

Draft: February 22, 2004

EMBEDDED OPTIONS AND THE CASE AGAINST COMPENSATION IN CONTRACT LAW

*Robert E. Scott**
*George G. Triantis***

Abstract

Compensation is the governing principle in contract law remedies. The principle shapes the doctrines that specify the consequences of breach: particularly the default provision for expectation damages and the restriction on the parties' ability to stipulate damages. Yet, the compensation principle has tenuous historical, economic and empirical support. Moreover, a promisor's right to breach is a subset of a family of termination rights that serve important risk management functions. The termination rights, in turn, can be characterized as call options on the subject matter of the contract. Where a buyer can terminate her obligation to purchase a good, she effectively holds a call option defined by its option price and exercise price. The option price is the buyer's sunk investment, which may be in the form of the prospective damages liability under the contract, and the exercise price is the additional sum needed to acquire the good. The contracting choice among the many pairs of option price and exercise price is subject to a variety of factors identified in this essay. In light of the heterogeneity among optimal option prices, we argue not only against the penalty rule restriction on liquidated damages, but also against having even a default damages provision to begin with. We propose that parties be forced to agree explicitly with respect to all termination rights, including breach damages, either by the threat of specific performance of their contemplated exchange or by the court's refusal to enforce contracts that fail to do so.

INTRODUCTION

Compensation is the governing principle in contract law remedies. This principle shapes the key doctrines that specify the consequences of breach. Expectation damages, the default measure of damages, aim to put the promisee in the position she would have occupied had the promisor performed, while specific performance is available at the option of the promisee when the court believes that money damages are inadequate to compensate for her loss. Although parties may agree to liquidated

* David and Mary Harrison Distinguished Professor of Law, University of Virginia

** Perre Bowen Professor of Law and Horace W. Goldsmith Research Professor, University of Virginia

Draft: February 22, 2004

damages, contract doctrine instructs them to abide by the compensation norm. We argue in this article that, although the compensation principle has a profound influence on contemporary contract law doctrine, it has tenuous historical, economic and empirical support. Its evolution in the common law was more accidental than deliberate; compensation is virtually ignored in the theoretical analysis of efficient contract design and it plays little role in the contracts expressly negotiated by commercial parties and by consumers. As a result of an unfortunate turn in history, contract damages are analogized to compensation for wrongs, and this has impeded the efficient evolution of both default remedies and the regulation of liquidated damages.

It is well known that contract damages effectively give the promisor an option between performing the promise or breaching and paying damages.¹ Over the past decade, legal scholars have begun to analyze contract remedies explicitly in terms of options.² But contract damages are only a subset of the broader category of termination rights that give one party an option to walk away from the contemplated exchange.³ A firm offer or unilateral promise, for example, grants the promisee such an option. Broad warranties, such as satisfaction-or-your-money-back provisions, give buyers similar

¹The classic statement is by Oliver Wendell Holmes:

Nowhere is the confusion between legal and moral ideas more manifest than in the law of contract. Among other things, here again the so called primary rights and duties are invested with a mystic significance beyond what can be assigned and explained. The duty to keep a contract at common law mean a prediction that you must pay damages if you do not keep it—and nothing else.

Oliver Wendell Holmes, *The Path of the Law*, 10 Harv. L. Rev. 457, 462 (1897).

²Robert E. Scott, *The Case for Market Damages: Revisiting the Lost Profits Puzzle*, 57U. Chi. L. Rev. 1155 (1990); George G. Triantis, *The Effects of Insolvency and Bankruptcy on Contract Performance and Adjustment*, 43 Univ. Tor. L. J. 679; Alexander J. Triantis and George G. Triantis, *Timing Problems in Contract Breach Decisions*, 41 J. Law & Econ. 163 (1998); Paul G. Mahoney, *Contract Remedies and Options Pricing*, 24 J. Legal Stud. 139 (1995); Victor P. Goldberg, *The Net Profits Puzzle*, 97 Colum. L. Rev. 524 (1997); Victor P. Goldberg, *Discretion in Long-Term Open Quantity Contracts: Reining in Good Faith*, 35 U.C. Davis L. Rev. 319 (2002); Avery Wiener Katz, *The Efficient Design of Option Contracts: Principles and Applications*, (mimeo 2004).

³See TAN *infra*.

options. Requirements, output or installment contracts grant one party substantial discretion to determine the contract quantity. And a contract may provide that one party has the right to terminate, to cancel, to renew, or to return or redeem goods.

Either or both parties to a contract typically enjoy the right to terminate at some cost. For the purposes of our analysis and argument, we focus on the option held by a buyer of goods or services. The buyer holding an option has the right to avoid the exchange by paying either a termination fee or damages. Several previous articles characterize the right to breach as if the buyer held a put option on the agreed upon exchange that she could exercise by paying damages to the seller.⁴ In this essay, however, we analyze the buyer as effectively incurring an obligation to pay damages in exchange for a call option on the subject matter of the contract. The *exercise price* of the option is the difference between the contract price and the damages liability. The buyer's prospective liability for damages is effectively the price of the call option (the *option price*), as if the buyer makes a nonrefundable deposit or payment for the call option and pays an additional price to exercise it. The option price is a function of the exercise price: the higher the exercise price, the less valuable the call option and, generally, the lower the option price. We label the exercise price of the buyer's call option, x , and the option price, d (to remind the reader of the link with damages), and we describe the option created by the termination provision by the pair (d,x) .

The price of an embedded call option is determined just as the price of any other product: it is a function of the option's value to the option holder, the cost to the option writer and the competitiveness

⁴ By exercising the put, the buyer, in effect, sells the contract good or service back to the seller for the stipulated damages liability. Triantis, *Effects of Insolvency and Bankruptcy*, supra note –, at 680-4; Mahoney, supra note –; Triantis and Triantis, supra note –, at 168-9, 169 n15.

of the market in which they transact. In light of the great variety in the conditions under which parties contract for this option, it should not be surprising that commercial and consumer contracts contain a wide range of call option prices. These prices are rarely equivalent to the measure of the seller's expectation in a completed sale.

Consider the electronics store that sells television sets for \$400 and offers full refunds for any returns made within 30 days. This contract gives the buyer a free 30-day call option to purchase the television set for \$400. Recall that we characterize the right of termination as a call option described by the pair of an option price and exercise price, (d, x) , where the contract price is $P = d + x$. This contract's pairing of option price and exercise price is $(0, 400)$. The option is valuable to the buyer because she is uncertain as to the value of the television set to her family. The retailer in this case bears significant costs in accepting returns, including the cost of receiving, inspecting and reselling the returned goods, often through a discount outlet or internet sale. Nevertheless, options to return without charge are frequently given to buyers for free. Such "free" options are remarkably common in both commercial and consumer contracts.⁵ They are particularly interesting for our purposes because these contracts make no attempt to compensate the seller for losses it suffers when the buyer walks away from the contemplated exchange. These return options provide free insurance to buyers, which is puzzling at first blush because of the adverse selection and moral hazard problems they are likely to create. But buyers do pay an exercise price in the cost of bringing the product back to the store. If the

⁵It is inaccurate to conclude that the cost of the option is included in the overall contract price because the buyer does not pay this price if she walks away from the option. Where the seller does not charge the buyer for an unexercised option, the seller recovers the cost from buyers who exercise the option. This cross-subsidization leads to adverse selection and moral hazard issues discussed later in the paper. Another example of a free option is the right of consumer borrowers to prepay their loans or mortgages without compensating their lenders for consequential losses that might be caused, for example, by declines in the market rate of interest. Similarly, a borrower who defaults is liable to pay the accelerated balance owing, but not compensation for the lender's foregone opportunities when market rates have dropped since the loan.

Draft: February 22, 2004

consumer would pay up to \$25 to avoid the trip, we might characterize the return option as (25, 400), even though the \$25 is not payable to the store.⁶

Some buyers enjoy no call option because they pay (or owe) the entire purchase price and cannot escape this obligation by cancelling the transaction.⁷ Some retailers, for example, offer refunds for ordinary course sales but do not accept returns of sale merchandise. Although sale merchandise may be more difficult to resell, it is likely that the retailer is “undercompensated” in the former case and “overcompensated” in the latter. Additional examples of overcompensation can be found in the penalties consumers pay in other transactions. For example, economy fares on airlines are typically conditioned on a penalty of \$100 if the passenger chooses to not travel or to change her itinerary: if the fare is \$250, the passenger’s call option is (100, 150).⁸ The penalty applies even on flights that are overbooked and almost certain to be full.

Indeed, our casual observation reveals that the termination provisions affecting consumers regularly departs from the predictions of the compensation principle. Examples of such embedded options abound in commercial contracting as well. Retailers often have the right to return unsold

⁶If we think briefly about the seller’s option, we note that liquidated damages are often pegged significantly below the compensatory level. For example, repair and replacement are common remedies for defective performance. Contracts also limit recovery for packages lost by couriers or luggage lost by airlines. Recently, distributors of electricity have issued interruptible electricity contracts that allow the distributor some flexibility to interrupt electric service to commercial customers, which they exercise in times of demand spikes. Some contracts provide for financial compensation at the time of the interruption, but others provide for a payment in advance. Neither payment is intended to reflect the loss suffered by the customer. See Ross Baldick, Sergey Kolos and Stathis Tompaidis, Valuation and Optimal Interruption for Interruptible Electricity Contracts (Univ. of Texas Working Paper 2003).

⁷Of course, fresh start in bankruptcy (and limited liability of some organizations) creates the well known option held by debtors, which we set aside in this essay.

⁸Indeed, in the event of a change in plans, the passenger also may have to pay any increase in the fare from the time she bought the ticket.

Draft: February 22, 2004

merchandise to the wholesaler or distributor. Rights of return are common, for example, in the retailing of books, journals, newspapers, musical compact discs, jewelry, and cigarettes.⁹ Aircraft manufacturers permit purchasers to cancel orders or to change the type of aircraft even after the manufacturer has made significant investment in production.¹⁰

Why are these contracting patterns so much at odds with the compensation principle that governs contract law? In this article, we offer the explanation based on the risk management objectives of many contracting parties. It is well known that compensatory remedies insure the buyer against the seller's breach and thereby against the risk of fluctuations *in the seller's cost of performance*. Options created by termination rights provide another type of insurance that benefits risk averse buyers and contributes to the risk management objectives of business contractors.¹¹ The ability of a buyer to breach and pay expectation damages partially insures the buyer against decreases *in the value of the seller's performance*. The buyer can avoid the loss to the extent that it is greater than the seller's

⁹Eugene Kandel, *The Right to Return*, 39 J. L. & Econ. 329, 330 (1996).

¹⁰For example, Airbus allows buyers to choose between the aircraft in its family of A318, A319, A320 and A321 with a very short lead-time. These aircraft are built on the same production line and have many common components, so the decision as to which aircraft is built can be delayed. Airbus explicitly markets contract options and offers guidance to buyers in the valuation of these terms. John Stonier and Alexander J. Triantis, *Natural and Contractual Real Options: The Case of Aircraft Delivery Options*, in A. Micalizzi and L. Trigeorgis, eds., *Real Options Applications* 159-195 (1999). John Stonier, *The Change Process*, in Tom Copeland and Vladimir Antikarov, *Real Options: A Practitioner's Guide* 47 (2001).

¹¹In his Nobel lecture, Robert Merton stated: "When [an option is] purchased in conjunction with ownership of the underlying asset, it is functionally equivalent to an insurance policy that protects its owner against economic loss from a decline in the asset's value below the exercise price for any reason." Robert C. Merton, *Applications of Option-Pricing Theory: Twenty Five Years Later*, 88 Am Econ. Rev. 323, 336-7 (1998). In Merton's terms, the buyer in a contract "owns" the contract and holds a put on that asset. Although framing the termination right as insurance (or a put option) may fit better with the conventional language of contract theory, we prefer to speak of the equivalent call option (equivalent by virtue of put-call parity).

Draft: February 22, 2004

expected profit under the contract.¹² But if the buyer has the right to cancel and pay less than the seller's expectation, then the seller effectively insures the buyer against more severe decreases in the value of the exchange to the buyer. This insurance is efficient if the seller enjoys a comparative advantage in bearing the risk by hedging or diversifying against it. The seller's actions in performing the underlying contract itself might also reduce the variance in valuation. Through such contract options, therefore, buyers can manage some of the risks of their operations by outsourcing risk management to their sellers.

Most contract scholars appreciate that remedial provisions affect contract price: the higher the prospective damages liability, the less the buyer is willing to pay in the contract. The options analysis, however, frames the determination somewhat differently. The damages liability is the price of a call option and it depends on the exercise price of the option. For any given exercise price, the option price divides between the parties the surplus created by the option: namely, the difference between the option's value to the buyer and its cost to the seller. The parties should choose a pairing of option price and exercise price that maximizes this surplus. Given the close link between options and insurance, it should not be surprising that the optimal terms of embedded options are a function of considerations that determine insurance contracts: risk bearing capacity, adverse selection and moral hazard. Although we discuss the role of these considerations in the structuring of embedded options, our principal contribution is to highlight the heterogeneity of optimal terms, which is consistent with our observation of a wide variety of termination provisions in practice.

¹²As explained in Part II, this insurance against downside risk is a product of the abandonment option created by nature. Termination rights in a contract allocate between the parties the abandonment option that an integrated firm would enjoy in any venture. Compensatory damages promote the optimal abandonment of the parties' joint venture (what is conventionally known as efficient breach).

Draft: February 22, 2004

This diversity has important normative implications for the default rules of contract damages. The characterization of breach damages as an option price separates damages from the compensation principle and reinforces the criticisms of the penalty rule that have been raised by contract scholars.¹³ A theory of embedded options suggests that, in many contexts, a default rule may be no more appropriate with respect to damages than with respect to the price of any other good, service or contract term for which there is no established market price. In thick markets, courts should simply enforce the risk allocation specified in the contract by awarding market damages. Outside thick markets, we argue in favor of a bargain-forcing approach of judicial restraint. The courts could refuse to recognize any option by specifically enforcing the contract exchange unless the parties clearly state the price and exercise price of the option. Or, the courts might refuse to enforce the contract as a whole if it fails to make express provision regarding termination rights.

The article proceeds as follows. In Part I, we contrast the contemporary dominance of the compensation principle of contract damages against its tenuous historical roots, its marginal relevance in contract economics and its limited use in commercial and consumer contracts. Part II characterizes the termination provisions available to a buyer – including the right to breach and pay damages -- as call options that serve an important risk management function. We specify factors that explain the variations in the price paid for these options. Part III draws the implications for optimal default rules for breach of contract summarized above. The Conclusion summarizes and proposes an agenda for future research.

¹³ See TAN *infra*. The compensation principle is also responsible for leading courts as well as scholars to search in the wrong place for solutions to lost-volume seller cases by debating alternative measures of the seller's loss. See Scott, *The Case for Market Damages*, *supra* note – .

I. RECONSIDERING THE COMPENSATION PRINCIPLE IN CONTRACT LAW

A. The Dominance of the Compensation Principle in Contract Law Doctrine

The doctrinal view of contract breach is that it is a wrong that is remedied by compensating the victim for her loss.¹⁴ This view is reflected in the default provision for damages and the constraints imposed on the freedom of parties to stipulate other measures of damages by contract. The Restatement of Contracts, for example, states that “[t]he traditional goal of the law of contract remedies has not been compulsion of the promisor to perform his promise but compensation of the promisee for the loss resulting from breach.”¹⁵ Since the classic article by Fuller and Perdue in 1936, courts and scholars have typically compared three alternative bases for evaluating the promisee’s entitlement to compensation: restitution, reliance and expectation.¹⁶ The dominant measure is expectation damages:

¹⁴ See, e.g. Uniform Commercial Code § 1-106: “The remedies provided by this Act shall be liberally administered to the end that the aggrieved party may be put in as good a position as if the other party had fully performed...”; Restatement(Second) of Contracts §347. Although the compensation principle is firmly enshrined in the Code and the Restatement, we suggest in the next section that contract law embraced the compensation principle relatively late in its common law development.

¹⁵Restatement (Second) of Contracts, Intro. Note to Ch. 16, Remedies. See also, E. Allan Farnsworth, *Legal Remedies for Breach of Contract*, 70 Colum. L. Rev. 1145, 47 (1970) (describing the common view that the central purpose of contract damages is compensation); UCC § 1-106, Comment 1 (same). For Oliver Wendell Holmes, at least, the task of determining the proper remedies for breach is simply the gap filling analysis courts use to allocate other contractual risks . See, for example, the following from *Globe v. Landa, Cotton Oil Co.*, 190 U.S. 540 (1903):

When a man commits a tort, he incurs, by force of the law, a liability to damages, measured by certain rules. When a man makes a contract, he incurs, by force of the law, a liability to damages, unless a certain promised event comes to pass. But unlike the case of torts, as the contract is by mutual consent, the parties themselves , expressly or by implication, fix the rule by which the damages are to be measured. It is true that, as people when contracting contemplate performance, not breach, they commonly say little or nothing as to what shall happen in the latter event, *and the common rules have been worked out by common sense, which has established what the parties probably would have said if they had spoken about the matter.* (Emphasis added).

