court will find no other reason to assist the breaching party. This latter assumption does not accurately reflect the legal system in some jurisdictions and cases. Where the assumption does not presently apply, our analysis generates a legal policy recommendation as well as a contracting recommendation. It suggests that courts should be skeptical of such suits in quasi-contract, and if they are not, that the parties should pay the up-front payment in a separate contract with separate consideration. These prescriptions eliminate a race to breach, and

---

9. The quantity \( S^* \) is socially optimal in that it maximizes the sum of the payoffs to the buyer and seller. \( S^* \) would not necessarily maximize the sum of payoffs to all three if some third party were behaving strategically—e.g., by pricing as a monopolist. See Chung (1995) and Spier and Whinston (1995) for such a framework.

10. The possibility of suing “in quantum meruit,” one type of “quasi-contractual” suit, is a long-established way to recover benefits conferred on another who the court deems would otherwise be unjustly enriched. Even where there is a contract and one party has clearly breached the contract, courts have often allowed that party to recover the value of benefits he conferred upon the victim of the breach. See, e.g., Britton v. Turner, 6 N.H. 481, 26 Am. Dec. 713 (1834), where the plaintiff breached a 12-month labor contract by working only 9 1/2 months, but was allowed to recover 9 1/2 month’s wages. However, courts sometimes refuse to allow such recoveries. See Berke & Co. v. Griffin, Inc., 116 N.H. 760 (1976), where the court notes that, “generally quantum meruit recovery will not be awarded where the conduct has been ‘wilful’ [sic].”

11. If quantum meruit recovery is allowed for the breacher, extreme perversities may result. Suppose that unjust enrichment means receiving more than one’s expectancy resulting from another’s breach, and suppose further that a contract breacher can always sue to prevent unjust enrichment. Then both buyer and seller will be in a race to breach when trade is inefficient, since the breacher would get the surplus generated by doing the efficient thing. Each will attempt to make announcements of his refusal to perform in advance of the other’s announcements. More peculiar still, each will assert in court that the announcement of the other did not amount to an anticipatory repudiation or breach! These Alice-in-Wonderland phenomena occur if the breacher can always “pin” the other
allow the parties to neatly overcome the overinvestment problem by controlling who breaches.

To be concrete about the sort of unjust enrichment suit that will prevent the parties from solving their contracting problem, consider the lease example in the introduction. Suppose that while the economics faculty is attending the winter American Economics Association meetings, the landlord removes their belongings and re-rents their offices. The faculty can sue for breach of contract to recover any damages they suffer, but they may suffer no damages: they may find other accommodations for less rent and not be much inconvenienced. Would the landlord, then, be able to sue them to recover the rent for the two free months they enjoyed at the beginning of their lease, or for their rental savings from their involuntary relocation? Were they “unjustly enriched”? In this analysis, we assume the answer is no: the landlord who breached has no cause of action. This assumption is probably realistic in this example. In other examples, where the structure of the contract and flavor of the breach differs, however, some jurisdictions will allow the breacher to sue. This analysis provides reasons not to.

This issue of whether the contract breacher can sue can be reformulated as a question involving the penalty doctrine. If a buyer leaves a deposit when he orders a good, is it punitive to make the deposit nonrefundable? Courts often view it so, though the analysis that follows provides a reason not to, since a nonrefundable deposit may be used to provide efficient incentives.

If the seller breaches, the buyer can sue for expectation damages equal to \( v(\theta) - p \), which would give him the same payoff as if the seller had performed. Since we follow Shavell (1980) and Rogerson (1984) in ignoring litigation costs, a breach victim will sue whenever damages are positive.\(^\text{12}\) For now, we also assume that the court has sufficient information to properly calculate damages.\(^\text{13}\)

No one forces the buyer to sue, so he will not sue if \( v(\theta) - p < 0 \). Thus, if the seller breaches, she must pay the buyer

\[
\max[0, v(\theta) - p].
\]

Likewise, if the buyer breaches, he must pay the seller

\[
\max[0, p - c(S) - \bar{a}(S)].
\]

This assumes that the buyer notifies the seller of the breach before she incurs the variable costs \( c(S) \). Again, the “max” takes into account the fact that no one forces the seller to sue. The seller’s expectancy is \( p - c(S) \), but \( \bar{a}(S) \) is party to his expectancy by suing for disgorgement of any surplus from not trading when that is efficient.

