

ISSN 1045-6333

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

IS BREACH OF CONTRACT IMMORAL?

Steven Shavell

Discussion Paper No. 531

11/2005

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center/

The Social Science Research Network Electronic Paper Collection:
http://papers.ssrn.com/abstract_id=#####

Is Breach of Contract Immoral?

Steven Shavell*

Abstract: When, and why, might it be thought immoral to commit a breach of contract? The answer to this fundamental question is not obvious, because, as is stressed, and as has been overlooked in addressing the question, contracts do not usually provide explicitly for the particular events that are observed to occur. When a contract does not expressly address a contingency that occurs, the morality of breach is assumed here to depend on what the contract would have said had it addressed the contingency. This assumption is explained to imply that breach is not immoral if expectation damages would have to be paid for breach, but that breach might be immoral if damages are less than the true expectation, as is probable. This conclusion is related to the results of a survey that was conducted of individuals' attitudes toward the morality of breach. The conclusion is also related to the views of commentators on the morality of breach and of those on the "efficiency" of breach.

* Samuel R. Rosenthal Professor of Law and Economics, Harvard Law School. I wish to thank Ian Ayres, Oren Bar-Gill, Omri Ben-Shahar, Louis Kaplow, and Eric Posner for comments, David Katz, Gary Lee, Michael Simkovic, and Jane van Lare for research assistance, and the John M. Olin Center for Law, Economics, and Business at Harvard Law School for research support.

Is Breach of Contract Immoral?

© 2005. Steven Shavell. All rights reserved.

1. Introduction

The question posed here is when, and why, it might be thought immoral to commit a breach of contract. The view that there is something wrong with a person's breaking a contract, or, equivalently, that a person ought to meet his or her contractual obligations, is widely recognized. The Restatement of Contracts, for instance, refers to the "sanctity of contract and the resulting moral obligation to honor one's promises,"¹ and a well-known commentator avers in a typical statement that a "contract must be kept because a promise must be kept."²

Yet it is manifest that contracts are often disobeyed and that the law permits this without imposition of rigorous sanctions – the party in breach generally is obligated to pay only damages to the victim of the breach; the party is not usually punished or compelled to perform.³ The ability to commit breach and pay damages stands in apparent opposition to, or at least raises questions about, the moral duty to perform. The Restatement and commentators seem to be of the opinion that breach and payment of

¹ RESTATEMENT (SECOND) OF CONTRACTS ch. 16, Introductory Note (1981).

² CHARLES FRIED, CONTRACT AS PROMISE: A THEORY OF CONTRACTUAL OBLIGATION 17 (1981).

³ RESTATEMENT (SECOND) OF CONTRACTS ch. 16, Introductory Note (1981) states that "The traditional goal of contract remedies has not been compulsion of the promisor but compensation of the promisee for the loss resulting from breach...punitive damages have not been awarded for breach of contract, and specific performance has not been granted where compensation in damages is an adequate substitute for the injured party." *See also*, for example, JOHN D. CALAMARI & JOSEPH M. PERILLO, THE LAW OF CONTRACTS 611 (4th ed., 1998) stating "The primary relief that the Anglo-American legal systems offer is substitutionary relief, normally damages....Specific performance is an extraordinary remedy...."

damages is tolerable, or sometimes is even desirable, for practical, economic reasons.⁴ Some authorities have seemed almost to celebrate the option to commit breach despite its negative moral aspect, notably Holmes, who wrote “The duty to keep a contract ... means ... that you must pay damages if you do not keep it, – and nothing else ...” even though “such a mode of looking at the matter stinks in the nostrils of those who think it advantageous to get as much ethics into the law as they can.”⁵

Against this background, I ask in section 2 whether it is immoral to break a contract if damages would have to be paid as a consequence. To consider the question, one must, of course, define moral behavior, and I assume that performance is morally required in a contingency that arises if *either* the parties explicitly considered the contingency in their contract and agreed that there should be performance in it – *or*, if they did not explicitly consider the contingency in their contract, that they know that they would have agreed on performance had they provided for the contingency. Given this definition of moral behavior, the primary conclusion that I reach is that it is not immoral to commit breach if expectation damages would have to be paid for breach, but that it might be immoral to commit breach if damages are less than the true expectation measure, as is probable.

The kernel of the argument that I elaborate is simple. A breach that occurs in a contingency that was *explicitly* addressed in a contract is obviously immoral under the

⁴ RESTATEMENT (SECOND) OF CONTRACTS ch. 16, Introductory Note (1981) (observing that circumstances may change, leading a party to want to breach to avoid losses, and that this economic reason for breach is consistent with traditional common law decisions). *See also* E. ALLAN FARNSWORTH 755-756, 761-763 CONTRACTS (3rd ed., 1999).