¹⁶Lon L. Fuller and William R. Perdue, Jr., *The Reliance Interest in Contract Damages:1*, 46 Yale L.J. 52 (1936).

Draft: February 22, 2004

the payment necessary to put the promisee in as good a position as if the promisor had performed.¹⁷ Over the past 150 years, the courts have expanded the availability of compensation for the promisee's consequential losses, subject only to relatively modest constraints of mitigation, foreseeability and uncertainty.¹⁸ Specific performance is available at the option of the promisee when money damages are thought inadequate to compensate for her loss.¹⁹

The dominance of the compensation principle explains the legal regulation of parties' ability to stipulate damages expressly in their contract. Although most of contract law provides default rules from which parties are free to contract away, remedial defaults carry heavier presumptive weight than other provisions. Contract breach terms may not deviate from the compensation principle. According to doctrinal statements, the only permissible ground for stipulating damages is the anticipated difficulty of measuring the promisee's loss. For instance, the UCC provides that stipulated damages must be an amount that is reasonable in light of the anticipated or actual loss caused by the breach and the difficulties of proving the amount of loss.²⁰ In practice, the courts strike down supercompensatory

¹⁷Restatement (Second) of Contracts §344, 347. "The initial assumption is that the injured party is entitled to full compensation for his actual loss." Intro. Note to Topic 2, Enforcement by Award of Damages.

¹⁸In a recent article, Richard Craswell observes that courts appear willing to soften or disregard these traditional limitations when they find that the breach was willful. Richard Craswell, *Against Fuller and Perdue*, 67 U. Chi. L. Rev. 99, 138-43 (2000). Craswell cites Arthur Corbin as noting that "a lesser degree of certainty will be required as against one whose breach is described as 'willful' or is motivated by malice or avarice than against one whose breach was due to misfortune and whose efforts to perform were honest and in good faith." Id. at 140. Craswell also speculates that courts may choose reliance damages as means of effecting some sharing of losses between contracting partners (for example, in response to an unforeseen contingency).

¹⁹Restatement (Second) of Contracts §359. See ROBERT E. SCOTT & JODY S. KRAUS, *CONTRACT LAW AND THEORY* 116-118, 992-95 (3D. ED. 2003).

²⁰UCC § 2-718; Restatement (Second) of Contracts §356 (same). The modern penalty rule found in the Code and the Restatement is derived from a line of common law cases invalidating any stipulated damages where the amount specified exceeded the "just compensation for the loss or injury actually sustained." *Jaquith v. Hudson*, 5 Mich. 123, 133 (1858). For discussion, see Charles J. Goetz & Robert E. Scott, *Liquidated Damages, Penalties and*

liquidated damages far more often than they police undercompensatory limitations on damages.²¹ In cases in which it is especially difficult to measure the promisee's loss, a court may order specific performance or an injunction even in the face of a contractual provision for liquidated damages.²² This additional basis for setting aside liquidated damages provides further evidence that the law regards measurement difficulty as the sole legitimate reason for deviating from the default of expectation damages.

B. Historical Roots of Compensation in Contract Breach Remedies

Although now firmly entrenched in doctrine, the compensation principle is a recent development in contract law. As we describe in some detail below, the courts did not demonstrate a strong inclination to compensate the promisee for the losses caused by the promisor's breach and they enforced the parties' agreement respecting the consequences of breach until the 19th Century. Therefore, our critique of the compensation principle does not attack a deep-rooted principle of contract law.

the Just Compensation Principle, 77 Colum. L. Rev. 554, 555-557 (1977). Restatement (Second) of Contracts § 347, Comment a states that "The parties to a contract may effectively provide in advance the damages that are to be payable in the event of breach as long as the provision does not disregard the principle of compensation." UCC § 2-718, Comment 1 states: "A term fixing unreasonably large liquidated damages is expressly made void as a penalty. An unreasonably small amount would be subject to similar criticism and might be stricken under the section on unconscionable contracts or clauses."

²¹ See e.g., 5 ARTHUR CORBIN, CONTRACTS, § 1068 (1964) ("public policy may forbid the enforcement of penalties against a defendant, but it does not forbid the enforcement of a limitation in his favor."); Justin Sweet, *Underliquidated Damages As Limitations of Flexibility*, 33 Tex. L. Rev. 196, 203-06 (1954). Although we would expect that undercompensatory liquidated damages are in fact at least as common as overcompensatory provisions, a review of reported cases between 1998 and 2003 reveals that the incidence with which courts strike down penalties far exceeds their rejection of undercompensatory damages provisions. [summarize results here]

²² Restatement (Second) of Contracts § 361, Comment a states:
Merely by providing for liquidated damages, the parties are not taken to have fixed a price to be paid for the privilege not to perform. *The same uncertainty as to the loss caused that argues for the enforceability of the provision may also argue for the inadequacy of the remedy that it provides.* (emphasis added)"

Draft: February 22, 2004

1. Forms of Action under the Early English Law of Contract: The Evolution of the Compensation Principle.

At early common law, there was no cause of action for breach of an informal (unsealed) executory promise.²³ The only actions available for breach of contract were the action for debt and the action in covenant (for instruments under seal). The notion of compensation was foreign to either action. The action for debt lay only for the recovery of a sum certain. One party was seeking relief for a debt that was due and owing, fixed by the parties prior agreement, and the court in no sense awarded compensation for breach of contract. For example, where a seller tendered goods to a buyer and the buyer refused to accept delivery, the seller was able to recover the purchase price and force the buyer to take delivery of the goods (for which title had passed under the contract).²⁴ Alternatively, if the buyer tendered the purchase price and the seller refused to transfer goods that were then available, the buyer had no action at law. Here, the buyer's only recourse was to bring an action in equity for specific performance since the remedy at law was inadequate.

The evolution of commercial exchange during the late middle ages increased the pressure on the legal system to supplement the self-enforcement mechanisms of the law merchant and medieval trade fairs with a legal mechanism that permitted promisors to make credible promises to strangers.²⁵ The response of the English common law courts was to recognize a promisee's right to recover in assumpsit for breach of promise. Assumpsit was originally a tort action ("he undertook") that

²³ JAMES BARR AMES, *LECTURES ON LEGAL HISTORY* 92 (1913); *Id.* at 92, 122-23; A. W. B. SIMPSON, *A HISTORY OF THE COMMON LAW OF CONTRACT* 47-48 (1986).

²⁴ AMES, *LECTURES ON LEGAL HISTORY* at 88-89.

²⁵ Avner Grief, *Informal Contract Enforcement: Lessons from Medieval Trade* in 2 *THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND LAW* 287 (Peter Newman, ed. 1998).

developed in response to the need to provide an action in tort for the negligence of a bailee or carrier for hire.²⁶

The principle of compensation that supported the action in assumpsit was thus a distinctly tort notion premised on the idea of ex post redress for a harm committed by the defendant.²⁷ Over time the action in assumpsit was seen as an attractive avenue for seeking recovery for promissory undertakings. Initially, a plaintiff was allowed to bring assumpsit only where the defendant performed his promise unskillfully (i.e., a carpenter who undertook to build a house for the plaintiff and performed poorly). Subsequently, the English courts held that a plaintiff could recover in assumpsit for the promisor's failure to act altogether. In such a case, the plaintiff could recover damages based upon the principle of compensation for the injury done (in this case the nonperformance of a promise to act).²⁸ The action in assumpsit for breach of promise thus lay for plaintiffs who had either conferred benefits and or undertaken action in preparation for performance in reliance on the defendant's promise. In either case, a plaintiff who was seeking compensation via assumpsit was asking for compensation under a theory of *reimbursement* for the loss of that which had been given (directly or indirectly) to the

²⁶SIMPSON, supra note — at 210-215. The traditional tort action of trespass on the case would not permit recovery for negligent bailment because the plaintiff could have been equally careless in entrusting a third party with his property. Id. In the *Humber Ferryman Case*, for example, the plaintiff alleged that defendant bailee *undertook* to carry his goods safely. The failure to perform this undertaking was the gravamen of the action, and, as was traditional in tort actions, the resulting injury to the plaintiff's property required *compensation*. 22 Lib. Ass., Edw. III pl.41 (1348).

²⁷ AMES, SUPRA NOTE -- at 130. A parallel line of cases permitted recovery in deceit for a false warranty for goods sold and delivered. This action was also, in its origin, a pure action in tort. Id. At 136-7.

²⁸ Id.

promisee.²⁹

2.. *The Enforcement of Executory Promises: Compensation as Market Damages.*

The final stage of the development of assumpsit was the recognition of the enforceability of a purely executory exchange of promises. For many years, the orthodox view among legal historians had been that this final stage of development matured in England sometime in the 16th or 17th centuries. But Morton Horwitz has argued convincingly that the enforceability of executory contracts occurred much later, at the end of the 18th century (in England) and the beginning of the 19th century (in America). This development coincided with a period of commercial expansion and with the development of markets for the sale of commodities and subsequently for the sale of stock.³⁰

As late as the 18th century, therefore, contract law was still dominated by the action in debt and a title theory of exchange. Recovery was based on the price of goods tendered or delivered (or, in the case of a breach by the seller, an action for specific performance in equity). Throughout the 18th century, exchange was not conceived in terms of future returns and thus expectation damages were not recognized by 18th century courts on either side of the Atlantic.³¹

²⁹ During this early period of the action in assumpsit, a plaintiff could bring an action for breach of promise independent of the doctrine of consideration and the concept of exchange. The early notion of *special assumpsit* (the contract action) did not require a quid pro quo as was required for an action for debt which was explicitly tied to the notion of exchange. AMES, supra note – at 143-144.

³⁰ Morton Horwitz, *The Historical Foundations of Modern Contract Law*, 87 Harv. L. Rev. 917 (1974).

³¹ Horwitz cites only two English cases in the 18th century that raise the issue of expectation damages. Horwitz, supra note – at 921. In *Fleureau v. Thornhill*, 96 Eng. Rep. 635 (C.P. 1776), the court limited the plaintiff to restitution damages, holding that “plaintiff could not be entitled to damages for the fancied goodness of the bargain which he supposes he has lost.” In the United States, only a few actions for breach of executory contracts were brought before the Revolutions. See, e.g., *Boehm v. Engle*, 1 Dall. 15 (Pa. 1767) where the seller was allowed to sue for the contract price of breach of a contract for the sale of land. *Id.* at 922.

Draft: February 22, 2004

This narrow conception of contract as enforcing only partially executed exchange transactions conceived of compensation in equally narrow terms. Chancellor Kent articulated this principle of contract damages as compensation being fixed by the jury “with a moderation agreeable to equity and good conscience and where the claims and pretensions of each party can be duly attended to...”³² During this pre-market period, courts almost universally declined to instruct juries on damages defaults or to revise damage judgments (whether excessive or inadequate). During the 18th and well into the 19th centuries, therefore, common law courts enforced only partially executed transactions, and saw the compensation principle as requiring either reimbursement or, where the promisee’s performance was tendered, specific performance. These courts did not assure a plaintiff “the value of his bargain,” but rather would specifically enforce the actual bargain that the parties had struck.

Executory contracts were not enforced in the United States until the decision in *Sands v. Taylor* in 1810.³³ Under the older rule, where a buyer breached a contract to purchase goods, the seller would have been required to tender the contract goods and sue for the contract price. But in this case, the seller covered on the market by reselling the goods to a third party and then sought damages based upon the contract-market differential. The court conceded that this was a case of first impression in America and granted the plaintiff market damages.³⁴ The subsequent emergence of a market damages default rule coincided with the increased use of fixed-price forward contracts for the delivery of commodities.³⁵ Disputes over stock transactions were also common during this period;

³²Seymour v. Delancy, 6 Johns. Ch. 222, 232 (N.Y. Ch. 1822).

³³5 Johns. 395 (N.Y. 1810).

³⁴ Id.

³⁵See e.g., *Shepherd v. Hampton*, 16 U.S. (3 Wheat.) 200 (1818).

courts awarded market damages for failure to deliver stock certificates in a rising market.³⁶ This link between commodities and stock transactions and contract law was thus the major step in the development of a market damages default rule for breach of thick market contracts. Contract thereafter became an instrument for managing the exogenous risks that affected changes in price in well-developed markets.³⁷

3. The Penal Bond as the 18th Century Alternative to Executory Contracts.

We now return to the 18th century and ask how commercial transactions could proceed in the face of a legal system that did not enforce executory contracts. To be sure, self-enforcement was widely relied upon by many commercial parties (using merchant juries and other extra-legal processes).³⁸ Ultimately, however, commercial parties sought legal mechanisms that would permit them to conduct exchange transactions and to make credible commitments to each other.

The most important of these forms was the *penal bond*. Initially, the bond developed as a debt instrument to circumvent restrictions on interest under rigid usury laws.³⁹ For example, a debtor would promise under seal to pay a creditor \$1,000 in six months. This promise was subject to an express

³⁶See e.g., *Groves v. Graves*, 1 Va. (1 Wash.) 1 (1790).

³⁷ *Id.* at 941. This principle of market damages was generalized in England in 1826 with the publication of the first treatise announcing a general rule of damages for failure to deliver goods:

In an action for assumpsit, for not delivering goods upon a given day, the measure of damages is the difference between the contract price and that which goods of similar quality and description bore on or about the day when the goods ought to have been delivered.

Horwitz, p.941. It was at this point that contract fully separated from property and courts, for the first time, granted promisees a property right in the contract itself.

³⁸Avner Grief, *supra* note – at –.

³⁹THEODORE SEDGWICK, *A TREATISE ON THE MEASURE OF DAMAGES* 392-93 (1847); THEOPHILUS PARSONS, *THE LAW OF CONTRACTS* 433-34 (1855).

Draft: February 22, 2004

condition subsequent that it would be null and void if the debtor paid the creditor \$900 (the amount of the loan) on the day following the execution of the sealed promise. The advantage of the formal instrument was that it was conclusively enforceable at its face value by an action in covenant and was an absolute obligation once the defeasing condition failed to occur.⁴⁰ Subsequently, the penal bond was expanded and made conditional on executory promises to perform a specific service or action or to deliver property at a time certain.

During the 17th and 18th centuries, courts of equity began to hear claims for relief from the bond where there was only a minor breach of the stipulated conditions. By the end of the 18th century, common law courts had adopted the equity rule of relief from the bond where the amount owed greatly exceeded the loss to the plaintiff from the breach of the condition. American courts followed this tradition and “chancered” these bonds.⁴¹ Nevertheless, according to Horwitz, at the beginning of the 19th century virtually all large business transactions took the form of two independent bonds, each of which stipulated damages for failure to perform an executory promise.⁴²

From the beginning of the 18th century, English courts adjudicating disputes over the terms of bonds had sought to distinguish between penalties—where they would grant relief from the bond in appropriate cases--- and liquidated damages, which the parties were free to stipulate in the bond without interference from the courts.⁴³ By the time of Lord Mansfield, English courts were predisposed

⁴⁰ A.W.B. Simpson, *The Penal Bond with Conditional Defeasance*, 82 L.Q. Rev. 392, 411-12 (1966); Sedgwick at 393.

⁴¹ Sedgwick, *supra* note – at 394.

⁴² Horwitz, *supra* note -- at 928.

⁴³ W. S. HOLDSWORTH, *A HISTORY OF ENGLISH LAW* 663 (1924).

Draft: February 22, 2004

to interpret damages provision in bonds as liquidated and enforceable. Mansfield announced as settled the proposition that “where the covenant is to pay a particular liquidated sum, a court of equity cannot make a new covenant for a man...”⁴⁴ Presumably, American courts adopted a similar presumption favoring liquidated damage clauses in bonds. In any event, at the beginning of the 19th century the number of bonds used to effect commercial transactions in America still greatly exceeded the number of contracts that sought to enforce mutual promises.⁴⁵ The dominance of bonds, bills of exchange and other sealed instruments meant that commercial parties had little reason to take their transactional disputes to common law courts. The ubiquitous use of bonds and the corresponding absence of markets thus caused a delay in the development of a modern default rule of contract damages. Because of its liquidated damages provisions, the bond delayed the development of a damages default until the mid-nineteenth century.