\(^\text{12}\) Ignoring litigation cost limits the descriptive power of this essay, but it allows us to further develop the influential benchmark case considered by both Shavell (1980) and Rogerson (1984).

\(^\text{13}\) As Shavell (1980) emphasizes, this requires the court to observe the value \( v(\theta) \) but not the state \( \theta \) nor the functional relationship \( v(\cdot) \). Section 5 of this essay considers imperfectly informed courts that cannot observe even \( v \).
deducted because the seller is obligated to "mitigate" her damages by selling the investment at its salvage value.\textsuperscript{14,15} Although the duty to mitigate damages can induce efficient actions ex post, this feature is inconsequential in a symmetric information context where renegotiation is possible. The duty to mitigate damages is nonetheless quite important for efficient investment incentives, and Section \ref{sec:general} explains that in a more general setting the duty to mitigate damages should be broader than the law often recognizes.

When contractual quantity is not an available instrument, the key to inducing efficient investment is to set the price low, so that it is the seller-investor who commits any breach (compare with Edlin and Reichelstein, 1995, where quantity is used as an instrument).

**Proposition 1.** Let $p = \inf v(\theta)$. Then the seller invests and breaches efficiently. However, the seller must be paid an up-front payment, or she will not agree to such a low price, unless trade is always efficient.

**Proof.** It makes little difference whether the buyer or seller gets the first report of $\theta$ and has the opportunity to be the breaching party. The contract is designed to ensure that only the seller has the incentive to breach. To see this, suppose the buyer does not breach, but stands ready, willing, and able to buy. Then either the seller performs and the buyer pays $p$ and enjoys $v(\theta)$ from consuming the good, or the seller breaches and must pay $v(\theta) - p$ in damages. Either way, the buyer gets payoff $v(\theta) - p$, which is nonnegative since $p = \inf v(\theta)$. Performance is therefore a weakly dominant strategy, since breach would yield the buyer at best 0, and even less if the seller sued. (Recall that we assumed the buyer cannot breach and sue.)

If the seller breaches, she will be sued. Her payoff will be the salvage value minus damages, that is, $a(S) - [v(\theta) - p]$. If the seller performs, she receives $p - c(S)$, so she breaches if and only if\textsuperscript{16}

$$a(S) - v(\theta) \geq -c(S).$$

\textsuperscript{14} If the buyer wants to breach, he should announce his intention early, repudiating the contract. Otherwise his damage payment might increase by $c(S) + a(S)$ (if the seller produced and lost her opportunity to salvage her investment).

\textsuperscript{15} The duty to mitigate generally bars a plaintiff from recovering costs incurred after cancellation of the contract. A canonical example is *Rockingham County v. Luten Bridge Co.*, 35 F.2d 301 (4th Cir. 1929), where Luten Bridge Co. continued building a bridge after notice of cancellation was given by Rockingham County. Rockingham successfully argued that it owed Luten only the "damages which the company would have sustained, if it had abandoned construction at that time." Note, however, that some expenses incurred after cancellation are recoverable. For instance, advertising expenses are generally recoverable when they might reasonably increase salvage value; in a typical view, such expenses are recoverable even "where the result is an aggravation of the damages rather than a mitigation," if "expenses are the result of a prudent attempt to minimize damages." See, for example, *Mr. Eddie, Inc. v. Ginsberg*, 430 S.W. 2d 5 (Tex. Ct. App. 1968), a breach-of-contract case, where a wrongfully dismissed employee under a three-year employment contract spent $1,340 in an unsuccessful job search.

\textsuperscript{16} Ties, where $a(S) - v(\theta) = -c(S)$, leave the seller indifferent.
This inequality matches inequality (2), so the parties trade if and only if trade is efficient ex post. Since expectation damages is a liability rule allowing unilateral breach, no renegotiation is necessary for efficient trade.