⁵ Oliver W. Holmes, Jr., *The Path of the Law*, 10 HARV. L. REV. 457, 462 (1897). Although this oft-cited quotation suggests that Holmes found breach and payment of damages morally unobjectionable, I note in section 5 below that that is not a proper interpretation of his meaning.

definition of moral behavior that I employ.⁶ However, I suggest that parties often do not explicitly provide for contingencies for a variety of reasons – owing to the cost of so doing, difficulties courts would face in verifying the occurrence of contingencies, the ability to commit breach, and the opportunity to renegotiate terms. Typically, the conditions that are mentioned in contracts are not in fact singleton contingencies, they are instead events, groupings of individual contingencies, so that what the parties really would want in a specific contingency tends to be left unaddressed.

Hence, I view the case of breach that occurs in a contingency that was not expressly considered as important, if not central.⁷ In this context, we must determine what the parties would have agreed to had they provided for the contingency in order to ascertain whether the breach was immoral. But the fact of breach and the payment of expectation damages allow us to make an important *inference*: the party in breach would not have been willing to commit breach unless the cost to him of performing (or the benefit of not performing) exceeded expectation damages. That the cost of performing would be this high is shown to imply that, had the parties bargained about a term addressing the contingency, they would have agreed that there would be no obligation to perform in the contingency. In other words, the parties should know, and we can deduce, that the breach was not immoral. This inference cannot be made if the measure of damages for breach is less than the true expectation. If that is so, the fact of breach does

⁶ Since the definition says that one has a moral obligation to obey an explicit contractual provision. Still, the reader might ask how it can be immoral to breach an explicit provision yet also, as I claim, not be immoral to breach and pay expectation damages? The resolution of this apparent conflict is given below; see note 22 *infra*.

⁷ A reader who views this statement skeptically is probably not interpreting an explicit provision for a contingency in the manner that I do, and such a reader should reserve judgment until he or she has considered section two below, especially the subsection entitled “*The observed incompleteness of contracts.*”

not imply that the parties would have agreed to nonperformance in the contingency; they might or might not have.

In section 3 of the article, I ask whether breach in reality is likely to be immoral, given the analysis in section 2. The main point that I make flows from the observation that in practice, damages for breach are likely to be less than fully compensatory, and perhaps substantially so. Therefore, the willingness of a party to commit breach does not imply that nonperformance would have been agreed to had the contracting parties considered the contingency that occurred, and breach might be immoral. Hence, it is necessary to consider how strong the reason for breach was to know if breach was immoral, when a contingency was not explicitly addressed in a contract. (But this inquiry is not necessary when a contingency is explicitly addressed, for then the breach is definitely immoral.)

In section 4, I report on a survey that I conducted about individuals' attitudes toward breach of contract. The survey shows that if individuals are asked the general question whether breach is unethical, they usually respond that it is. If, though, individuals are told that the breach occurred in a contingency that was not provided for in the contract and that, had the contingency been addressed, the contracting parties know that performance would not have been stipulated, individuals no longer tend to view breach as wrongful.

In section 5, I review literature considering the morality of breach and also that on so-called efficient breach. These literatures do not consider the point stressed here, that contracts are frequently incomplete and when so do not exemplify the kind of promise to

which a moral duty of performance would necessarily attach.⁸ Relatedly, literature concerned with the morality of breach does not recognize the inference that can be drawn from the willingness of a party to commit breach and pay expectation damages. Because of these omissions, that literature accords the moral duty to obey contracts different weight from the weight that is suggested by the logic of this article.

In section 6, I conclude with remarks on the implications of this article for the moral and legal advice that commentators on contracts sometimes offer us.

2. The Morality of Breach: In Principle

Here I develop the main analytical argument outlined above. Namely, I define moral behavior by reference to detailed contractual provisions, explain why contracts tend to omit explicit mention of many contingencies, and then analyze breach in unprovided-for contingencies given that expectation or lesser damages must be paid.

Definition of moral behavior in a contingency – that delineated in an actual or a hypothetical contractual provision that explicitly addresses the particular contingency. I assume that if a contract provides explicitly for a particular contingency, then the moral duty of a party is unambiguous if that contingency arises: the party is obligated to perform a given act if and only if the contract provision addressing the contingency states that the act is to be performed. Consider a contract under which a seller is supposed to clear a person's driveway if it snows and the contingency that the seller's snow-clearing equipment is stolen. If the contract mentions this possibility and specifies that if it occurs,

⁸ However, two of my own writings emphasize that under the expectation measure, breach occurs in contingencies when parties would not have specified performance in a completely detailed. See originally Steven Shavell, *Damage Measures for Breach of Contract*, 11 BELL J. ECON. 466 (1980), and see also LOUIS KAPLOW & STEVEN SHAVELL, FAIRNESS VERSUS WELFARE ch. IV (2002).

the seller still has an obligation to clear snow (perhaps because he can rent snow-clearing equipment), then the seller is assumed to have a moral duty to clear snow if his equipment is stolen.⁹ But if the contract says that the seller is excused from having to clear snow if his equipment is stolen, the seller would not have a duty to perform if his equipment is stolen.