3. The Evolution from Market Damages to Expectancy.

During the early decades of the 19th century, the use of bonds declined in both England and America as parties turned to executory contracts and the damage rules that had developed in the context of forward commodities contracting. Clearly, liquidated damages in formal bonds were not well-suited to contracts premised on the allocation of market risk.⁴⁶ Executory contracts began to

⁴⁴ *Lowe v. Peers*, 98 Eng. Rep. 160, 162 (K.B. 1768). Lord Eldon declared in 1801 that he “could not but lament any supposed principle that even an enormous and excessive damage provision, to which the parties agreed, should be void as a penalty.” *Astley v. Weldon*, 126 Eng. Rep. 1318, 1321 (C.P. 1801).

⁴⁵ Horwitz, *supra* note -- at 929.

⁴⁶ See *Graham v. Bickham*, 2 Yeates 32 (Pa. 1795) (sharp market fluctuations justify a recovery for market damages in excess of the stipulation in a bond).

dominate the bond for most 19th century contracts between commercial parties.⁴⁷ In the early decades of the 19th century, courts enforcing executory contracts in commercial transactions generally limited damage awards to market damages or its equivalent. The market damages default that had developed in the futures market context was adopted for relation-specific investment contracts in thin markets as well. Sedgwick, writing in 1847 states that: “The law does not aim at complete compensation for the injury sustained; it seeks rather to divide than to satisfy the loss. In cases of contract,... the direct pecuniary damages...form the measure of relief.”⁴⁸ Thus, he concludes, “both the English and American courts have generally concurred in denying profits as any part of the damages to be compensated for [breach of contract].”⁴⁹ In short, prior to the decision in *Hadley v. Baxendale* in 1854, courts generally refused to take into consideration any damages that “remotely” resulted from the breach.⁵⁰ Generally that meant that plaintiffs were limited to market damages and generally they were denied recovery for other consequential losses.⁵¹ This principle was repeated early and often by American courts.⁵² Justice

⁴⁷Horwitz, supra note – at 932.

⁴⁸SEDGWICK, supra note – at 57.

⁴⁹ Id. at 69.

⁵⁰Remoteness, in turn, was interpreted to mean that contracting parties could only recover that which *both parties could have contemplated at the time of contract*. Id. at 65-67.

⁵¹ See cases cited in SEDGWICK, supra note – at 68-69 et seq., and in PARSONS, supra note – at 458-59.

⁵²In New York, the Court of Appeals reviewed an action brought for the price of a steamboat. The buyer showed that the machinery of the vessel was defective and that those and other defects caused considerable delay in putting the boat into operation in his established ferry business. The buyer sought to deduct from the contract price not only the cost of repairs but also the lost profits for the trips that the vessel might have run during the period of the delay, having proved that the net profits lost would be \$100. The Court of Appeals disallowed the recovery of lost profits. S at 71. In another New York case, a seller contracted to deliver a steam engine to a buyer who owned a mill for manufacturing oil. The seller failed to deliver the machine at the time for performance and the buyer sought the profits it could have made from manufacturing oil had the machine been delivered on time. The court disallowed the claim, limiting damages to the loss of full use of the mill and other machinery and interest on the additional stock purchased in anticipation of the installation of the engine. *Freeman v. Clute*, 3 Barb. S.C. R. 424. In a similar case from Massachusetts, in an action for breach of a contract to repair a mill-dam, the plaintiff was

Draft: February 22, 2004

Story, in a widely-cited opinion, held that “an allowance of damages upon the basis of a calculation of profits, is inadmissible. This would be a calculation upon conjecture and not upon facts.”⁵³

By the mid 1840's treatise writers had announced a general default rule governing consequential damages for breach of contract: The party in default is liable for all losses that may fairly be in contemplation of the parties at the time of contract; that is, the “plaintiff must have turned the mind of the defendant to the consequences likely to ensue from default.”⁵⁴ The early crystallization of this rule suggests that *Hadley v. Baxendale*, decided in 1854, announced a rule that had already developed in America. Perhaps, then, the primary effect of *Hadley* was not to limit recovery of consequential damages that previous courts would have allowed, but rather to extend damage recovery by granting the plaintiff the right to recover consequential damages in those instances where he had “communicated special circumstances” to the defendant.⁵⁵

In addition to formulating a market damages default, during this period the courts developed more clearly the principle that the “contract itself furnishes the measure of damages.”⁵⁶ The mid-19th century contract law thus distinguished and rejected earlier cases giving the jury wide latitude and discretionary authority to determine the measure of damages, either by reducing or enlarging the award. The amount of compensation was now regulated by the

entitled to recover only the expense incurred in repairing the dam, but not the loss of profits in its business.

⁵³The Schooner Lively, 1 Gallison 314, 325.

⁵⁴SEDGWICK, supra note – at 112.

⁵⁵Hadley v. Baxendale, 9 Exch. 341, 156 Eng. Rep. 145 (1854).

⁵⁶SEDGWICK, supra note – at 200.

direction of the courts, and the sole object was to ascertain the agreement of the parties, which agreement controlled the measure of damages.⁵⁷ To be sure, this did not mean that the plaintiff could always recover the contract price (or specifically enforce the contract). Rather, the plaintiff was allowed to recover the “price” for breaching (or terminating) the contract and that price was to be determined by the parties. Only in default of their determination was a court empowered to set the price of termination as the (narrowly defined) “loss” that was caused by the breach.

4. Distinguishing Between Liquidated Damages and Penalties

During the early decades of the 19th century, common law courts found themselves with a vexing interpretive problem. On the one hand, they had to police penal bonds. The restrictions on enforcing these bonds that had developed initially in courts of equity were subsequently enacted into statutes in many states.⁵⁸ As a result of these provisions, the general rule evolved that no other sum could be recovered on a bond than that which compensated the plaintiff for his actual loss.⁵⁹ On the other hand, as bonds decreased in use and executory contracts took their place, the courts had to interpret liquidated damages clauses which were increasingly common terms in executory contracts between commercial parties. Promisors who were sued for breach of an executory contract which

⁵⁷ Among other things, this principle was an explicit rejection of the concept of breach as “fault.” The motives behind the breach were “irrelevant.” PARSONS, *supra* note – at 443.

⁵⁸ These statutes had two principal provisions: 1) they gave defendants on bonds for the payment of money, the right to pay into court the principal sum of the debt with interest and costs even though the condition was in default; and 2) they required plaintiffs suing to enforce bonds conditioned on a particular performance to assign specific breaches by the defendant and to prove the amount of loss for each breach. SEDGWICK, *supra* note – at 393-96.

⁵⁹ Technically, since the bonds were sealed instruments and thus enforceable on their face, the plaintiff was awarded judgment for the face amount of the bond (since the “penalty” was the debt), but the statute only allowed execution to issue for the amount of the loss actually suffered and proved by the defendant.

Draft: February 22, 2004

contained a liquidated damage clause argued that these damages agreements were, in fact, disguised “penalties” and should be subject to the same legal restrictions that had come to narrow the scope of the so-called “penal bonds.”

The courts seized on the “intent of the parties” as the key to this interpretive dilemma. An invalid penalty in a performance bond was an amount that was not intended to be paid but rather was intended only as security for the performance of the executory promise in question. Enforceable liquidated damages, on the other hand, were intended to be paid by the promisor if it elected not to perform the agreement (a termination option, as it were).⁶⁰ Complicating the analysis further was the fact that, during this period, the courts also recognized that a contract to do or refrain from doing an act *or in the alternative* to pay a given sum of money was neither liquidated damages nor a penalty, but rather was an enforceable alternative contract (well-known to the civil law at this time).⁶¹ The understandable result of these diverse holdings was a universally decried confusion.

A much applauded “solution” to the confusion of the earlier case law was proposed in *Jaquith v. Hudson* in 1858.⁶² By the time *Jaquith* was decided, the widespread use of bonds had largely

⁶⁰ See cases cited in SEDGWICK, *supra* note – at 398-420.

⁶¹ Sedgwick, *supra* note -- at 398. Thus, by the mid-century the following generalizations could be advanced: First, the primary objective of courts in scrutinizing stipulated damage clauses was to determine the true intent of the parties, although the language they used was not conclusive evidence of that intent. Second, in England there was a presumption in favor of enforcing liquidated damage clauses in executory contracts, while American courts were more reluctant than their English counterparts to admit the agreement of the parties as conclusive. Third, when the agreement was in the alternative, courts would uniformly enforce the obligation to pay money. Fourth, where a court determined that the intention of the parties was to evade the statutory restrictions on penal bonds, the clause would be treated as a penalty and damages limited to actual losses. Fifth, independent of the above, if damages were uncertain and incapable of being measured except by conjecture, then any stipulated damages would be enforceable. *Id.* at 424

⁶² 5 Mich. 123 (1858).

disappeared. Thus, the only question courts faced routinely was the proper interpretation of clauses in executory contracts that purported to stipulate the payment of a sum certain in lieu of performance. Unhappily, the “intent of the parties” test that the courts had developed to distinguish between valid liquidated damages and invalid penalties simply did not square with the outcome of the cases. Many courts found an “implied intention” to create a penalty even though the parties insisted otherwise by their express language. *Jaquith* purported to reconcile these contradictions.

In *Jaquith*, the court upheld a liquidated damage clause enforcing a covenant not to compete in a contract for the sale of a partnership interest in a mercantile business. The court explicitly rejected the prevailing intent test of enforceability. Instead, the court held that the governing principle reconciling the cases was that damages for breach of any contract must be based on “the *principle of just compensation* for the loss or injury actually sustained; considering it no greater violation of this principle to confine the injured party to the recovery of less, than to enable him, by the aid of the court, to extort more.”⁶³ The compensation principle, the court declared, was a mandatory rule, a “principle of natural justice” and not a default rule, and thus the intention of the parties was irrelevant. Since the compensation principle was the law of the contract, parties were not permitted by express stipulation, however clear the intent, to set it aside.⁶⁴

But if the compensation principle is a mandatory rule, then what role do the parties have in

⁶³ Id. at —.

⁶⁴ The court cited no authority for the proposition that the compensation principle of expectation damages was an immutable rule. This is particularly surprising given that this “principle” had only been generally recognized for less than thirty years. The court reverted to maxims rather than to authority, holding that courts could set aside the parties intention on the familiar ground: “*conventus privatorum non potest publico juri derogare.*” Id. at —.

Draft: February 22, 2004

setting damages defaults? The court in *Jaquith* had a ready answer to this question. The only task for the parties was to specify just compensation ex ante in those instances where they had a comparative advantage over a court seeking to do so ex post. Such comparative advantage would exist where the anticipated loss from the breach of the contract was uncertain, remote or speculative. Thus, while *Jaquith* purported to restrict party sovereignty over stipulated damages, the *Jaquith* rule actually gave greater latitude to 19th century contracting parties than would be true under modern damage rules. Given the then prevailing view that lost profits could not be recovered without evidence that the parties had bargained with those potential consequences in mind, the range of ex ante uncertainty was substantial, especially in the case of specific-investment contracts in thin markets. Consequently, parties had considerable freedom to stipulate damages under the *Jaquith* rule.

5. From Jaquith to the Present: The Expansion of Expectation Damages

The 150 years following *Jaquith* reveals a consistent pattern in the evolution of remedial default rules for breach of contract. In *Globe v. Landa*, Justice Holmes sought to generalize the older common law rule of consequential damages as requiring that the parties must have made a “tacit agreement” that the breacher would be liable for losses caused by special circumstances.⁶⁵ damages. But this limitation on consequential damages came under substantial pressure over time. One source of pressure was the increasing sophistication of expert testimony in establishing with some specificity the quantum of lost profits that might result from breach. Meanwhile, as commercial enterprise expanded during the mid to late 19th century, the notion of lost profits for a business buyer became more commonly accepted, and courts came to regard lost profit claims as a natural and direct result of a

⁶⁵ *Globe Ref. Co. v. Landa Cotton Oil Co.*, 190 U.S. 540 (1903).

breach rather than as a special circumstance that required communication.⁶⁶ Eventually, the constraint on the recovery of lost profits was limited to the so-called “new-business rule” which denied recovery of profits in the case of a non-established business enterprise.⁶⁷

On its own terms, the trend toward granting lost profits as part of expectation damages was both understandable and defensible.⁶⁸ The unintended consequence of this development, however, was a corresponding reduction in the freedom of parties to stipulate damages. A further doctrinal shift occurred with the erosion of the tacit agreement test for the recovery of consequential damages. By the time the UCC was adopted in the 1950's, the tacit agreement test was explicitly rejected and a plaintiff was entitled to recover any loss resulting from the particular requirements and needs of which the seller had *reason to know* at the time of contracting.⁶⁹

The universal adoption of Article 2 of the UCC completed the expansion of the expectation damages default. In addition to relaxing the restrictions on recovery of consequential damages, the Code (unlike its predecessor the Uniform Sales Act) invites lost profit claims by volume sellers.⁷⁰ Even

⁶⁶See e.g., UCC §2-715. Comment 6.

⁶⁷The new-business rule is now treated by many courts merely as a presumption in favor of the defendant rather than as an absolute bar to the recovery of lost profits for start-up businesses. See *Drews Company, Inc. v. Ledwith-Wolfe Assoc., Inc.*, 371 S. E. 2d 532 (S.C. 1988).

⁶⁸Since the denial of lost profits was based on claims of uncertainty, the ability to establish losses with reasonable certainty justified recovery under an expectation default.

⁶⁹UCC §2-715(1) and Comment 2; Restatement (Second) of Contracts §351, Comment a (same). Compare §67 of the Sales Act of 1906 which retained the traditional rule.

⁷⁰The Code abandons the common law preference for market damages, which had been enshrined in the Uniform Sales Act, and instead explicitly authorizes the recovery of lost profits against a breaching buyer so as to “eliminate the unfair and economically wasteful results arising under the older law when fixed price articles were involved.” UCC §2-708(2) and Comment 2.

Draft: February 22, 2004

more significantly, the Code elevates the compensation principle to an overarching norm and a fundamental principle of interpretation in §1-106.⁷¹ Section 1-106 seemingly trumps the arguments for market damages even in thick market contexts. Thus, most courts under the Code have concluded that market damages should not be used where the award departs from the economic gain the promisee would have enjoyed had the contract been performed.⁷²

The process which has resulted in a broad expectation damages default and a corresponding reduction in parties' freedom to stipulate breach damages requires some explanation. What explains the current dominance of the compensation principle as a quasi-mandatory rule? One answer is that commercial parties have been largely successful in side stepping the compensation principle and its associated doctrines in important respects. While underliquidation is technically subject to the same constraints as overliquidation,⁷³ courts have generally enforced undercompensatory damage provisions and termination options. In response, commercial sellers have designed limitation of remedy provisions such as the ubiquitous "repair and replacement" clause. Moreover, both commercial and consumer buyers typically enjoy rights of return and other undercompensatory termination options. Finally, the alternative contract remains available as a means of designing embedded options that might otherwise be subject to the penalty rule.⁷⁴ Contracts with "take or pay" clauses or "play or pay" clauses are

⁷¹UCC §1-106 ("The remedies provided in this Act shall be liberally administered to the end that the aggrieved party may be put in as good a position as if the other party had fully performed...."). See Scott, *The Case for Market Damages*, supra note— at 1169.

⁷²See cases cited in note – infra.

⁷³See UCC §2-718, Comment 1.

⁷⁴Restatement of Contracts §334 (1932).

routinely enforced by courts.⁷⁵

On the other hand, it would be incorrect to suggest that the compensation principle and the penalty rule no longer have practical consequences (other than as a trap for the unwary who fail to come within one or more of the recognized “safe harbors”). Contemporary American courts continue to police liquidated damage clauses and a significant number of stipulated agreements between commercial parties are invalidated as penalties.⁷⁶ Contract terms that provide for non-refundable deposits or pre-payment of the contract price and franchise termination fees are particularly vulnerable to collateral attack in subsequent litigation. The linkage of a quasi-mandatory compensation principle with both the penalty rule and the (recently developed) expansive default of full expectation damages seems to have resulted from the anomalies of the penal bond and the early common law forms of action. But once entrenched, these rules have proved remarkably durable despite the claims of scholars that they preclude (or at least constrain) parties from writing efficient contracts.

C. Economic Contract Theory and Compensation

Given that compensation is now entrenched doctrinally as the motivating principle behind default damages rules and the regulation of liquidated damages, it is perhaps not surprising that scholars have advanced economic justifications for the rules in each case. However, the literature in the economic analysis of contracts reflects a growing academic consensus against the penalty rule and raises substantial doubts about the desirability of expectation damages. Indeed, the scholarship of the

⁷⁵ [cites]

⁷⁶A search for contracts cases of the past five years that invoke liquidated damages and the penalty rule returned 96 recorded decisions. A detailed examination of these cases revealed [describe results].

last decade scarcely mentions compensation either as an end or a means to maximizing the joint welfare of contracting parties.