The seller's investment problem is to choose

\[ \hat{S} \in \arg \max \int \max[p - c(S), \tilde{a}(S) - v(\theta) + p]dF - S. \]  

(7)

This optimization differs from the social optimization given by (3) only by adding \( p - v(\theta) \) in each contingency. Investment incentives are unaltered, so we have shown that choosing \( p = \inf v(\theta) \) induces efficient investment and breach.

Nonetheless, the buyer must pay the seller an up-front payment to induce her to sign such a contract, unless trade is always efficient. That is, consider a case where trade is sometimes inefficient: that is, for some \( \theta' \), \( v(\theta') - c(S^*) < \tilde{a}(S^*) \). Since by construction \( p \leq v(\theta') \), substitution yields \( p - c(S^*) < \tilde{a}(S^*) \). Without an up-front payment, the seller would be better off avoiding the contract, investing \( S^* \), and selling it as scrap.\(^{17}\)

Thus, the seller invests efficiently if the buyer and seller agree to trade at the favorable price given in Proposition 1. However, the buyer must often pay the seller some up-front payment \( T \) to convince her to sign the contract under these favorable terms. In fact, an up-front payment is necessary in cases where Rogerson–Shavell overinvestment occurs, that is, in cases where trade is sometimes inefficient.

A similar efficiency result holds if the buyer invests before the seller's costs are known.

Proposition 2. Suppose the buyer invests, and let the valuations be \( v(S) \) and \( c(\theta) \). Then if \( p = \sup c(\theta) \), the buyer invests and breaches efficiently.

Proof. The proof is left to the reader, since it is essentially the same as when the seller invests.

The efficient contracts in Propositions 1 and 2 are Cadillac contracts combined with up-front payments. Recall that since \( Q = \{0, 1\} \), any contract to trade \( q = 1 \) is a Cadillac contract. This method of solving the parties' contracting problem stands in contrast to the method described in Edlin and Reichelstein (1995). In their analysis of one-sided investment, they advocate signing a contract to trade some intermediate quantity \( q \in (0, 1) \), or if this is not possible, having the contract specify that the parties trade only in certain contingencies (where those contingencies arise with appropriate probability). The difference in approach results because Edlin and Reichelstein restrict attention to a case where \( p \) is high so that it is the noninvestor who breaches, just as in Shavell (1980) and Rogerson (1984); the balancing approach presented

17. If the optimal investment \( S^* \) is not unique, read \( S^* \) as the seller's choice from among the optima.
by Edlin and Reichelstein would not be efficient in the more general context of Section 4.

The legal assumption that the contract breacher cannot sue may appear punitive to the breacher. For instance, if the buyer invests and the seller fails to deliver, the buyer keeps any up-front payment. As a result, the buyer may get more than his expectancy out of the deal. We previously pointed out that in cases such as our lease example, it is realistic to think that the party who made the up-front payment has no recourse if she breaches. Another example is a CD club whose membership consists of an initial enrollment purchase of 12 CDs for one cent together with the promise to pay a high price for some number of additional CDs at a later date. If the club later refused to sell the additional CDs (even at the high price), it is doubtful that the initial discounted sale would be voided, or that the buyer would owe additional money.

On the other hand, suppose a buyer puts down a deposit on a couch, only to cancel the purchase subsequently. Under the Uniform Commercial Code Section 2-718, a court may insist that some of the deposit be returned to the buyer, if it is deemed to be unreasonably large “in light of the anticipated or actual harm caused by the breach, the difficulties of proof of loss, and the inconvenience or nonfeasibility of otherwise obtaining adequate remedy.” Yet, suppose the couch were custom-made and that the maker had to order special fabrics and foams that have a low resale value. In consideration of this specific investment, the customer makes what we call an up-front payment; or perhaps he simply buys the materials from the maker, and a separate contract is written to build the couch, so that if the buyer cancels the construction (breaching), he should be entitled to no refund on his purchase of materials (or payment for investment). When the buyer is seen to be purchasing the investment in a separate deal, the up-front payment should not be refundable to a breaching buyer unless the investment were not made. If the up-front payment were made in a separate agreement, the penalty doctrine might not apply across the agreements.