I assume also that if a contract is not explicit about a contingency, then the moral duty of a party if the contingency arises is determined by what the contract would have said had it provided explicitly for the contingency – supposing that what the contract would have said is known to the parties.¹⁰ Thus, if the contract for clearing a driveway of snow does not address the contingency that the seller’s snow-clearing equipment is stolen but, had the contract included a clause covering theft of the equipment, the parties know that that clause would have stated the seller still has an obligation to clear snow, the seller would have a moral duty to clear snow if his equipment is stolen; and so forth.

The assumption that I am making – that a person has a moral duty to perform an act given a contingency if and only if the contract did state explicitly, or would have stated in an explicit provision, that the person was to perform the act in the contingency – derives from the view that a contract is a species of a promise, for which there are, of

⁹ By a particular contingency I mean a fully described contingency, or at least one leaving out no aspects of possible relevance to contracting parties. Hence, the statement that snow-clearing equipment is stolen might not be considered fully described, for it does not say whether or not the seller can, for instance, rent snow-clearing equipment. However, for expositional simplicity, I will assume that the statement that snow-clearing equipment is stolen is a full description of the contingency. On the notion of a completely specified contingency (or “state of the world”), a notion that is basic to the theory of probability and of expected utility, see the classic discussion in LEONARD J. SAVAGE, FOUNDATIONS OF STATISTICS ch.2 (2nd rev.ed., 1972).

¹⁰ In reality, what a contract would have said about a particular contingency might not be known with confidence by the parties, implying that they might not know their moral duties with confidence.

course, well known grounds for finding moral obligations.^{11,12} For my purposes, however, it is not necessary to discuss grounds rationalizing the morality of promise-keeping; it is enough to suppose that, if a person makes a promise to do something in a given contingency, he has imposed on himself a moral obligation to do that thing if the named contingency arises.

The nature of obligations to perform in a completely detailed contract. Given the assumption that the moral duty of a person to perform an act in a contingency rests on whether the contract did, or would have, required the act to be performed in the contingency, it is of interest to identify the character of the contingencies under which contracting parties would be likely to specify that an act is to be performed.

Let us focus on a contract to produce something or to provide a service.¹³ It will now be argued that we would expect *the buyer and the seller to agree to a contractual*

¹¹ See, e.g., DAVID HUME, TREATISE OF HUMAN NATURE § III.II.V (1739), IMMANUEL KANT, GROUNDWORK OF THE METAPHYSICS OF MORALS 15, 32, 38 (Mary Gregor trans. & ed., Cambridge Univ. Press 1997) (1785), W.D.ROSS, THE RIGHT AND THE GOOD ch. 2 (1930), John Rawls, *Two Concepts of Rules*, 64 PHIL. REV. 3 (1955), John Searle, *How to Derive "Ought" from "Is"*, 73 PHIL. REV. 43 (1964), and Thomas Scanlon, *Promises and Practices*, 19 PHIL. & PUB. AFF. 199 (1990).

¹² A different definition from mine of moral contractual behavior is that contractual behavior is moral as long as it was anticipated by both parties – so that breach would be moral in a contingency as long as breach was expected to occur by both parties (even if performance would have been agreed to had the parties provided explicitly for the contingency). Under this definition, breach would tend to be moral more often than when I find it to be moral. If one employs this definition of moral contractual behavior (or some variant of it), the thrust of this article would not be altered, namely, that what is facially a breach of a contract might not be immoral. That in turn would be so, because (as will be discussed) contingencies are usually not explicitly provided for in a contract, meaning that what is anticipated in a particular contingency cannot be read directly from the contract.

¹³ There are other important types of contract, such as contracts to convey property. The analysis of other types of contract would be similar in its general structure: (a) one would examine the nature of a completely detailed contract, ascertaining when performance and when nonperformance would be provided for; and then (b) one would compare the contingencies in which incomplete contracts are breached to the contingencies in which nonperformance would be allowed in a completely detailed contract. If this analysis were undertaken for contracts to convey property, the particular conclusions reached would be somewhat different from those obtained here. The main reason is that the parties would not have a strong affirmative reason to allow for nonperformance. See Steven Shavell, *The Design of Contracts and Remedies for Breach*, 99 QUART. J. ECON. 121 (1984) and Steven Shavell, *Specific Performance versus Damages for*