The principal economic justification for expectation damages is that they compel the promisor to internalize the costs that her breach inflicts on the promisee. The promisor consequently has the incentive to make the efficient breach decision.⁷⁷ By internalizing the promisee's loss, the promisor also has the incentive to take the efficient precautions against contingencies that threaten to increase the cost of performance.⁷⁸ In addition, expectation damages provide insurance to the promisee against the risk of increases in the promisor's cost of performance that might lead her to breach. This insurance may be efficient if the promisor is risk neutral and the promisee is risk averse.⁷⁹

The literature has identified a large number of countervailing considerations against expectation damages. Our purpose is not to present a comprehensive review of these factors, but simply to

⁷⁷ See e.g., Goetz & Scott, *supra* note – at 558-59:

Generally, breach will occur where the breaching party anticipates that paying compensation and allocating his resources to alternative uses will make him “better off” than performing his obligation. As long as the compensation adequately mirrors the value of performance, this damage rule is efficient. It induces a result superior to performance, since one party receive the same benefits as performance while the other is able to do even better. Under the current damage rule, all of these net gains from breaching are retained by the breacher. In order to maintain the efficiency value of the rule, however, it is only necessary that some minimal amount of benefits are retained by the breacher in order to induce him not to perform.

⁷⁸ *Id* at 579-83.

⁷⁹ E.g., A. Mitchell Polinsky, *Risk Sharing Through Breach of Contract Remedies*, 12 J. Legal Stud. 427 (1983); Samuel A. Rea, Jr., *Efficiency Implications of Penalties and Liquidated Damages*, 13 J. Legal Stud. 147, 151-63 (1984). To many, the insurance objective seems inapt in contracts for goods or services. Several decades before Holmes wrote his contract options statement (*supra* note –), Pollock wrote: “A man who bespeaks a coat of his tailor will scarcely be persuaded that he is only betting with the tailor that such a coat will not be made and delivered to him within a certain time. What he wants and means to have is the coat, not an insurance against not having the coat.” F. Pollock, *Principles of Contract* xix (1881).

Draft: February 22, 2004

illustrate that they present a very significant challenge to the case for expectation damages and indeed compensatory damages in general. Expectation damages ensure efficient breach only if the court can accurately and predictably measure the promisee's losses at no cost. Yet, the practical difficulty in verifying the promisee's losses makes the court's assessment of expectation damages uncertain and costly, and this may lead either to excessive or insufficient breach. For example, even if the court is accurate on average, a risk averse promisor may engage in inefficient performance (i.e. where the cost exceeds the promisee's anticipated benefit) out of a desire to avoid the uncertainty in her prospective damages liability.⁸⁰ Moreover, the expected enforcement costs may outweigh the incentive and insurance gains from expectation damages, so that the parties may be better off with undercompensatory damages that reduce the likelihood of litigation. At least in part, the common law doctrines of foreseeability and uncertainty may be explained as striking such a compromise.

More prominent in economic contract theory is the observation that expectation damages give rise to well known moral hazard incentives for the promisee to overrely on promises and to fail to take steps to avoid losses from breach.⁸¹ Several common law doctrines cut back on compensation in order to address this problem as well. In particular, the mitigation doctrine deprives the promisee of

⁸⁰If extralegal sanctions supplement expectation damages, even accurately calculated damages will over deter breach. [cite].

⁸¹This is the well-known overinvestment problem that is created by reliance or expectation damages (as well as specific performance). See Steven Shavell, *Damage Measures for Breach of Contract*, 11 Bell J. Econ. 466 (1980); William P. Rogerson, *Efficient Reliance and Damage Measures for Breach of Contract*, 15 Rand J. Econ. 39 (1984). Assume, for example that the seller contracts to produce a specialized good for the buyer and the buyer makes a specific investment (one that has no value other than with the good) in order to increase its valuation of the good. The buyer's expectation damages if the seller breaches will be the difference between its realized valuation and the price. The buyer will always receive the return on his investment, even if trade does not occur and the investment is effectively wasted. Thus, the buyer's marginal return for investment exceeds the social return, leading to inefficiently high investment.

Draft: February 22, 2004

compensation for post-breach losses that she reasonably could have avoided.⁸² The foreseeability and uncertainty doctrines remove unusual losses from the measure of damages and thereby dampen the incentive to rely more than the average promisee.⁸³ Moreover, the fact that plaintiffs tend not to be compensated for waiting for payment until after trial for payment and that they do not recover their legal costs, restores some of the promisees' incentives to protect themselves from breach. Nevertheless, the problem of excessive reliance and insufficient precaution by the promisee remains whenever the damages are compensatory in the sense that they vary at the margin with the losses of the promisee.

Economists in this field have embraced the dual objective of optimizing both the ex post decision whether to trade and the ex ante incentives to make specific investments. The challenge is posed by the fact that the conditions for efficient trade and efficient investment cannot be fully specified in the contract in light of transaction costs: the cost of foreseeing and describing all possible future states of the world, and the cost of verifying the realized state to a court. Even if expectation damages were sufficient to deliver ex post efficiency, the parties could not condition the damages award on ex ante

⁸² The doctrine of avoidable consequences has both negative and affirmative aspects. The affirmative branch of the doctrine permits recovery of all reasonable expenses that the plaintiff incurs in seeking to avoid damages. See e.g., *Rench v. Hayes Equip. Mfg. Co.*, 134 Kan. 865, 8 P.2d 346 (1932). The negative branch of the doctrine precludes a plaintiff from passively incurring losses which she could reasonably avoid or from actively increasing such losses where prudence would require an adjustment. See, e.g., Restatement (Second) of Contracts §350; *Rockingham County v. Lutten Bridge Co.*, 35 F.2d 301 (1929). For discussion, see Charles J. Goetz & Robert E. Scott, *The Mitigation Principle: Toward a General Theory of Contractual Obligation*, 69 Va. L. Rev. 967, 973-76 (1983).

⁸³ In general, speculative losses cannot be recouped. These include lost good will, reputation or emotional distress. See SCOTT & KRAUS, *CONTRACT LAW & THEORY* at 1050 et seq. The foreseeability limitation has been considerably eroded over time. Under UCC §2-718, for example, the plaintiff can recover all losses resulting from the general or particular requirements and needs of which the seller at the time of contracting had reason to know....” See also, Restatement (Second) of Contracts §351.

Draft: February 22, 2004

efficient investment.⁸⁴ This applies not only to the aforementioned tendency of the promisee to overrely, but also to the fact that when the promisee is insured against breach, she lacks the incentive to incur costs that help or induce the promisor to avoid breach.⁸⁵

Yet, even when information is not verifiable, ex post efficiency is easy to achieve through implementation mechanisms, as long as the parties themselves are symmetrically informed.⁸⁶ These mechanisms elicit from the parties their information about the state of the world that can then be used to identify to the court efficient payoffs for that state. The availability of these solutions diminishes the significance of compensatory damages in forcing promisors to internalize the losses caused by their breach. Ensuring efficient investment incentives, especially on both sides of the contract, is a more difficult task.

The contribution of expectation damages in promoting efficient decisions to trade is even less valuable if contracting parties are able to renegotiate their agreement. The premise that parties can often renegotiate to efficient ex post outcomes is sufficiently well accepted that economists have largely

⁸⁴Aaron Edlin and Stefan Reichelstein resuscitate expectation damages in their model to show that they can lead to efficient investment where only the promisee's investment is relevant and where the parties can renegotiate their agreement. Aaron S. Edlin and Stefan Reichelstein, *Holdups, Standard Breach Remedies, and Optimal Investment*, 86 Am. Econ. Rev. 478 (1996)..

⁸⁵Yeon-Koo Che & Donald Hausch, *Cooperative Investments and the Value of Contracting*, 89 Am. Econ. Rev. 125 (1999).

⁸⁶Jean Tirole, *Incomplete Contracts: Where Do We Stand?*, 67 *Econometrica* 741 (1999); Benjamin Hermalin and Michael Katz, *Judicial Modification of Contracts Between Sophisticated Parties: A More Complete View of Incomplete Contracts and their Breach*, 9 J. Law, Econ. & Org. 98 (1993). Under some conditions, an option contract might reveal even asymmetrically held private information. Georg Noldeke and Klaus M. Schmidt, *Option contracts and renegotiation: a solution to the hold-up problem*, 26 *Rand J. Econ.* 163 (1995)

Draft: February 22, 2004

set aside the concern with efficient breach to focus on the problems caused by renegotiation.⁸⁷ In particular, the division of the renegotiation surplus does not reflect the fact that one or both the parties have made specific investments contributing to the surplus, and therefore the investing party is compelled to share the payoffs from her investment. Thus, renegotiation of contractual commitments undermines attempts to establish efficient investment incentives.⁸⁸ Given that states of the world and performance obligations cannot be fully described in the initial contract, the courts often cannot discern when a party is acting strategically in forcing renegotiation to capture investment payoffs by refusing to carry out an efficient trade. One set of solutions to this problem proposes that the parties raise ex ante legal or practical obstacles to renegotiation, such as bargaining through organizations whose internal regulations require many layers of consent.⁸⁹ A more popular theoretical approach is to try to predetermine the allocation of the renegotiation surplus by effectively assigning bargaining power to one party or the other.⁹⁰

⁸⁷Oliver Hart and John Moore, *Incomplete Contracts and Renegotiation*, 56 *Econometrica* 755 (1988).

⁸⁸Oliver E. Williamson, *Transaction Cost Economics: The Governance of Contractual Relations*, 22 *J. Law & Econ.* 233 (1979); Benjamin Klein, Robert G. Crawford and Armen A. Alchian, *Vertical Integration, Appropriate Rents, and the Competitive Contracting Process*, 21 *J. Law & Econ.* 297 (1978). The key to the hold-up threat is the presence of specific investment that, by definition, increases the value of contract performance but is worthless if the deal falls through. For discussion, see Schwartz & Scott, *Contract Theory*, supra note – at —.

⁸⁹See e.g., Alan Schwartz & Joel Watson, *Economic and Legal Aspects of Costly Contracting*, 20 *J. L. Econ. & Org.* #2 (2004); Tirole (1999). This precommitment generally cannot be made simply by contract. Under current law, an agreement not to modify a contract is not enforceable. Restatement (Second) of Contracts §311 comment. A (1979); *Beatty v. Guggenheim Exploration Co.*, 122 N.E. 2d 378, 387 (N.Y. 1919) (“Those who make a contract can unmake it. The clause that forbids a change can be changed like any other ... Whenever two men contract, no limitation self-imposed can destroy their power to contract again.” Cardozo, J); See also, *Zumwinkel v. Legget*, 345 S.W. 2d 89 (Mo. 1961); Christine Jolls, *Contracts as Bilateral Commitments: A New Perspective on Contract Modification*, 26 *J. Leg. Stud.* 203 (1997); Robert E. Scott, *Rethinking the Default Rule Project*, 6 *Va. J.* 84 (2004).

⁹⁰E.g. Tai-Yeong Chung, *Incomplete Contracts, Specific Investments, and Risk Sharing*, 58 *Rev. Econ. Stud.* 1031 (1991); Philippe Aghion, Mathias Dewatripont & Patrick Rey, *Renegotiation Design with Unverifiable Information*, 62 *Econometrica* 257 (1994); Georg Noldeke and Klaus M. Schmidt, *Option contracts and renegotiation: a solution to the hold-up problem*, 26 *Rand J. Econ.* 163 (1995). The idea in these two articles is to

Draft: February 22, 2004

In sum, contemporary economic contract analysis indicates that expectation damages are neither sufficient nor necessary to achieve efficient breach. Parties have the choice between more complete contracts that yield both efficient trade and efficient investment but can only succeed if there is no prospect of renegotiation, or simple contracts which the parties may renegotiate to achieve efficient trade without compromising investment efficiency.⁹¹ More to the point, this theoretical literature raises doubts not only about the doctrinal constraints on liquidated damages, but also about the wisdom of compensatory damages defaults.⁹²

D. Contract Damages and Termination Provisions

The irrelevance of compensation in the economic theory of contracts is reflected in the patterns of commercial and consumer contracts in practice. Many, if not most, contracts provide that at least one party may walk away from the exchange by paying an amount that is different from the expectation

set the correct investment incentives for one party through the setting of her default point in renegotiation and for the other party by assigning her bargaining power over the division of the renegotiation surplus. Property ownership rights may also be used to determine the parties' bargaining position in ex post renegotiation. Sanford J. Grossman and Oliver D. Hart, *The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration*, 94 J. Pol. Econ. 691 (1986); Oliver Hart and John Moore, *Property "Rights and the Nature of the Firm*, 98 J. Pol. Econ. 1119 (1990). Yet another approach is to permit bargaining power in renegotiation to be exogenously determined, but to set the contract quantity so as to balance the likelihood of renegotiation to a lower or higher value, in such a way as to also cancel out under- and overinvestment tendencies. Aaron S. Edlin and Stefan Reichelstein, *Holdups, Standard Breach Remedies, and Optimal Investment*, 86 Am. Econ. Rev. 478 (1996).

⁹¹See, e.g., Tirole, *supra* note –; Schwartz and Watson, *supra* note –.

⁹²An excellent review of the arguments for and against the penalty rule is Aaron S. Edlin and Alan Schwartz, *Optimal Penalties in Contracts*, 78 Chi.-Kent L. Rev. 33 (2003). [acknowledge economic arguments against penalties; e.g. Tai-Yeong Chung *On the Social Optimality of Liquidated Damage Clauses; An Economic Analysis*, 8 J. L. Econ. & Org. 280 (1992); Philippe Aghion & Patrick Bolton, *Contracts as a Barrier to Entry*, 77 Am. Econ. Rev. 388 (1987)].

damages measure. In some cases, the amount is liquidated damages fixed by contract.⁹³ In truth, the widespread deviation from the compensation principle is evidenced in a host of termination provisions, of which breach damages are but a subset. Although the freedom to contract over liquidated damages is constrained by the compensation principle (particularly the penalty rule), contracting parties are at liberty to structure substantive obligations so as to define breach in any way they deem appropriate. To a significant degree these substantive and remedial provisions are substitutes.⁹⁴ For example, a termination or cancellation clause with a fee permits the promisor to avoid performance with the payment of the fee instead of damages.⁹⁵ The right to return or to walk away from a contract with the loss of a deposit has the same effect as the option to breach and pay damages. In some industries, buyers frequently have similar choices between purchasing and paying a fixed amount under take-or-pay or play-or-pay clauses.⁹⁶ A requirements or output contract permits one side to avoid the

⁹³See TAN *supra*. Examples abound in connection with seller termination (including breach) as well. Richard Epstein observes that the warranty provisions in many sales contracts limit the buyer's remedies for breach either to liquidated damages that are undercompensatory or to the repair or replacement of the damaged goods at the option of the seller. Richard A. Epstein, *Beyond Foreseeability: Consequential Damages in the Law of Contract*, 18 J. Legal Stud. 105, 114-21 (1989). He also describes the standard limitation on damages in delivery services such as Federal Express, which undercompensate even the average consumer. Consistent with the analysis summarized above, Epstein concludes that, "[a]gainst the backdrop of express contractual provisions, there is ample reason to doubt that the expectation measure of damage of the classical common law maximizes the joint gains of the parties *ex ante*. If it did, we should expect to observe it frequently in practice, which is decidedly not the case."

⁹⁴Restatement §368 states that specific performance or an injunction will not be granted against a party who can substantially nullify the effect of the order by exercising a power of termination or avoidance. In contrast, §361 provides that the court can order an injunctive remedy in the face of a liquidated damages term.

⁹⁵Many such options are categorized doctrinally as alternative contracts. Traditional analysis has distinguished the alternative provision designed to secure performance of the primary promise (a liquidated damages clause) from two promised alternatives between which the promisor can choose, each an agreed exchange for the consideration given by the promisee (an embedded option in our terminology). *Garrett v. Coast & S. Fed. Sav. & Loan Ass'n*, 9 Cal. 3d 731, 738, 511 P.2d 1197, 1201 (1973); Restatement of Contracts §339 Comment f (1932). The former are at risk of being found unenforceable as a penalty, while the latter would be an enforceable alternative contract. The Restatement of Contracts, in classic understatement, recognizes that enforceable alternative contracts may easily be confused with invalid liquidated damage provisions. §334 Comment c (1932).

⁹⁶[cite Goldberg]

Draft: February 22, 2004

obligation to purchase or deliver larger quantities of goods.⁹⁷ These terms, and many others,⁹⁸ share the common feature that one party has the option to terminate the contemplated exchange by paying some amount or incurring some cost. We label the group of alternative mechanisms, “termination provisions,” recognizing that breach damages are but one subset.