Again, we remind the reader that the above analysis is valuable even in those cases or jurisdictions where the breacher could sue. In that case, the analysis provides a policy reason why a breacher should not be able to sue. Allowing an up-front payment to stand even when it appears punitive lends the parties considerable contracting power. Notice, though, that efficiency is not driven simply by allowing what in some cases is tantamount to high stipulated damages; efficiency comes from the combination of these “high damages” preventing one party from breaching, and the expectation damages remedy giving the other party appropriate breach and investment incentives.

3.1 No Up-front Payments: Overinvestment

For the sake of contrast, we show below that if up-front payments are impossible (say, because of liquidity constraints), then the overinvestment problem returns. Shavell (1980) considers up-front payments but only payments made to the noninvestor. Unless an up-front payment is made to the investor-seller, she will not agree to the low price that drives Proposition 1; instead, she will demand a
price \( p \) such that
\[
p - c(\tilde{S}) \geq \overline{a}(\tilde{S}),
\]
where \( \tilde{S} \) is whatever investment the seller chooses given the contract. If the inequality were violated, the seller would always lose money. With such a "high" price, as we shall see, the seller always performs and the buyer becomes the potential breacher.

In order to derive the overinvestment effect, we must impose additional structure on the model. Assume:
\[
\begin{align*}
(A1) & \quad -c'(S) - \overline{a}'(S) > 0; \\
(A2) & \quad S^* \text{ as defined by (3) is unique, and } S^* \in (0, \overline{S}); \\
(A3) & \quad \int_{\theta: u(\theta) - c(S^*) < \overline{a}(S^*)} dF > 0.
\end{align*}
\]
Assumption (A1) states that marginal investment lowers production costs by more than it increases salvage value. Assumption (A2) is self-explanatory, and (A3) states that when \( S = S^* \), salvage is efficient with positive probability. These assumptions about functional form and the uniqueness of \( S^* \) were unnecessary to prove Proposition 1, but we need them here to replicate the traditional overinvestment result.

**Proposition 3.** Assume (A1), (A2), and (A3). Then, unless the investor-seller receives an up-front payment, she overinvests.

**Proof.** As argued above, when up-front payments to the seller are impossible, the seller will agree only to a price \( p \) such that \( p - c(S) \geq \overline{a}(S) \), when evaluated at the investment level she intends to take. Consequently, if the buyer stands ready to perform, the seller also performs and receives \( p - c(S) \). (The seller’s other alternative is to salvage, which yields \( \overline{a}(S) \) if the buyer does not sue, and even less if the buyer does sue.) If the buyer breaches, the seller will sue to recover damages
\[
D = p - c(S) - \overline{a}(S).
\]
After salvaging her investment for \( \overline{a}(S) \), the seller’s net payoff is her expectancy \( p - c(S) \), just as when the buyer performs. The seller’s investment problem is therefore
\[
\max_{S} p - c(S) - S.
\]
To prove that the seller overinvests, compare this problem to the joint surplus maximization problem (3) and
(i) observe that (A3) guarantees that when \( S = S^* \), salvage is efficient with positive probability;

---

18. This assumption corresponds to Assumption 3 in Rogerson (1984), except that Rogerson wrote it as a weak inequality and proved weak overinvestment.

19. This proposition is essentially like those of Shavell (1980) and Rogerson (1984).
(ii) subtract the social planner's objective function from the seller’s and observe that the difference has a positive derivative with respect to $S$, because of observation (i) and assumption (A1);


Intuitively, the seller is overcompensated for her incremental investment by the expectation damages remedy, because when trade is inefficient, she receives the full cost savings that incremental investment would yield during production, $-c'(S)$, instead of the lesser "social return" from salvage, $\bar{a}'(S)$. Overinvestment results when up-front payments to the investor are impossible and the value of trade to the noninvestor is the source of uncertainty. However, as we have seen, sufficiently large up-front payments made to an investing seller combined with a commensurately low subsequent price ensure efficient investment. Since the buyer has already almost fully or even fully performed, he will not breach no matter what his value. This leaves the decision to breach in the hands of the investor. Under an expectation damages remedy, a combined breacher-investor gets the full residual surplus ex post, and so has the incentive ex ante to invest to maximize this surplus. This principle applies to more general contexts than the Rogerson–Shavell model, as shown in the following two sections.