This observation underscores the importance of improving our understanding of how contracting parties make their choices among alternative termination provisions and, specifically, how they select a termination fee. In Part II, we characterize termination provisions as options embedded in contracts. In this framework, the termination fee is effectively the price paid for the option enjoyed by the promisor. We focus on the termination right held by a buyer of a good. In the case of termination by breach, this option price is the amount of damages owed by the breaching buyer. We then examine the factors that determine the option price to which the parties agree. Although the protection of specific investment remains an important contract objective, our analysis shifts the focus to risk management. We conclude that contracts set option prices equal to the ex post loss from termination only by coincidence, thereby raising serious doubts about the compensation principle in the measure of

⁹⁷UCC § 2-306(1). Discretion under an output or requirements contracts is effectively constrained only by the duty of good faith, which limits moral hazard risk. A number of courts have recognized that under demand in the case of a requirements buyer (or under supply for an output seller) need only be justified by a “valid business reason” to pass the test of good faith imposed by UCC §3-601(1). See e.g., *Empire Gas Corp. v. American Bakeries Co.*, 840 F.2d 1333 (1988). Thus construed, a requirements contract can be characterized as an unconstrained option to take as small a quantity of the good as the buyer chooses. The buyer’s discretion is regulated in the over demand situation to prevent the buyer from opportunistically exploiting the contract price and reselling the goods in a favorable market. See *Empire Gas Corp. v. American Bakeries Co.*, 840 F.2d 1333 (7th Cir. 1988). Many other contracts grant one of the parties discretion to determine a significant exchange term. For example, the seller under an installment contract has discretion over the content of any installment. Breach does not occur unless nonperformance substantially impairs the value of the whole contract. UCC § 2-612. The UCC default rules grant the buyer an option relating to assortment of goods and the seller an option relating to shipment. UCC § 2-311.

⁹⁸For example, a unilateral promise invites the promisee to accept by rendering a performance rather than by a promissory acceptance. The promise creates an option contract once the promisee begins the invited performance. Restatement (Second) of Contracts § 45.

contract damages.

II. EMBEDDED OPTIONS

A. *Termination Rights and Call Options*

For expositional purposes, we focus on the termination right of a buyer of a good and assume that the seller's obligation is specifically enforceable. We define the buyer's termination right by the pair (d,x) , where d is the termination fee and x is the difference between the contract price, P , and the termination fee. For example, if the buyer's promise is enforced by damages, then d is the damages liability and x is the difference between the contract price and those damages. Perhaps the most intuitive way to think about termination is to view the buyer as having purchased the good for the price P and holding a put on the good with exercise price, x . In this paper, however, we use the alternative characterization in which the buyer pays an *option price*, d , to acquire a call option to purchase the good with an *exercise price* of x . Under the put-call parity rule, the call option is equivalent to the combination of a long position on the contract and a put.⁹⁹ We prefer the call option characterization because it is more parsimonious and analytically revealing than the contract-plus-put combination. The put characterization implicitly assumes that the buyer will enter into the contract and asks whether the parties would agree to a put (and on what terms). In contrast, framing the buyer's position as a call removes the decision one stage earlier in time. It acknowledges that the buyer has the alternative of deferring contracting and emphasizes that the contract offers an opportunity as well as insurance.

⁹⁹Specifically, a call option is equivalent to the combination of holding the underlying asset, borrowing the exercise price and a put option on the same asset. E.g., Hans R. Stoll, *The Relationship between Put and Call Option Prices*, 24 J. Fin. 801 (1969). We assume a zero discount rate.

The ways in which the price of the call option is paid are as diverse as the forms of termination fees. The buyer may make a nonrefundable deposit, agree to a cancellation fee, or assume a commitment to pay damages in the event of breach. The common feature is that the buyer pays for the call option by incurring an initial cost specified by contract and can subsequently choose to incur an additional cost to execute the contract exchange.¹⁰⁰ The sum of the option price and the exercise price is the contract price, $P = d + x$.¹⁰¹

The parties have the choice among a set of alternative option contracts, (d, x) .¹⁰² Consider first a contract in which the buyer agrees to buy a widget for \$12 with no termination right: (12,0). Now, suppose that the seller offers the buyer an alternative contract under which the buyer makes a deposit, d , and holds an option to purchase the widget for \$1. The buyer pays d to acquire a call option with exercise price \$1. The deposit, d , is effectively the price the buyer pays for the option. If there is some possibility that the buyer will value the widget for less than \$1, she will be prepared to pay a deposit --

¹⁰⁰The seller's option to terminate can be similarly presented. If the seller has a termination right, then she effectively incurs a liability for the termination fee and can pay the additional cost of delivering the good to the buyer in order to receive the contract price. The case of bilateral buyer's and seller's options is substantially more complex. See Triantis and Triantis, *supra* note –, at 184-94.

¹⁰¹The payment of the termination fee or damages by the buyer or seller, respectively, is subject to a further option if there is a risk of insolvency. Each party can obtain a discharge from the obligation by surrendering their assets in a bankruptcy proceeding. Thus, the option enjoyed by a contracting party is essentially a compound option because it includes both the termination option in the contract and the bankruptcy option held by all debtors. Triantis, *Effects of Insolvency and Bankruptcy*, *supra* note –, at 682-7. For the purposes of this essay, however, we assume that both parties cannot avoid their debts by bankruptcy or dissolution.

¹⁰²Some of the economic models of contracting analyze a contract that provides for two prices to be paid by the buyer. The buyer pays p_1 if there is trade and p_0 if there is not, where p_0 might be negative. The critical feature is the difference, $p_1 - p_0$, and the parties are otherwise free to set p_0 where they wish. See, e.g., Nöldeke and Schmidt, *supra* note –; Hart and Moore, *supra* note –. “[O]nly the difference $p_1 - p_0$ matters for efficiency. Therefore, we have one degree of freedom in choosing p_0 which can be used to share the expected surplus *ex ante* between the parties.” Klaus M. Schmidt, *contract renegotiation and option contracts*, 1 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND LAW 433 (Peter Newman, ed. 1998).

Draft: February 22, 2004

effectively an option price -- greater than \$11.¹⁰³ Writing this option may also be costly for the seller, but the parties will prefer to trade the option if this cost is less than the value of the option to the buyer. If the seller has some bargaining power, she will be able to capture some of the surplus. It is therefore likely that the aggregate contract price greater than \$12 when the seller writes a call option with exercise price of \$1. As a general matter, the sensitivity of the option price to changes in the exercise price is such that, for each dollar increase in the exercise price, the price of the option decreases by less than a dollar.¹⁰⁴ Therefore, the aggregate contract price rises asymptotically with the exercise price as the parties approach $(0, P^{\wedge})$.¹⁰⁵

If we begin our analysis instead at the $(0, P^{\wedge})$ end of the spectrum, suppose that $P^{\wedge} = x = \$18$. The seller effectively has given the buyer a free call option with an exercise price of \$18. We later discuss the possible motivations for giving a free option. For now, note that if the seller offers to reduce the exercise price from 18 to 17, the buyer will be willing to pay an option price in the form of a deposit -- but in an amount less than \$1. Similarly, if the seller offers to reduce the exercise price further, the buyer will be prepared to increase her deposit, but by less than the reduction in the exercise price. Of course, the seller does not seek to maximize the contract price, but rather to maximize the difference between the option price and the cost of writing the option for the buyer. In the discussion that follows, we identify two sources of value of the call options: (a) preserving the abandonment option given by nature and (b) insuring the buyer against the risk of fluctuations in her valuation of the seller's performance. At the end of this Part, we invoke some basic

¹⁰³Of course, if the time value of money is not zero, the buyer derives an additional benefit in deferring payment of even a small portion of the contract price.

¹⁰⁴JOHN HULL, INTRODUCTION TO FUTURES AND OPTIONS MARKETS.

¹⁰⁵In nominal (dollar) terms, the option price is most sensitive to changes in exercise price when the option is near or at the money (exercise price equal to the buyer's valuation).

insurance analysis to identify factors that determine the optimum (d,x) pairing. The factors are unlikely to yield an optimal value for d that is equivalent to the expectation measure of damages. Moreover, they are sufficiently heterogeneous and context dependent to support our argument in Part III against default damages provisions.

B. The Abandonment Option

A primary source of the value of a termination right in a thin-market contract is the real option of abandonment.¹⁰⁶ A given venture may seem profitable at the outset, but as new information is acquired over time, there may be good reason to abandon it in mid-stream in order to limit losses. When the venture is pursued by a single decision-maker, this valuable flexibility is referred to as an abandonment option. When the venture is carried out by contract, the agreement implicitly splits the value of the abandonment option between the parties and assigns the abandonment decision to one or both of the parties.

To illustrate this point, consider a firm whose research and development has yielded the design of a novel product for sale to consumers. To exploit this opportunity, the firm must first design and build a machine to manufacture the product. Suppose that the firm makes some preliminary investments before facing the final decision to spend C to construct the machine. The cost, C, is uncertain because it is a function of changing technology and input prices. The value of the machine, V, is also uncertain because it varies with the likely demand for its product. If the value of V falls below C, the firm may walk away from its sunk investment. This flexibility to abandon is valuable to the extent that the distributions of V and C are not strongly positively

¹⁰⁶The following analysis of the abandonment option in contractual relations follows Alexander Triantis and George Triantis, *Timing Breach in Contract Breach Decisions*, 41 J. Law & Econ. 163, 184-7 (1998). Other valuable real options are the option to delay, accelerate or decelerate investment in a project, or to switch to a different project. These other options may also be protected by contract: for example, we discuss the option offered to Airbus customers that permits them to switch the aircraft model they purchase. See *infra*.

Draft: February 22, 2004

correlated; it is especially valuable when they are negatively correlated. The option to abandon may also lead the firm to begin construction earlier than if the abandonment option were not available; that is, even if $E(V) < E(C)$. It is important to appreciate that this flexibility is one given by nature, in the sense that its exercise produces value without imposing a cost on any other person.

Suppose instead that the firm decides to outsource the construction of the machine to another firm (the “seller”). In light of the significant specific investments that can be made by each firm, the parties may enter into an executory contract. If the contract consists of rigid commitments by each side to perform -- the seller to deliver the specified machine and the buyer to pay the contract price -- the abandonment option of the integrated firm would be lost. This lost flexibility would therefore be a cost of contracting. Termination rights can reestablish flexibility by giving one or both parties the right to trigger the abandonment of the contract. They are most valuable when the abandonment option is significant; as noted above, when the distributions of V and C are loosely or negatively correlated.

Conventional efficient breach theory holds that expectation damages promote efficient abandonment decisions by compelling the buyer to internalize the cost of her decision to the seller. The buyer will then trigger abandonment only when $V > C$. As discussed earlier, however, contract theory suggests that expectation damages are neither necessary nor sufficient to ensure optimal abandonment. Moreover, a promisee protected by expectation damages has the incentive to make excessive specific investments. The seller in the foregoing example might accelerate the manufacture of the machine to effectively deprive the buyer of her termination option.

The put option characterization of the buyer’s termination right is more consistent with the

Draft: February 22, 2004

abandonment option of the integrated firm. However, the analysis can be reframed in terms of call options. Consider first the integrated firm. The preliminary investments in the construction of the machine give the firm a call option to spend C and reap a payoff of V . When this option is split between a discrete buyer and seller, the buyer incurs the expected termination cost under the contract in order to hold an option to pay the balance of the contract price and acquire V . If the option price is equivalent to expectation damages, the buyer's option simply represents its portion of the call of the integrated firm. However, if the option price is invariant to expectation damages, the call option plays an insurance role described in the remainder of Part II.

C. Two Insurance (Risk Management) Functions

When insurance concepts are invoked in the analysis of contract damages, they concern the risk of the promisor's breach. Thus, a risk averse promisee (say, the buyer) would pay a premium to be insured against increases in the seller's costs that lead the seller not to perform; a risk neutral promisor (say, the seller) is willing to accept that risk in return for the premium. The buyer, however, bears another risk that has been neglected in the economic analysis thus far: changes in the *value* of the seller's performance. A risk averse buyer, therefore, is willing to pay a further premium for the ability to avoid the contract if her valuation should fall below the contract price.

This distinction can be seen in the example of the previous section that compared contracting to an integrated firm that exploits the same opportunity. Recall that the cost of the project (setting aside ex ante investment) is C and its value is V , yielding a net payoff of $V-C$. The firm bears the risk of exogenous shocks to the joint distribution of C and V that cause $V-C$ to rise or fall. Suppose the project is exploited by a contract under which the seller incurs the cost, C , and the buyer realizes the value, V . If the contract is specifically enforceable and its price is a function of C , the buyer bears the risk of changes in C and V . Under

Draft: February 22, 2004

a specifically enforceable contract with a fixed price, P , the seller bears the cost risk ($P-C$) and the buyer keeps the value risk ($V-P$). If the fixed price contract is enforced by expectation damages, each party can shed some of their respective downside risk (in particular, the risk that V falls at the same time as C rises) without passing any additional risk to the other party. This is the abandonment option described in the previous section. Neither the seller or buyer suffers from conditions that increase C and decrease V so that $C > V$. Finally, if the buyer can walk away from the contract by paying less than expectation, she shifts to the seller some of the risk of decreases in V . Where the seller sells the option to the buyer, the seller bears the risk that V will fall below the exercise price. If the buyer is risk averse and the seller risk neutral, the parties might agree to shift the downside risk of fluctuations in V to the seller.

Some scholars believe that risk aversion can explain very little of the contracting that occurs in practice.¹⁰⁷ Although risk aversion may be accurate in the description of consumer parties, it seems inappropriate where business entities are concerned, particularly in light of capital market asset pricing theory. The theory predicts that managers of business firms cannot increase firm value by buying insurance because investors are at least as capable of hedging or diversifying nonsystematic firm risks in their financial portfolios. Moreover, firms also cannot create value by hedging systematic risks because shareholders can choose their desired portfolio risk by altering their asset allocation between risky and risk-free assets.¹⁰⁸ Corporate managers, however, actively hedge interest rate, exchange or commodity price risks, and risk management is the topic of much literature both by serious academics and practitioners.

¹⁰⁷[E.g. Goldberg]

¹⁰⁸See, e.g., Rene Stulz, Mastering Risk I: Diminishing the threats to shareholder wealth, *Financial Times* (April 25, 2000).

Draft: February 22, 2004

Reducing risk at the firm level has several advantages that cannot be replicated by investors and that might therefore lower the firm's cost of capital.¹⁰⁹ First, if the firm's tax rates are progressive, its tax liability is lower if income is stable rather than volatile. Second, the risk aversion of managers raises the cost of realigning their incentives by increasing the portion of compensation they derive from stock or stock options (particularly restricted stock). By dampening the variance in firm profitability, risk management reduces the cost of such compensation and allows firms to load up on incentive compensation.¹¹⁰ Third, risk reduction encourages shareholders to concentrate their investments in the firm, improving their incentive to monitor management.

Fourth, volatility in cash flow raises the probability of financial distress that would impose restructuring costs and disrupt relations with suppliers and customers. The higher risk of financial distress increases the cost of debt capital and lowers leverage. Debt financing, however, has several advantages: notably, interest payments are tax favored and debt may reduce managerial agency costs by compelling the disgorgement of free cash flow.¹¹¹ Therefore, by increasing debt capacity, risk management restores the firm's ability to realize these gains. Fifth, volatility also increases the conflict between stockholders and debtholders over capital budgeting and investment decisions. If the probability of insolvency is high, shareholders are more likely to induce their managers to invest in unprofitable risky ventures.¹¹² And, they are more hesitant to invest in lower risk, profitable projects because the payoffs from these projects will accrue to debtholders in the insolvency

¹⁰⁹For a summary of the empirical articles supporting this proposition, see Charles W. Smithson and Clifford W. Smith, Jr., *Managing Financial Risk* 505-9 (1995).

¹¹⁰E.g., Rene Stulz, *Optimal Hedging Policies*, 19 *J. Fin. Qu. Anal.* 127 (1984).

¹¹¹[cite Michael Jensen]

¹¹²[E.g. Jensen and Meckling]

Draft: February 22, 2004

states of the world.¹¹³ Sixth, volatility in cash flows affects the availability of internal financing of new projects. If information asymmetries impede the raising of external capital, fluctuations in the available internal funds may prevent efficient capital budgeting.¹¹⁴

Of course, the goal of risk management is not the reduction of risk. After all, the cheapest way to minimize risk might be simply to hold Treasury Bills. A firm must examine the distribution of payoffs from its operations and determine its overall risk profile. Risk management is a dynamic, portfolio strategy that encompasses myriad interrelated decisions about capital investments, technology, financing, marketing and supply relationships.¹¹⁵ Each strategy presents both benefits and costs. For example, diversifying operations reduces risk, but there is considerable evidence that it distorts investment decisions by permitting inefficient cross-subsidization of projects.¹¹⁶ A firm may also adapt to its risks by holding a buffer of cash or a line of credit, but this also may lead to overinvestment in unprofitable projects. Flexible manufacturing techniques may lower the vulnerability to market risks. For example, Hewlett Packard recently implemented a flexible assembly process that allows printers to be customized at the location of the customer rather than at its principal manufacturing plants.¹¹⁷ A firm may look to its capital structure and the design of its financial instruments to reduce the danger of insolvency. It may, for example, shorten the maturity of its outstanding

¹¹³[cite Myers]

¹¹⁴Kenneth Froot, David Scharfstein and Jeremy Stein, *A Framework for Risk Management*, Harv. Bus. Rev. 91-102 (Nov-Dec 1994).

¹¹⁵E.g., Culp, *Contingent Capital: Integrating Corporate Financing and Risk Management Decisions*, 15 J. App. Corp. Fin. 46 (2002).

¹¹⁶[cites]

¹¹⁷Alexander J. Triantis, *Real Options and Corporate Risk Management*, 13(2) J. App. Corp. Fin. 64-73 (Summer 2000).

debt or condition its interest obligations on the market prices of its critical inputs or products.¹¹⁸ Insurance companies have stepped up their competition against financial instruments by providing tailored insurance against bundles of risks, including earnings insurance.¹¹⁹

Forward contracts and options are perhaps the most widely discussed tools of risk management. Like the other mechanisms, they come with a downside too: for example, shareholders or managers may use them to speculate rather than reduce risk.¹²⁰ It is common to think of risk management narrowly in the sense of transactions in traded derivatives (particularly in foreign exchange, interest rates or commodities), where counterparties have a comparative advantage in hedging or diversifying the relevant risks. Yet, a firm's contracts with its customers and suppliers are relatively underappreciated instruments of risk management.¹²¹ As discussed in the following section, embedded options play a key role in the risk allocation of these thin-market contracts.

D. Forward Contracts, Option Contracts and Embedded Options

The important distinction we wish to emphasize in the context of risk management tools is the one between thick-market and thin-market contracts. A thick-market contract (forward or option) is motivated purely by speculation or risk allocation. Two important features of this contract are (1) the risk is exogenous so that the parties behavior has no impact on the payoffs from the contract, and (2) ex post, the value of the

¹¹⁸See Corey Billington, Blake Johnson and Alex Triantis, A Real Options Perspective on Supply Chain Management in High Technology, 15(2) J. App. Corp. Fin. 32 (2003).

¹¹⁹Neil A. Doherty, *Integrated Risk Management* ch. 14 (2000).

¹²⁰[E.g. Metallgesellschaft,...]

¹²¹Recent developments in FASB standards have reflected the realization that such contracts, particularly if long-term, have a significant impact on a firm's risk profile, and should be disclosed in financial statements. [cite]

Draft: February 22, 2004

commodity to the buyer (V) is equal to the cost of performance to the seller (C), because they are each equal to the market price. A forward contract is a simple purchase and sale contract in which the parties exchange risks, either because they have different expectations about future spot market prices or they have different abilities to manage the market risk. Forward contracts in thick markets are not option contracts, nor do they embed options the way we shall describe that thin-market contracts do. Neither party has any meaningful termination or cancellation rights because of the symmetry between the value of a thick market forward contract to the buyer and the performance cost of the seller. There is no abandonment option for the parties to share. The substance of pure thick-market contracts renders inapt any discussion of compensatory damages: The enforcement of such forward contracts is simply the specific enforcement of the market risk allocation.

Option contracts in thick markets have various motivations but again, they all relate to risk and insurance. For example, an optimistic buyer may purchase a call option from the seller in order to increase the leverage with which she can speculate and to obtain some downside protection. Indeed, the parties may enter into an option contract even if they share the same view of the expected future market price. If the buyer's subjective distribution has greater variance, she will attribute a higher value to the option than what the seller perceives to be the cost of writing the option. More to the point of our analysis, even if the parties observe the same distribution, the seller may write an option in favor of the buyer if risk is costly to the buyer and the seller has a comparative advantage in risk management, such as the opportunity to diversify or to hedge the market risk.¹²²

¹²²For example, the seller of a call option assumes the risk of the underlying asset value when it is out of the money; the buyer bears the risk while it is in the money.

Draft: February 22, 2004

The commodities seller may offer a wide range of alternative option contracts to the buyer. Option contracts are defined by the underlying asset (e.g., the commodity), the exercise price and the maturity date. For any given maturity and commodity (specifically, the probability distribution of the spot price of the commodity at maturity), the price of the option is a function of the exercise price. As before, we refer to this relationship by numerical pairings of option price and exercise price, (d, x) , and note in the case of a call option that there is a range of such pairings, ranging from $(d^{\wedge}, 0)$ to $(0+, x^{\wedge})$, where d^{\wedge} and x^{\wedge} are the maximum values for the respective variables. To our knowledge, there is no systematic explanation of the factors that determine which pairings are selected by options markets or become the most liquid. Casual observation, however, reveals that commodities options tend to be issued near the money; the exercise prices are near the corresponding spot prices. It may be that options are too risky and difficult to value when the exercise price lies at the tail of the distribution of the underlying asset (e.g. far out-of-the-money). Thus, they may be correspondingly more difficult for the writer to hedge.¹²³ Setting aside solvency concerns, therefore, the most common call option may look something like (d, v_0) , where v_0 is the spot price of the commodity at the time the option is issued.

The holder of a call option contract may walk away from the purchase of the underlying commodity without making any payment other than the option price. The ex post loss incurred by the option writer is irrelevant to the price. Moreover, the option price is not intended to be even ex ante compensatory in the sense envisaged by the legal regulation of liquidated damages. It is simply the price at which the seller agreed to bear the market risk shifted by the option. In a thick options market, the option price is determined by the value of the option to the marginal buyer which is equal to the cost of the option to the marginal seller. If the

¹²³We note also that, perhaps due to tax and incentive considerations, executive stock options are typically issued at the money. [cite Hall, Murphy]

Draft: February 22, 2004

option price is negotiated between writer and holder (even though the spot market may be thick), the price will fall somewhere between the value of the option to the buyer and its cost to the seller.

Thus, for the inframarginal option seller in the thick market and for the option seller in the negotiated deal, the option price will be supercompensatory even ex ante. It is supercompensatory in the sense that the seller may be overcompensated for its cost in producing any feature of a good or service delivered to the buyer. The penalty rule for liquidated damages does not, of course, apply to explicit option contracts. However, it does apply to bilateral contracts in thin markets, even if they contain embedded options that serve the same insurance purposes outlined above. Yet, even in thin markets, damages are effectively option prices that often have nothing to do with compensating the seller's ex post losses from breach.¹²⁴

Thin-market contracts differ from their thick-market counterparts in the following two respects. As noted above, the purpose of thick-market contracts is risk allocation. Literal performance is incidental and the parties may even choose to settle through monetary payment rather than physical delivery of the subject matter. Neither party, however, can cancel the contract. Accordingly, there is a sharp distinction between forward and option contracts in thick markets. A thin-market contract contemplates physical performance, and one or both parties may make investments specific to that performance. Yet, thin-market contracts embed options in the form of termination rights. These termination rights serve two purposes described earlier. First, they preserve the real option (e.g. abandonment or call option) that would be held by an integrated firm. Second, they provide an insurance function and may be part of the risk management strategy of one or both parties. The option price and strike price of the buyer's embedded call option are set to optimally allocate between the parties the risk of fluctuations in the buyer's valuation.

¹²⁴See Scott, *The Case for Market Damages*, supra note — at —.

Draft: February 22, 2004

For example, an aircraft manufacturer may be able to diversify the nonsystematic risk of fluctuating passenger demand better than the airlines if the manufacturer has a sufficiently diverse customer base. The manufacturer thereby enjoys a comparative advantage in salvaging the aborted exchange. In that case, the manufacturer may grant the airline an option to terminate during at least part of the manufacturing period.¹²⁵ The price for the option will be negotiated between the two parties and will therefore lie somewhere between its value to the buyer and cost to the seller, depending on their relative bargaining power.

The parties may also prefer that the seller use general manufacturing methods rather than cheaper specialized methods, so that the buyer can realize benefits from the option. The flexibility gains from general investments may thus outweigh the losses from foregone specific investment.¹²⁶ Specifically, the value to the holder of the option might exceed the cost of using general rather than specialized production methods. For example, when Airbus manufactured an increasing number of different aircraft using the same technology platform and common parts, it was able to grant a valuable option to its airline customers to convert their orders into orders for other aircraft.¹²⁷ The same flexibility could support an option to terminate because the platform could be used to service other contracts.

¹²⁵ “We had found that we could double-book some delivery positions, on the basis that not all options were exercised....Was there a way that we could diversify some of this risk by double booking options from two airlines that were perceived to be countercyclical? Our worldwide customer base suggested that we could.” John Stonier (Marketing Director Airbus Industrie), *The Change Process*, in TOM COPELAND AND VLADIMIR ANTIKAROV, *REAL OPTIONS: A PRACTITIONER’S GUIDE* 47 (2001).

¹²⁶ Conventional contract theory examines simply the dollar amount spent on specific investment. More recent scholarship recognizes that investment is specific or general to varying degrees. E.g., Eric Posner, Alexander Triantis and George Triantis, *Investing in Human Capital: The Efficiency of Covenants Not to Compete* (Working Paper 2003). A seller who uses general manufacturing techniques may be able to profit by selling the resulting flexibility in the form of embedded options. John Stonier and Alexander J. Triantis, *Natural and Contractual Real Options: The Case of Aircraft Delivery Options*, in A. MICALIZZI AND L. TRIGEORGIS, EDS., *REAL OPTIONS APPLICATIONS* 159-195 (1999).

¹²⁷ “This facilitated our [Airbus’] ability to allow customers to defer, until as late as possible, the decision as to which size of aircraft within the family they would actually take delivery of.” Stonier, *supra* note -- at 41.

Draft: February 22, 2004

The principal purpose of this discussion is to suggest that, in addition to the pursuit of other contracting goals such as efficient specific investment, thin-market contracts often serve a risk management function. This objective further complicates the structuring of termination rights, which we analyze here as embedded options. In particular, the pricing of embedded call options is likely to be quite heterogeneous across circumstances and this observation motivates our argument in Part III against a default damages provision. However, the pricing itself is an interesting avenue for future analysis and we offer a preliminary outline of the relevant factors in the next section.

E. Volatility as a Determinant of Embedded Options in Thin Markets

If the embedded call option simply entailed the transfer of exogenous risk from buyer to seller, the seller would write the option that maximized the surplus created by the difference between the option's value to the buyer and cost to the seller. As noted earlier, these measures are functions of the multi-variable risk profiles of each party. Given that the underlying exchanges occur in thin markets, the option price is the product of bargaining between the seller and buyer, who will divide the surplus. It is therefore likely that the price the buyer pays for the option will be higher than the cost of the option to the seller. Where the option price takes the form of damages liability, the seller might be viewed as "overcompensated" from this ex ante perspective.

Given the close connection between options and insurance, it should not be surprising that the determinants of the structure of embedded options should also invoke familiar considerations about moral hazard and adverse selection. In both contexts, the parties cannot control for these information problems because the courts cannot verify ex post the cause of any realized loss. Thus, the parties must choose whether to include an embedded option in the face of such concerns and, if so, its exercise price and option price by

weighing the various factors. In the traditional insurance contract, the relevant private information and hidden action lies with the insured party. In the context of an embedded option in a sales contract, however, there are concerns about adverse selection and moral hazard on both sides. We discuss first the seller's private information and control over buyer valuation. The seller's advantage in this respect reinforces the efficiency of shifting the risk to the seller and leads to a pairing of a lower option price and higher exercise price. We then turn to consider the buyer's private information and control that has the same effect as adverse selection and moral hazard in insurance. If it does not deter contracting altogether, it leads to a higher option price (analogous to the combination of a higher premium and/or higher deductible).

1. Seller's Private Information and Control

Even in the absence of superior risk bearing capacity, the seller may provide insurance to the buyer in the form of a call option if the seller has private information about the good that affects the buyer's valuation. For example, assume the buyer's value depends only on the quality of the good, which the seller knows perfectly. A simple contract would face a significant adverse selection problem if the range of possible quality was large. The buyer might offer a price associated with the average value of the product, but this may drive from the market the high quality sellers, leading the buyer to further discount the price. Warranties are well known as signaling devices that bridge the information asymmetry between seller and buyer.¹²⁸ Their enforcement, however, requires a court to determine whether the warranty has been breached. The costs of verifying breach may be large relative to the value of the signal, either because the evidence is indeterminate or because the warranty standard is opaque.¹²⁹ Therefore, the buyer may be given instead the right to return the

¹²⁸ See generally George Akerloff, *The Market for "Lemons:" Quality Uncertainty and the Market Mechanism*, 84 Q. J. ECON. 355, 366 (1970).

¹²⁹ In this respect, we note the vague default standards for warranty liability under the Uniform Commercial Code Under the Code's implied warranty of merchantability, for example, the buyer, in order to establish a breach of

good or cancel the deal without the burden of proving breach or incurring prospective litigation costs.¹³⁰

The ubiquity of sellers' private information is the most common explanation for the return policy of retailers of goods, particularly those whose value may not be transparent.¹³¹ If the only source of volatility in the buyer's valuation were the quality of the equipment, the seller could ensure that no buyer would wish to exercise her right of return. The seller thus would be prepared to issue the option for free. The buyer's valuation of the good is rarely due entirely to factors within the seller's private knowledge, however. To the extent that there are exogenous risks affecting the buyer's valuation, the seller bears a cost in writing the option. She may therefore be inclined to impose a fee or, in the alternative, to bear the cost in order to benefit from the favorable signal. One might think of the right to return in this respect as falling within a category of gifts which sellers give to prospective buyers. Instead of giving the prospective buyer a free dinner, a round of golf or a tote bag, the seller gives a free option.

Suppose that the value of the good to the buyer depends on effort or investment by the seller that is not contractable. In the parlance of economics scholarship, the buyer's valuation is a function of the seller's cooperative investment. For example, assume that the buyer commissions a portrait of herself and her children by a well-known artist whose prior work demonstrates great skill in capturing the essence of his subjects. The

warranty, must show that the good is not "fit for the ordinary purposes for which such goods are used." See UCC § 2-314(2).

¹³⁰In some cases where the seller has private information, the buyer may be able to acquire the information at some cost. But if the investigation requires an investment that the buyer will not recover if she fails to purchase the product, then the seller might hold-up the buyer during subsequent negotiations over price. There are several ways in which the buyer's investment in research may be protected. One way is for the seller to give the buyer an option on the good at a specified exercise price.

¹³¹For example, sellers of electronics equipment. Other examples of seller private information where a right of return is common include oriental rugs, works of art—such as paintings and sculpture—and similar goods where value is a combination of provenance and aesthetic values.

Draft: February 22, 2004

artist might be induced to make the efficient level of investment by giving the buyer a call option with an exercise price equal to the value that results when the seller exerts the efficient level of effort. However, as noted earlier in the essay, economists have shown that this solution is undone when the buyer can force the renegotiation of the exercise price and can thereby hold-up the seller with respect to her specific cooperative investment.¹³² If the seller could precommit not to renegotiate, then the option serves the purpose of inducing the efficient investment or effort by the seller. Retailers, for example, have largely been successful in preventing renegotiation by constraining the authority of their sales agents to modify sale prices. Specifically, an agent who receives a return is not authorized to give a partial refund of the original purchase price.

2. Buyer Private Information and Control

If buyers are heterogeneous in the volatility of their valuations, sellers may not write costly call options. For example, suppose that the fickleness of a buyer's tastes varies among buyers and that the seller cannot discriminate among buyers. The seller would charge an option price that reflected the average sensitivity of its buyers. The less fickle buyers would therefore cross-subsidize the options of the more fickle buyers and the less fickle might exit as a result.

A competitor retailer (an entrant) may see an opportunity to deny its customers the right of return in order to attract the business of the less fickle buyers, thereby leaving the original retailer with a more fickle pool and higher costs that must be recovered by raising prices. There are many possible equilibria, depending on conditions such as the risk aversion of consumers in the market and the volatility of their tastes over time. The entrant might draw away all buyers, leaving an industry without returns. Or, there may be a separation of the market in which the competitors coexist; for example, some restaurants take reservations while their

¹³²Supra note –.

competitors do not. Finally, the original firm may be able to hold on to most of the buyers because of its comparative ability to bear the buyers' valuation risks or because of the presence of significant seller side private information or control. The seller, however, might offer different combinations of option price and exercise price in order to screen her buyers for the volatility of their valuations.