4. General Settings and a Broad Duty to Mitigate

We now consider the full model, where $Q$ is arbitrary and $S$ is multidimensional. We show that the Cadillac contract and up-front payment solution extends to this context. This reveals a significant advantage of expectation damages over specific performance.\textsuperscript{20}

Suppose the parties have written a Cadillac contract to "trade" $\overline{q}$ with the entire payment $T$ made up front, so that no payment is required when $\overline{q}$ is performed.\textsuperscript{21} Consider the situation after the investment $S$ is made and contingency $\theta$ realized. What will the seller supply? Suppose the seller supplies $q \neq \overline{q}$. We assume that the buyer can accept $q$ and still sue for damages of

$$ D = v(\overline{q}, \theta) - v(q, \theta), $$

as provided in Section 2-714 of the Uniform Commercial Code. Since it is a Cadillac contract, the expression in (12) is nonnegative.

Will the buyer be entitled to higher damages if he rejects $q$, or should he be induced to mitigate damages by accepting $q$? Typically, the duty to mitigate damages does not extend to require a buyer to accept inferior performance. This may be a sound rule if courts have difficulty evaluating the consequences of nonconforming delivery. In some cases, however, the parties may want a

\textsuperscript{20} Edlin and Reichelstein (1995) show that a balancing technique works under either expectation damages or specific performance in a model like the one considered in Section 3. That technique, however, does not provide efficient incentives when the investment problem is multidimensional, as it is here.

\textsuperscript{21} It is possible to have some payments made at performance, but that analysis requires a detailed discussion of the importance of a divisible contract.
broader mitigation rule to solve their multidimensional contracting problem. If mitigation requires accepting partial or inferior performance, the buyer can recover at most $v(q, \theta) - v(q, \theta)$, which motivates the following definition.

**Definition 4.** Suppose a buyer has paid up front and been promised $q$. Under a **broad duty to mitigate damages**, if offered $q$, the buyer can collect only $v(q, \theta) - v(q, \theta)$ in damages.

A broad duty to mitigate prevents the breach victim from securing more than the benefit of his bargain by threatening to refuse $q$ unless the seller agrees to pay larger damages. Such opportunism would tend to distort the seller’s investment decision.

This duty to mitigate is broader than usually obtains. Consider, for instance, *Shirley MacLaine Parker v. Twentieth Century Fox*, 3 Cal. 3rd 176, 89 Cal. Rpts. 737 (1970). The California Supreme Court held that Shirley MacLaine Parker did not need to accept Twentieth Century’s offer to star in a western titled “Big Country, Big Man” to mitigate damages for Twentieth Century’s breach of the contract in which she was to star in a musical titled “Bloomer Girl.” Alternatively, consider the Uniform Commercial Code Section 2-601, under which the buyer has the right to “reject the whole” if “the goods or the tender of delivery fail in any respect to conform to the contract.” Moreover, under Section 2-711 a “rightful” rejection by the buyer leaves him with the same remedies as if the seller had not performed at all. On the other hand, the buyer does have some duty to accept an inferior performance: for example, the Uniform Commercial Code Section 2-508(2) allows a seller to cure her breach by delivering conforming goods late, but not too late. The next proposition suggests, however, that a broader duty to mitigate damages would have some advantages.

**Proposition 4.** A Cadillac contract to supply $q$ for up-front payment $T$ leads to efficient investment and breach when an expectation damages remedy is applied and the buyer has a broad duty to mitigate damages.