The risk in the buyer's valuation may also be a function of the buyer's efforts and investment. The insurance written by the seller who gives an option gives rise to moral hazard in the buyer's actions, both during and before the contract period. Of course, where the buyer has possession of the good during the option term, she will not have the efficient incentive to take care of it. Moral hazard comes also in more subtle forms. During the contract, the buyer who is protected by insurance will be motivated to forego measures that dampen (or take actions that increase) the volatility of her valuation of the product in order to maximize ex post the value of her option. Thus, for example, a buyer from a retailer with a policy of free returns may forego investing in research during the contract period to determine how well the product fits with the buyer's tastes. Alternatively, when the buyer is an agent of her family, the buyer may fail to invest efficiently in learning whether the product suits the tastes of other family members.¹³³ Because sellers cannot discriminate against buyers on the basis of the risk of return, buyers have the incentive to underinvest in actions that will reduce volatility. The seller, however, can require the buyer to bear some portion of the cost of returns though the expenditure of time in exercising the option even though it is nominally free. Moreover, some returns are limited to a "store credit" which further constrains the buyer's choice and thus motivates the buyer to internalize some of the cost of exercising the option.

¹³³This observation is distinct from Richard Craswell's theory of the effect of damages on search. His focus is on a party's pre-contract search for the appropriate contracting partner. See Richard Craswell, *Offer, Acceptance and Efficient Reliance*, 48 Stan. L. Rev. 481 (1996). We are focusing on the buyer's effort in figuring out the value of the good itself: e.g., whether the family really needs the television.

3. Summary

The preceding discussion reveals that the embedded options we observe in commercial and consumer contexts are a function of the trade-offs that results from the interaction of exogenous shocks and the problems of hidden action and hidden information that affect valuations for sellers and buyers respectively. The volatility in valuations that results from the interaction of these factors stimulates the seller to choose from an continuum ranging from no options ($d^{\wedge}, 0$) to embedded options that are free to the buyer ($0, x^{\wedge}$). The benefits of an option are that it (1) exploits the seller's comparative advantage in insuring the buyer against fluctuations in valuation, (2) signals the seller's private information about product quality, and (3) induces the seller to exert more effort necessary to increase the value of the good to the buyer, at least where the parties can precommit not to renegotiate the exercise price. The costs of using the option are that it (1) deters specific investment by the seller, (2) places exogenous risks on the seller, and (3) may lead to adverse selection and moral hazard. Where the efficiency costs of an option outweigh its benefits (e.g. most of the buyer's volatility comes from factors within the buyer's control), we are unlikely to see options at any price. Where the benefits outweigh costs we might see either a free option or a priced option. A free option is used where the predominant source of uncertainty is private seller information. A priced option is used where the risk is exogenous, but managed by the seller. The central point, however, is that the price of termination options vary widely. We turn now in Part III to a consideration of the ways in which that central insight informs our understanding of optimal damages default rules.

III. OPTION THEORY AND OPTIMAL DAMAGE DEFAULTS

A. The Conditions for Efficient Damages Defaults.

The case for default rules in contracts rests on the assumption that state institutions, such as courts and

legislatures, sometimes can design contract provisions at lower cost than the parties could themselves. The source of this advantage may be economies of scale. Private parties who might bear the cost of creating efficient terms to address contracting problems would not capture the full social gain because they cannot prevent later parties from copying those terms (nor require that they pay a fee).¹³⁴ As a result, contracting parties may underinvest in contract design. The efficiency of default provisions relies on the satisfaction of two key conditions. First, parties must have the freedom to contract for their own terms because agreements occur in heterogeneous circumstances and the state cannot feasibly write contracts that suit all environments. Second, the state's advantage in internalizing the benefit of default provisions to all parties in the economy must outweigh the parties' informational advantage concerning their particular circumstances.

This article focuses on the default damages provisions in contract law.¹³⁵ The state must decide whether to specify a default term or whether to induce parties to create their own damages term by refusing to enforce contracts that fail to do so. As we discussed in Part I(B) above, courts and legislatures over the past 150 years have created a default of full expectation damages. Although various authors have questioned the merits of the default,¹³⁶ no one thus far has doubted the case for a default damages provision or the salience of

¹³⁴For analyses of how legal practice innovations spread, see Gerald F. Davis and Henrich R. Greve, *Corporate Elite Networks and Governance Changes in the 1980s*, 103 *Am. J. Sociology* 1 (1997); Gerald F. Davis, *Agents Without Principles? The Spread of the Poison Pill through the Intercorporate Network*, 36 *Admin. Sci. Q.* 583 (1991).

¹³⁵It has long been recognized doctrinally that, in principle, determining the optimal remedies for breach of contract is no different from the analysis courts use to allocate any other contractual risk. See, e.g., the opinion by Justice Holmes in *Globe v. Landa*, *supra* note —.

¹³⁶See Charles J. Goetz & Robert E. Scott, *Enforcing Promises: An Examination of the Basis of Contract*, 89 *Yale L. J.* 1261, 1281-86 (1980). Recently, Richard Craswell summarized the difficulty in selecting an economically efficient default in his critique of the categories of compensatory damages advanced by Fuller and Perdue. After reviewing the various decisions made by contracting actors (search, precautions, reliance, trade), he suggests that “[t]here is... no reason to suppose that the totality of economic effects will *always* favor an award of expectation damages...[W]henever different amounts of damages would be optimal for each of the different incentives to be optimized, the measure that is optimal when *all* of the relevant incentives are considered will often be some hybrid or

Draft: February 22, 2004

the compensation principle in shaping the preferred default. Yet, the theory of embedded options advanced in this article implies that optimal contract damages have little to do with compensation. Moreover, optimal damages are as context dependent as the price of goods in sales contracts. Therefore, the wisdom of having any default needs to be considered explicitly.

The first condition of an efficient default -- the ability of parties to contract for their own, different term -- is generally satisfied in contract law, but it is violated without good reason in the case of damages for breach. As noted in Part I, a liquidated damages provision is void as a penalty unless the damages are “reasonable in light of the anticipated or actual harm caused by the breach and the difficulty of proof of loss.”¹³⁷ Under this rule, parties must structure their liquidated damages provisions, including alternative termination mechanisms such as pre-paid deposits, so that they accord with the compensation principle. This regulation of liquidated damages would be less problematic if, as courts conventionally assume, parties had no legitimate reason for deviating from the compensation norm. To the contrary, however, a theory of embedded options adds to the existing justifications in the academic literature for liquidated damages that are under- or overcompensatory, thereby bolstering the already strong case against the penalty rule.

The second condition -- the state’s cost-efficiency advantage in designing defaults -- is violated with respect to a wide range of contract terms. In light of the heterogeneity and complexity of commercial and consumer contracts, the economies of scale in fashioning default terms are outweighed by the informational

intermediate number [of the expectation, reliance and restitution interest measure].”

Craswell nevertheless believes that the optimal damages would have some significant link to compensation. Craswell comes to a similar conclusion in his critique of Fuller and Perdue’s three-way classification of expectation, reliance and restitution interests. Craswell, *Against Fuller and Purdue*, supra note -- at 109-10.

¹³⁷ See UCC §2-718; Restatement (Second) of Contracts §356.

Draft: February 22, 2004

advantage enjoyed by the parties themselves. Simple default rules are often inadequate in a complex world.¹³⁸ Not surprisingly, therefore, a relatively small number of default rules have evolved at common law and a much larger set of default standards, addressing a wide range of contracting contexts, have been proposed in Article 2 of the UCC and the Restatement of Contracts. Contract law typically provides broad standards when the efficiency conditions for default rules are not met. The strategy of adopting default standards rather than rules, however, gives rise to a variety of undesirable effects noted below.

In the following discussion we elaborate our argument that the default damages rule in contract law violates the fundamental conditions of efficient defaults. We make several proposals for reform. First, thick market contracts simply allocate market risk and therefore should be enforced by a default based on market damages. Second, in many thin-market contracts, breach damages are effectively a price paid by the promisor for insurance against fluctuations in her valuation of the contract. This price varies widely depending on the source of the volatility in valuation, and a variety of other factors that have no connection to compensating the promisee for her losses. Contracting parties should have the freedom, therefore, to agree to liquidated damages of their choosing. Third, given the heterogeneity of the option price or insurance premium, any default is unlikely to satisfy the second efficiency condition. Thus, in the absence of a negotiated option, the alternatives for the courts are either to specifically enforce thin market contracts that lack options to terminate or to decline to enforce such contracts where they fail to specify termination rights.

B. Optimal Damages in Thick-Market Contracts.

¹³⁸Moreover, state-created default rules may not even be possible where there is asymmetric information. Defaults must condition on information that the enforcing authority is able to observe. A default rule that conditions on unverifiable information would create moral hazard. Consequently, parties will routinely contract out of these possible defaults. For discussion, see Schwartz & Scott, *Contract Theory*, supra note -- at 605-08.

Draft: February 22, 2004

Why do parties enter into fixed-price contracts for future delivery of goods that are traded in a thick market? After all, one can always acquire the goods on the spot market at the prevailing price without negotiating a forward contract. In the pure thick market contract, there is no specific investment, there is no efficient breach, there are no information asymmetries and there is no moral hazard. As we discussed earlier, parties enter fixed price contracts in these markets simply to shift market risks. In the most liquid markets, the parties typically do not contemplate physical delivery, but only a closing of their respective positions through the payment of money. In this light, there is no reason for the court to take any action other than to enforce the risk allocation by awarding market damages.¹³⁹

Consider a contract to deliver at a future date 1,000 tons of a stipulated commodity at a contract price of \$200,000. Suppose that the seller later contracts to fill the order from a third-party supplier for \$150,000. The buyer then breaches and, at the time of performance, the market price of the commodity is \$100,000. The third party supplier subsequently releases the seller from his contract. The market measure of damages is the \$100,000 difference between the market and contract price. The buyer argues, however, that the seller's loss is only \$50,000 in foregone profits and that this compensatory amount is the appropriate measure of damages.¹⁴⁰

Before we determine the optimal default rule to govern the buyer's breach, we must first understand the motivations that lead parties to write this contract. If the buyer only wanted to guarantee a supply of the commodity at the date of delivery, the parties could have contracted to pay the market price on that date. The

¹³⁹ The following analysis of optimal damage measures in thick market contracts draws on the text in Scott, *The Case for Market Damages*, supra note – at 1160-1179.

¹⁴⁰ This facts of this hypothetical are loosely drawn from *Nobs Chemical, USA, Inc. v. Koppers Co.*, 616 F.2d 212 (5th Cir. 1980).

Draft: February 22, 2004

fixed price contract, however, has effectively shifted the risk of price fluctuations. After the contract the seller bears the risk of price increases and, in turn, has purchased the reciprocal opportunity to hold the buyer to the risk of a price decline. The seller can deal with this contract risk in a number of ways. The seller may choose to self-insure and bear the entire risk internally. If so, the seller may wait and purchase the contract goods on the spot market before the contract delivery date. Alternatively, the seller can lay off the risk, as in our example, by contracting to acquire the goods from a third party at a fixed price. This alternative allows the seller to hedge its risk by selling a portion of the contractual opportunity to the third party. The seller will choose the risk bearing strategy that maximizes its expected return over a range of similar contracts. This analysis thus argues for granting the seller damages based on the full \$100,000 market risk.

Driven by the compensation principle, however, most courts have declined to award market damages in similar contexts.¹⁴¹ An especially attractive focal point has been § 1-106 of the UCC, which expressly incorporates the compensation principle.¹⁴² Section 1-106 seems to trump the various arguments for market damages. Thus the conclusion reached by many (but not all) courts is that market damages are not appropriate where the market fluctuation deviates substantially from the ex post economic loss to the promisee. These holdings, however, are based on a misunderstanding of the motivation for thick-market contracts, which is simply the allocation of risk. A blind application of the compensation principle in these thick market contracts encourages inefficient behavioral adjustments by both parties.

¹⁴¹See *Nobs Chemical, USA, Inc. v. Koppers Co.*, 616 F.2d 212 (5th Cir. 1980); *Allied Cannery & Packers, Inc. v. Victor Packing Co.*, 209 Cal. Rptr. 60 (1984); *H-W-H Cattle Co., v. Schroeder*, 767 F.2d 437 (8th Cir. 1985); *Coast Trading Co. v. Cudahy Co.*, 592 F.2d 1074 (9th Cir. 1979); *Union Carbide Corp. v. Consumers Power Co.*, 636 F. Supp.1498 (E.D. Mich. 1986). But see *Tongish v. Thomas*, 840 P.2d 471 (Kan. 1992) (compensation principle does not trump market damages simply because contract-market differential exceeds promisee's economic loss); *Transworld Metals, Inc. v. Southwire Co.*, 769 F.2d (2d Cir. 1985) (same).

¹⁴² "The remedies provided by this Act shall be liberally administered to the end that the aggrieved party may be put in as good a position as if the other party had fully performed...." UCC §1-106(1).

In short, any use of the compensation principle to depart from market damages in thick market contracts is inconsistent with the contracting goal. In these thick-market contexts the courts should simply enforce (specifically in most cases, because the parties do not contemplate actual delivery) the parties risk allocation. If parties prefer an alternative risk shifting mechanism offered by option contracts, the court should similarly simply enforce the risk allocation that they have bargained for.

C. Optimal Damages in Thin-Market Contracts

Most contracting does not occur in the paradigmatic thick market context. The issue of optimal damages rules is more complex, therefore, when we turn to thin-market contracts (or hybrid thick-thin markets). The central insight of the embedded options approach is that in thin or hybrid markets there will be a wide variety of termination provisions. Recall that the measure of damages for breach is merely a subset of a larger family of termination options. Like other termination fees, contract damages are essentially the nonrefundable portion of the contract price. That portion is nonrefundable because it is the price paid for the termination option which, by definition, the buyer does not "return". Whether or not the buyer exercises the option, she pays that price. We would normally expect the price of the option to lie somewhere between the option's value to the buyer and its cost to the seller.

1. The costs of the compensation principle

a. The penalty rule. Once we see the damages award as essentially a price for a service (insurance against fluctuations in buyer's valuation), then the penalty doctrine becomes a major impediment to efficient contracting. Contract scholars have previously demonstrated that the premise underlying the regulation of liquidated damages—that parties have no good reason to write damages clauses that depart from the

Draft: February 22, 2004

compensation principle-- fails for several reasons.¹⁴³ For example, as a result of the doctrines of foreseeability and uncertainty, expectation damages are limited to the losses from breach that the promisee expects to be able to verify to a court. A liquidated damage clause that incorporates observable but nonverifiable values thus will be vulnerable to a penalty claim even when the clause accurately measures the promisee's lost expectation.¹⁴⁴

Option theory supplies an even more devastating critique of the penalty rule and, generally, of the doctrinal mandate that liquidated damages adhere to the compensation principle. If damages are merely termination fees, those fees do not even purport to reflect the loss suffered by the seller from the buyer's option to terminate. As we have seen, the seller may give the buyer a termination option in thin market contracts for several different reasons: to promote efficient abandonment and/or to insure the buyer against fluctuations in the value of the contract performance. When the contract insures the buyer against fluctuations in her valuation of contract performance, any stipulated damages may be supercompensatory. Even viewed ex ante, options to terminate will often reflect the profit (above cost) to the seller for providing insurance to the buyer against fluctuations in the buyer's valuation of the contract performance in much the same way that the seller of a car will make a profit on additional options such as a sunroof and CD player. If termination is

¹⁴³Recently, economists have tackled the dual objectives of efficient ex ante investment and efficient ex post trade under the assumption that the court lacks the information to calculate expectation damages and presumably, therefore, to strike down penalties. Several proposals contemplate very large liquidated damages. For example, Aaron Edlin and Stefan Reichelstein present a solution to the problem arising when both sides of the contract make specific investments, that calls for specific performance of the delivery obligation of the seller. Recognizing that specific performance might not be ordered in some cases, the authors propose that the parties stipulate very large damages in its place. Aaron S. Edlin & Stefan Reichelstein, *Holdups, Standard Breach Remedies and Optimal Investment*, 86 Amer. Econ. Rev. 478 (1996). See also, Aaron S. Edlin & Alan Schwartz, *Optimal Penalties in Contracts*, 78 Chi-Kent L. Rev. 33 (2003).

¹⁴⁴This difficulty with the penalty rule permits promisors to invoke the penalty doctrine strategically. As a consequence, sophisticated parties are discouraged from using liquidated damage clauses even when these clauses would otherwise be optimal. See Charles J. Goetz & Robert E. Scott, *Liquidated Damages, Penalties and the Just Compensation Principle: Some Notes on a Theory of Efficient Breach*, 77 Colum. L. Rev. 554 (1977).