**Proof.** Joint surplus is maximized by solving the following iterative program: Ex post, given assets $S$, and the realized contingency $\theta$, choose $q^*$ to solve

$$Z^*(S, \theta) \equiv \max_{q \in \Omega} v(q, \theta) - c(S, q, \theta). \quad (13)$$

Ex ante, joint surplus is maximized by choosing $S$ to solve

$$\max_S \{E[Z^*(S, \theta)] - I(S)\}, \quad (14)$$

where $E$ denotes the expectation operator.

Under a Cadillac contract, with payment $T$ made entirely up front, the seller

---

22. As other sections in the Uniform Commercial Code make clear, these rights of rejection should not be read overly broadly, but they do apply to nonconformities that go to "the heart of the agreement."
solves a nearly identical problem. Suppose the seller delivers \( q \neq \bar{q} \). Given the broad duty to mitigate damages, the buyer’s rights will be limited to damages of \( v(\bar{q}, \theta) - v(q, \theta) \) regardless of whether he rejects or accepts \( q \). Since the contract is a Cadillac contract, \( v(\bar{q}, \theta) - v(q, \theta) \geq 0 \), so the buyer will in fact demand, and if necessary sue for, payment of \( v(\bar{q}, \theta) - v(q, \theta) \). Therefore, given \( S \) and \( \theta \), the seller chooses \( q \) ex post to solve

\[
W(S, \theta) = \max_{q \in \mathcal{Q}} [-D(q, \theta) - c(S, q, \theta)],
\]

where damages \( D(q, \theta) = v(\bar{q}, \theta) - v(q, \theta) \).

The seller’s ex post objective function is the same as the joint surplus less \( v(\bar{q}, \theta) \), so the seller’s optimal choices of \( q \) match those from surplus maximization. The seller’s ex ante investment problem is to solve

\[
\max_S \{E[W(S, \theta)] - I(S)\}.
\]

Since \( W(S, \theta) = Z^*(S, \theta) - v(\bar{q}, \theta) \), the seller’s ex ante choices of \( S \) similarly match those under joint surplus maximization, and the Cadillac contract to trade \( \bar{q} \) for a payment made entirely up front leads to first-best efficiency.

The expectation damages remedy allows us to find a contract \( \bar{q} \) that leads to efficient choice of assets, efficient use of assets, and efficient performance or breach. Compare this with the specific performance remedy. Under a specific performance remedy, when the investment decision is a one-dimensional decision, fixed-price contracts can be efficient. These results require a continuous contractible variable \( q \) and a number of assumptions about how investment affects valuations (Chung, 1991; Aghion et al., 1994; and Edlin and Reichelstein, 1995). None of these assumptions is needed for the approach above. Moreover, even when those assumptions are met, the first-best cannot be implemented under specific performance when the investment decision is multidimensional; for instance, Edlin (1993) shows that multidimensional decisions contain a bias toward investing in overspecialized assets instead of assets with higher values in alternative uses. No such biases exist under an expectation damages remedy with up-front payments and a Cadillac contract. Provided the parties can arrange for one party to make all the ex ante investment expenditures and decisions together with ex post breach decisions, expectation damages is an ideal remedy. Of course, an expectation damages remedy may require the courts to have more information than specific performance requires.

The intuition behind Proposition 4 is as follows. A large up-front payment made by the noninvesting buyer to the investing seller ensures that the buyer performs the contract. A broad duty to mitigate damages, including accepting nonconforming performance, leaves the breach decision unilaterally in the hands of the investor. Otherwise some renegotiation would be necessary, because the buyer could refuse a nonconforming tender; such renegotiation typically would involve splitting the returns to investment, which could distort investment incentives. The Cadillac contract eliminates contingencies where overperformance is efficient, so that \( v(q^*, \theta) > v(\bar{q}, \theta) \). Such contingencies
also require bilateral negotiations and typically entail sharing the gains from extra performance, again distorting investment incentives. Williamson (1975) and Edlin and Reichelstein (1995) explain how such opportunism would reduce investment incentives.

In a model where the buyer invests, instead of the seller, the analysis proceeds similarly. For this “nonbalancing” approach, we would then want the buyer to effectively dictate the level of performance and capture the residual created beyond the seller’s expectancy. Two features are necessary. A Cadillac contract would ensure that the seller has a duty to deliver performance at least as valuable as is efficient. Second, if the buyer decides he wants to cancel (the inefficient) part of performance, this must not discharge the seller’s remaining duty to perform the noncancelled parts of the contract. This requires that the court construe the contract as “divisible”—essentially as a number of separate smaller contracts. Such a construction allows the buyer to cancel inefficient parts of the contract and still have the seller deliver the efficient parts or else pay compensatory damages. This construction thereby avoids negotiation and surplus-splitting, allowing the buyer-investor to capture all residual, just as the broad duty to mitigate allowed the seller-investor to capture all residual in Proposition 4.

5. Imperfectly Informed Courts

This brief section demonstrates that the analysis of Section 3 is essentially unchanged when courts do not observe the true damages. What is critical is that the courts impose an unbiased measure of damages. This point, which is made in Edlin 1993, has proven sufficiently provocative to warrant elaboration.

Suppose that when true damages are $D$, the court observes and imposes $\tilde{D}$, where

$$\tilde{D} = D + \varepsilon, \quad (17)$$

and $\varepsilon$ represents the court’s error or misperception. Assume that the parties cannot anticipate the direction of the court’s bias (i.e., the expected value of $\tilde{D}$ is $D$).

Since we are ignoring litigation costs, if the seller breaches, the risk-neutral buyer will still bring a case whenever the true damages $D$ are positive. Given the low trading price $p$, the buyer always brings the case. The seller’s payoff from breach is

$$\bar{a}(\delta) - \tilde{D} = \bar{a}(\delta) + p - v(\theta) - \varepsilon, \quad (18)$$

23. In some cases where the buyer invests, a divisible contract is unnecessary. For instance, in our lease example the economics department (the buyer) invests. The department has a property right to stay for the duration of the lease; moreover, it can move out early without needing any agreement from the landlord (though it would owe damages). Thus the law allocates decision rights over $q$ to the lessee (at least when $q \leq \bar{q}$), which means that no divisible contract is needed.
and from performance is

\[ p - c(S). \] (19)

The risk-neutral seller breaches whenever the expected payoff from breach exceeds that from performance. This leads to the same breach rule as when the court’s decision is predictable. Moreover, the seller’s expected payoff is the same as before, so the seller has the same investment incentives. Thus, the unpredictability of court outcomes does not alter the analysis as long as the court imposes an unbiased judgment \( D \) of damages, which establishes that Proposition 1 holds even if the court does not observe true damages, as long as the court observes some informative signal and enforces an unbiased estimate of damages. This goes a small way toward addressing the criticisms of Schwartz (1993: 406) about the information requirements of other mechanisms that might be used as problem-solving defaults. However, this still assumes more than the mechanism design literature does (see, e.g., Hermelin and Katz, 1993, or Rogerson, 1992).

Of course, courts introduce a number of biases. Principal among these is the “certainty” requirement. Courts will not guess at damages if their information is woefully inadequate to the task. For instance, the “new business” rule “prohibits recovery of lost profits resulting from a breach of contract that has prevented the plaintiff from establishing a proposed new business, on the ground that profits in such cases are too speculative” (Fuller and Eisenberg, 1990: 267). In parallel fashion, in Freund v. Washington Square Press, 34 N.Y. 2d 379, 314 N.E. 2d 419, 357 N.Y.S. 2d 857 (1974), the Court of Appeals of New York found that Freund was entitled to only nominal damages for Washington Square’s breach of a contract to publish his book; sales were unpredictable, so only token damages could be awarded. Such a result biases damages downward. Therefore this model lends support to what Fuller and Eisenberg (1990: 267) identify as a “definite trend toward abrogating” the new business doctrine, and similarly to the flavor of the Official Comment in the Uniform Commercial Code to Section 1–106, which states that “compensatory damages are often at best approximate: they have to be proved with whatever definiteness and accuracy the facts permit, but no more.” For a case in line with these trends, see Fera v. Village Plaza, Inc., 396 Mich. 639, 242 N.W. 2d 372 (1976), where the Michigan Supreme Court upheld a $200,000 jury verdict in favor of the plaintiffs who had signed a 10-year lease in order to open a “book and bottle” shop. (The plaintiffs were unable to take possession because the defendant subsequently leased the space to a third party.)