Draft: February 22, 2004

provided through a liquidated damages term, the damages are, in essence, the option fee. The option fee is a function of the value of the option to the buyer not a function of the ex post loss suffered by the seller. Thus, there is no penalty in any meaningful sense of the word. Moreover, our discussion of the ubiquity with which sellers offer undercompensatory options to return goods or services underscores the fact that contracting parties are not concerned with compensation when they structure termination rights. There is no reason why compensation should suddenly be relevant when the termination fee exceeds some measure of compensation.

b. The lost volume confusion. The costs of the compensation principle are not limited to the inefficiencies of the penalty rule. Consider the paradigmatic lost volume case in which the seller's economic loss may systematically exceed market damages.¹⁴⁵ In these cases, scholarly debate has focused on how much of the seller's selling costs were "consumed" by the breaching buyer and whether the default measure of damages ought to be the full profit lost by the seller (which is overcompensatory) or incidental damages (which are undercompensatory). But the focus on lost volume and selling costs is a red herring. Rather, the choice between market damages and lost profits is a choice between alternative prices for writing the buyer an option to terminate. As we have seen, the parties might rationally choose a free option but that would not mean that the seller is not harmed.¹⁴⁶ Rather the option to terminate could be priced below seller's cost as a means of increasing its market share. Moreover, whether a given volume seller would have chosen to write an option to a buyer and at what price the option would be offered simply cannot be determined a priori. And, in any event, the fee paid by the buyer to have the option to terminate would have nothing to do with compensating

¹⁴⁵In lost volume cases, market damages appear to undercompensate the seller for the buyer's breach. This is particularly salient when the market damage measure is zero. Sellers argue that, although they did in fact resell the contract goods, the second buyer would have purchased anyway. Therefore, if the breaching buyer had performed the contract, the seller would have realized two profits rather than just one. Thus, the seller claims the lost profit from the breached contract as compensation for his lost expectancy. SCOTT & KRAUS, CONTRACT LAW AND THEORY at 1091-1105.

¹⁴⁶[Goldberg]

the seller for his loss.

The compensation principle (and the penalty rule) corrupts the way courts apply damages in lost volume cases in even more subtle ways. In particular, the rule prevents sellers of special order goods from writing efficiently priced option contracts that require the buyer to pay a specified nonrefundable deposit. When the buyer breaches, the seller is precluded from retaining the deposit by the express terms of UCC §2-718(2), which effectively limits the deposit to either \$500 or the amount permissible as liquidated damages under the penalty rule.¹⁴⁷ In response to the regulation of deposits, many sellers then sue for lost volume damages.¹⁴⁸ While scholars have criticized the courts for granting lost volume damages in this circumstance, the impulse of the courts may well be to provide the seller a “make up” for the invalidation of the contracted-for termination fee.¹⁴⁹ In turn, the precedential effect of these decisions may lead courts to grant lost volume damages in other contexts, where it is likely that the parties would have preferred a low cost (even free) option to terminate.

2. Expectation Damages Cannot be Justified as a Majoritarian Default

Lost volume sales is just one area in which the measure of expectation damages are a point of controversy. There are many others. Consider, for example, the cases in which a buyer (or lessee) purchases

¹⁴⁷Section 2-718(2) provides that “Where the seller justifiably withholds delivery of goods because of the buyer’s breach, the buyer is entitled to restitution of any amount by which the sum of his payments exceeds

(a) the amount to which the seller is entitled by virtue of terms liquidating the seller’s damages in accordance with [the penalty rule], or

(b) ...twenty percent of the value of the total performance for which the buyer is obligated under the contract or \$500, whichever is smaller.

¹⁴⁸ The offset for lost volume damages is implicitly invited by UCC §2-718(3).

¹⁴⁹See, e.g., *Rodriguez v. Learjet*, 946 P.2d 1010 (1997); *R.E. Davis Chemical Corp. v. Disonics, Inc.*, 826 F.2d 678 (1987).

Draft: February 22, 2004

the right to mine the seller-lessor's land in return for a royalty on extracted ore and a promise to regrade the land upon conclusion of the lease. Assume the buyer-lessee breaches the agreement by failing to regrade the land. Courts have divided (bitterly) over the proper damages default.¹⁵⁰ Should the lessor be entitled to cost of performance damages when the economic loss from breach (as measured by the diminution in market value) is substantially less than the cost of regrading? Once again, the focus on expectation damages as a majoritarian default is misplaced. These grading contracts, just as lost volume sales, each have idiosyncratic circumstances that are critical to the choice of termination option/damages measures. Considering only the insurance features discussed in Part II, the decision of the parties to include a termination option in their contract depends on the interaction of five separate sources of volatility in the lessee's valuation (viz, exogenous shocks and informational asymmetries of both lessee and lessor respectively). The option price, if the termination option were explicitly negotiated between the parties, would depend upon the relative contributions from these various sources. As the parties trade off these factors, they will arrive at a wide range of option prices and, if the termination is a breach provision, an equally wide range of damages measures. Occasionally and only by coincidence will the option price reflect the amount of compensation for the lessor's loss. Then, add to the considerations that trigger the writing of embedded options the other factors that are relevant to the underlying contract, including the trade off between commitment (to encourage relation-specific investment) and flexibility (to encourage efficient abandonment decisions). It soon becomes clear that any attempt to generalize damage measures across these different contexts is doomed.

A straightforward implication of the preceding analysis is that expectation damages cannot be justified as a majoritarian default rule in thin market or hybrid market contexts. An efficient default rule is one that is

¹⁵⁰Compare *American Standard, Inc. v. Schectman*, 439 N.Y.S. 2d 529 (1981) with *Peevyhouse v. Garland Coal & Mining Co.*, 382 P. 2d 109 (1962).

relatively simple in form and suitable for a wide variety of contracting parties.¹⁵¹ The criterion of simplicity is a function of institutional competence. Courts cannot conduct investigations into the efficiency properties of all possible rules and rule combinations. The suitability criterion is even more difficult to satisfy because, as we have shown, contracting parties are exceptionally heterogeneous regarding termination provisions. Default rules covering all the various combinations of option and exercise prices would simply be too expensive to create. There would need to be as many legal rules as there are sets of contracting parties.

It is tempting to argue that a default measure for even a smaller plurality of parties would at least save the costs of contracting in those cases. But writing defaults for minority preferences imposes costs on all those parties who would prefer to contract out. These costs are not trivial. The gain of having a rule appropriate for a plurality of parties yields a lower social gain than one satisfying a majority. Against this limited gain, we must set out the added cost to the majority who must contract away from the default rather than draft their contract on a blank slate. Courts tend to regard state-created defaults as presumptively fair or efficient and this institutional bias raises the cost of contracting out.¹⁵² This is particularly true in the case of damage defaults because contract doctrine explicitly identifies the compensatory purpose of damages and the correspondingly acceptable rationale for liquidated damages: the difficulty of verifying the promisee's loss. Thus, expectation damages impose a cost on the majority of contracting parties who wish to adopt any among

¹⁵¹Schwartz & Scott, *Contract Theory*, supra note -- at —.

¹⁵²See, e.g., *Hayward v. Postma*, 312 Mich. App. 720, 724, 188 N.W. 2d 31,33 (1971) (parties must use clear and unequivocal language to shift liability for risk of loss from seller to buyer); *Caudle v. Sherrard Motors Co.*, 525 S.W. 2d 238, 240 (Tex. Civ. App. 1975) (same); *Davis v. Small Business Inv. Co.*, 535 S.W. 2d 740, 744 (Tex. Civ. App. 1976) (contractual provision purporting to allocate to debtor the burden of “all” expenses incurred in preserving collateral not an “agreement otherwise” sufficient to opt out of UCC §9-207(2)(a)). Moreover, judicial interpreters may be reluctant to give the express language of the contract a meaning that conflicts with the relevant default. See *Nanakuli Paving & Rock Co. v. Shell Oil Co.*, 664 F.2d 772 (9th Cir. 1981) (merger clause that excludes evidence of prior dealings does not bar evidence of usage of trade to alter the price term in the contract). For discussion, see Goetz & Scott, *The Limits of Expanded Choice*, supra at ---.

a range of liquidated damages as termination options.¹⁵³

3. The “No Contract,” Bargain-Forcing Default.

The project of creating optimal damages default rules founders on the costs of rule creation for heterogenous parties that function in complex commercial environments. These costs essentially preclude the creation of a majoritarian “rule.”¹⁵⁴ As with all default rules, one might say that the expectation default is both under- and over-inclusive. Thus, as Craswell has shown, the courts have over time effectively created a default standard based on the broad categories of over-expectancy, expectancy or under-expectancy.¹⁵⁵ But such standards seldom are good fits. Parties typically need specific guidance as to what they are supposed to do, and standards often create moral hazard.¹⁵⁶ The moral hazard difficulty is exacerbated by asymmetric information. Defaults that condition on behavior that parties or courts cannot observe will be exploited for private ends. Thus, parties contract away from them. As a consequence, inefficient default standards only raise transaction costs unnecessarily. Moreover, a standard that accepts the compensation principle as a baseline, neglects the insurance motivation for breach damages described above.

¹⁵³The same criticism may be leveled at the regulation of termination rights under the requirement of good faith. See Goldberg, *Discretion in Long-term Open Quantity Contracts*, supra note —.

¹⁵⁴A rare example of a successful default rule is the market damages rule that applies in thick market contracts; the requirement that a breacher pay the other party the difference between the contract and market prices. The rule applies in only one state of the world, when there is breach; it is simple to apply because the court only compares the contract and market prices; and it is efficient for many parties. Other good default rules are hard to find.

¹⁵⁵See Craswell, *Against Fuller and Purdue*, supra at —.

¹⁵⁶Parties reject standards because they increase the likelihood of moral hazard. Parties sometimes have incentives to take actions that are privately optimal but publicly undesirable. When a standard governs, the promisor who wants to behave strategically must ask herself what a court will later do if the promisor is sued. The vaguer the legal standard and the more that is at stake, the more likely is the promisor to resolve doubts in her own favor. The promisor will thus attempt to maximize private gains at the expense of joint welfare maximization. A standard is efficient, therefore, only when the party on whom it confers discretion would maximize joint returns in the course of maximizing her private gain. Schwartz & Scott, *Contract Theory* at —.

Draft: February 22, 2004

What, then, is the proper role of courts in resolving disputes over damages measures in incomplete contracts in thin markets? It would be appropriate to develop a damages standard if the standard would not create moral hazard. Otherwise, whenever a contract fails to make express provision regarding termination rights, courts have a choice: They could adopt a “no option” default and specifically enforce the exchange. Or, courts might adopt a “no contract” default and decline to enforce the contract on the ground that, lacking material terms, the agreement is too indefinite to enforce.

Contemporary courts exercise both choices in appropriate circumstances, sometimes refusing to enforce contracts that lack material terms on the grounds of indefiniteness, and, sometimes, awarding specific performance. To be sure, abandoning the expectation damages default would require an extension of existing doctrine in either case. The case for specific performance rests on the characterization of the option as a contract plus a put. Where the parties fail to negotiate the exercise price of the put, the “no option” solution would be compelled and the buyer would be required to pay the full purchase price for the goods or services. In this case, specific performance (or the equivalent seller’s action for the price) would be available in thin markets whenever the parties declined to negotiate an exercise price for the put option. The alternative, “no-contract” solution would be consistent with the characterization of the option as a call on the subject matter of the contract. Unless the parties price the call option, courts would lack “a reasonably certain basis for giving an appropriate remedy.”¹⁵⁷

In part because it fits more comfortably with existing doctrine, we argue here for an approach that relies on the indefiniteness doctrine and focuses on the bargain-forcing aspects of a no-contract default. Under

¹⁵⁷ See UCC §2-204(3) (indorsing the indefiniteness doctrine when courts lack the basis for imposing an appropriate remedy).

Draft: February 22, 2004

the common law, courts will declare contracts void for indefiniteness if the parties fail to specify readily verifiable measures of performance in realized states of the world. The indefiniteness rule might be understood as a kind of global bargain-forcing default, one that uses the threat of non-enforceability to encourage parties to specify the solution to certain contingencies themselves. Viewed from this lens, a no-contract, bargain-forcing default in thin or hybrid markets would require parties who negotiate embedded options in their contracts to think of termination and breach as unified issues. Even in consumer contracts, where the relevant termination options are written by merchant sellers, the no-contract result is superior to the current rule as it provides consumers the relevant information as to their contract rights at the time of contract, when they most need it. For example, consumers could then choose more efficiently between free option or no option sellers, depending on their preferences over volatility.

Contrary to conventional wisdom, contemporary courts continue to invoke the indefiniteness doctrine.

¹⁵⁸ A systematic review of the case law shows that American courts continue to dismiss claims for breach of contract on the grounds of indefiniteness, often without granting any relief to the disappointed promisee. If the parties appear to have discarded verifiable information that they might have used at relatively low cost to condition performance, the courts will decline to enforce the agreement legally.¹⁵⁹ On this view, a bargain-forcing default would compel the conclusion that any failure to specify a termination fee was either the result of inadvertence or the agreement was *deliberately* left indefinite. In either case, the agreement would be legally unenforceable.

¹⁵⁸ Conventional wisdom holds that courts should (and do) fill contractual gaps with general standards of reasonableness and good faith. See e.g., 1 ARTHUR CORBIN, CORBIN ON CONTRACTS §95 at 400 (1963). This conventional wisdom is misleading, however. In a recent article, one of us has shown that the indefiniteness doctrine has survived the influence of the UCC and the widespread academic support for filling gaps with vague standards. See Scott, *A Theory of Self-Enforcing Indefinite Agreements*, supra note – at ---.

¹⁵⁹ Id. at --.

Draft: February 22, 2004

The 'no contract' default is not as radical as it first appears. Once we see the damage measure in these negotiated contracts as a price for a service (insurance for the buyer), it seems obvious that there will be no default price for the service owing to the heterogeneity of such services in large economies. This rule, thus, is no different from the indefiniteness rule that courts apply when parties fail to provide a contract price for services for which there is no established market price.¹⁶⁰

CONCLUSION

Many consumer and commercial contracts contain explicitly negotiated termination provisions. We have analyzed these provisions as embedded options that serve a valuable risk management function. We have focused attention on the buyer's option to terminate as a type of insurance against decreases in the buyer's valuation of the exchange. The factors that determine the pairing of option price and exercise price depend on the sources of volatility of the buyer's valuation. The price of the option is thus a function of its value to the buyer and its cost to the seller. Only by coincidence will that option price reflect the ex post loss suffered by the seller if the option is exercised. Contract damages, in turn, are default terms that specify a termination fee in the absence of an agreement to the contrary; they are thus a subset of the array of termination provisions. But unlike other substantive termination provisions, both the default measure of damages and the freedom to contract away from it are constrained under current law by a rigid adherence to the compensation principle.

¹⁶⁰The Uniform Commercial Code authorizes courts to provide a reasonable price where the parties have not specified a price. UCC §2-305(1). This provision, however, simply follows the line of cases holding that price terms in sales contracts could be supplied from evidence of market prices. *Varney v. Ditmars*, 217 N.Y. 223, 111 N.E. 822 (1916). But where there is no relevant market price, courts continue to decline to enforce such agreements on the grounds of indefiniteness. *Cobble Hill Nursing Home, Inc. v. Henry and Warren*, 74 N.Y. 2d 475, 548 N.E. 2d 203 (1989). For discussion, see Robert E. Scott, *A Theory of Self-Enforcing Indefinite Agreements*, 103 Colum L. Rev. 1641, 1645-59 (2003).

Draft: February 22, 2004

The current disregard for the compensation principle in commercial and consumer practice, and its irrelevance to scholars working in contract theory, prompts an investigation into the emergence and development of the compensation principle in contract law. The common law courts adapted the action in assumpsit that had developed in tort in order to provide a remedy for breach of informal promises. This solution to a perceived gap in the common law brought with it the tort concept of ex post compensation for the plaintiff's injury. Until the development of commodities markets in the early 19th century, "compensation" in contract meant reimbursement for reliance expenditures for partially executed transactions. Fixed price forward contracting stimulated the development of a market damages default as the measure of the plaintiff's expectancy in the contract. This narrow conception of expectancy migrated to negotiated executory contracts as well. Over time, as parties were able to establish their consequential losses with greater certainty, the measure of expectation damages expanded to include the award of lost profits.

The penalty rule developed similarly along a separate but parallel track propelled by the peculiar institution of the penal bond. Not until the mid-19th century did courts explicitly link the regulation of liquidated damages provisions with the compensation principle. Thereafter, liquidated damages were enforceable only if ex post compensation was uncertain, remote or difficult to measure. But the expansion of expectancy to include lost profits and the erosion of the *Hadley v. Baxendale* rule reduced ex post uncertainty and thus narrowed the grounds for stipulating damages.

Option analysis severely undermines any contemporary justification for the expectation damages default and the penalty rule. A focus on the insurance function of termination provisions shows that the expectation default fails to satisfy the efficiency conditions for state-supplied contract terms and that the current regulation of liquidated damages impedes the ability of contracting parties to write welfare enhancing

Draft: February 22, 2004

agreements. In this light, our project in this article is to trim back the unruly branches of the common law. We are engaged in good husbandry not radical surgery.