6. Implications

The central result of this analysis is that the expectation damages remedy need not lead to overinvestment in relationship-specific assets. In fact, even when investment and trade are multidimensional, we found that all decisions were first-best. The conventional wisdom in the law and economics literature about distorted investment results from assumptions about which party breaches. If up-front payments to the investor are possible, the party who breaches is not
determined exogenously by which party has uncertain valuations. The parties themselves control who will want to breach in low-trade contingencies. When the noninvesting party makes a sufficiently large up-front payment, he will want to carry out the contract even if his valuation proves unfavorable. Any breach will be made by the investing party, who then has efficient investment and breach incentives under the expectation damages formula.

Thus, when one party needs to make a specific investment, it is natural that the other should make some payment early. An up-front payment is desirable even when neither party will flee or become bankrupt and the legal system costlessly enforces contracts.

From this analysis we learn a number of lessons about policy. Notice first that we assumed that the breaching party could not successfully sue. This argues that courts should be skeptical of unjust enrichment claims or at least should honor contracts that give up the right to sue for unjust enrichment. Viewed differently, it suggests that courts should not invoke the penalty doctrine to return deposits.

Another lesson is that the courts should take a broader view of the duty to mitigate—or at least should stand ready to take a broader view if requested in the contract. Such a view of mitigation was critical to the efficiency of Cadillac contracts in Section 4. When the buyer is the investor, the critical issue becomes the divisibility of the contract.

A third lesson is that specific performance is not always the best remedy when specific investments are involved. Edlin and Reichelstein (1995) argue that specific performance is superior when both parties make investments (but where the investments are simplistic). Yet when the investment and breach decisions can all be allocated to one party, the expectation damages remedy is superior because multifaceted investment can be efficient. Therefore, when parties neglect to specify a remedy, as they often do, expectation damages is a good default rule in one-sided investment problems. Its application implies that the ex ante incentives of the contract breacher were appropriate (at least the incentives arising from the given contingency). These first-best incentives may very well be gotten at the expense of incentives for the victim of breach (see Edlin and Reichelstein, 1995), but this will not matter if the victim does not make substantial reliance decisions.

The victim of breach may seek specific performance, in order to use the threat of forcing inefficient performance to increase her total payoff. The courts, however, should be loathe to grant it unless the victim can show some reliance or specific investment of her own. This proposal might be viewed as giving a new interpretation to the old rule that specific performance is granted only when the “legal remedy” of damages is “inadequate.”

The inadequacy of damages is traditionally viewed as meaning that damages are difficult to measure, such as with unique chattel. Section 5, however, showed that difficulty of measurement per se is not particularly important. The model indicates that the victim of breach should have to show that the remedy is “inadequate” because it does not provide parties in her position with appropriate incentives ex ante. Expectation damages would not provide
appropriate incentives to the victim of breach. When the breach victim also must invest, then, specific performance is better, as it provides more balanced incentives (see Edlin and Reichelstein, 1995).

Finally, we should ponder the implications of Section 5, which explained that the accuracy of the expectation measure is not so important as its lack of bias. This suggests that it is worthwhile to consider the efficacy of certain legal rules that tend to bias damages. For instance, the rule that damages must be "certain" biases damages downward. In accord with modern trends, damages should be approximated as well as the evidence allows. If estimates become wildly high, as some worry, the appropriate cure is to lower awards, not to categorically eliminate such damages.

Of course, the lessons from the theoretical analysis presented here must be taken with a grain of salt. Many factors were not considered. For instance, there are sensible reasons why a duty to mitigate should not be too broad. For one thing, it may require the court to know a great deal about valuations. Also, up-front payments will not solve all problems if the investing party can go bankrupt or can otherwise flee with the payments. In practice, courts are costly, and damage awards may diverge in a predictable way from true damages. These various complications are inevitable. One hopes they are not discouraging but simply provide fertile territory for new research.

References