provision to undertake a task in a particular contingency if and only if the cost would be less than the value of performance to the buyer. In particular, if the cost would exceed the value to the buyer in the contingency, we would expect the contractual provision to excuse the seller from having to perform. Suppose, for instance, that the value to a person of having his driveway cleared of snow is \$200 (say in terms of time saved and convenience) and that, in a normal contingency, the cost of clearing the driveway would be \$50, whereas in an unusual contingency in which the seller's snow-clearing equipment is stolen, the cost of clearing the driveway would be \$500 (imagine that this would be the cost of renting equipment). Then we would predict the outcome of bargaining would be a clause addressing the normal cost contingency that requires the seller to clear the driveway, and a clause addressing the unusual cost contingency that excuses the seller from having to clear the driveway. That this would be the predicted outcome is readily seen if we imagine that the parties bargain about a separate price for each contingency. Clearly, for the contingency with the \$50 cost, any price for clearing the driveway between \$50 and \$200, such as \$100, would be satisfactory to the two parties; the buyer and the seller would each prefer a provision requiring the driveway to be cleared for a \$100 price to a provision not requiring the driveway to be cleared and with no payment being made (if the price is \$100, the buyer would obtain a net gain of \$100 and the seller would obtain a profit of \$50). However, for the contingency with the \$500 cost, there is no price for clearing the driveway that would be satisfactory to the two parties; the buyer would be willing to pay at most \$200 for the provision requiring the driveway to be

Breach of Contract: An Economic Analysis, Harvard Law School (2005) (unpublished manuscript), for analysis of the nature of completely detailed contracts to convey property. A full discussion of the morality of breach of contracts to convey property would be distracting for the purposes of this article.

cleared, but the seller would demand at least \$500 for the provision; so the provision would not require the seller to clear the driveway.¹⁴

The general qualitative point, then, is that in a contractual provision devoted to a particular contingency, the seller would have a duty to perform if it would not be too hard to perform; but the seller would not have a duty to perform if it would be sufficiently difficult to do so, because the buyer would not be willing to pay enough to induce the seller to agree to perform in such a problematic contingency.

The observed incompleteness of contracts. In reality contracts are far from completely detailed. It is evident that they usually mention relatively few contingencies explicitly out of the multitude of possible relevance.¹⁵ A contract for removing snow from a person's driveway might mention several conditions – such as whether clearing is to be done if it snows on a holiday – but not a practically endless number of events that

¹⁴ Similar, but somewhat more complicated, logic leads to the same conclusion – that the provision for the \$50 cost contingency will specify clearing of the driveway but not the provision for the \$500 cost contingency – if we assume that there is a single contract price rather than a different contract price for each cost contingency (as in the text). Suppose, for instance, that the parties are considering a contract in which the contract price would be paid at the outset and they are bargaining over this price and the two contingent provisions. Imagine that they initially discuss a contract price of, for instance, \$100, and provisions that require the driveway to be cleared whether the cost is \$50 or \$500. Assume that the likelihood of the \$50 cost is 90% and that the likelihood of the \$500 cost is 10% (we need to be explicit about the likelihoods in this note). Then the value of the contract to the buyer would be $\$200 - \$100 = \$100$, since the buyer would be guaranteed performance. The expected value of the contract to the seller would be $\$100 - (90\% \times \$50 + 10\% \times \$500) = \5 . Suppose the seller proposes to the buyer that the candidate provision requiring that he perform in the high cost contingency be altered so as to excuse the seller from having to perform, and that the seller offers to lower the price from \$100 to \$70 to induce the buyer to agree to the change. The expected value of the new proposed contract to the buyer would be $90\% \times \$200 - \$70 = \$110$, so the buyer would be better off (by \$10). The expected value of the new contract to the seller would be $\$70 - 90\% \times \$50 = \$25$, so the seller too would be better off (by \$20). Since both would be better off under the proposed contract with the seller not having to perform in the high cost contingency, they would be likely to decide to switch to the proposed contract. More generally, it can be shown that, for any contract in which the seller does not perform if and only if the cost is less than the value of performance, there exists a substitute contract (generally with a different price) in which the seller is required to perform if and only if cost is less than the value of performance, such that *both* parties prefer this substitute contract. This is why it is said that parties will tend to elect contractual provisions of the claimed nature if they are making a completely detailed contract. For a proof, see Shavell, *supra* note 8.

¹⁵ Of course, if the contract is verbal, the number of contingencies that can be included is small, given the limited ability of individuals to remember what was said and, that aside, difficulties of proof.

could matter to the seller – theft of his snow-clearing equipment, illness of his crew, snow so deep that it makes roads impassable for the trucks transporting the snow-clearing equipment¹⁶ – or that could matter to the buyer – unexpected travel out of town over the winter, sale of home, inheritance of snow-clearing equipment.¹⁷ Although contracts that are the product of substantial legal effort will address numerous contingencies of potential significance, they will still not provide expressly for many that could be of consequence.¹⁸

An issue concerning the definition of an incomplete contract is very important to mention for clarity. Suppose that a contract states that “snow is to be cleared from the buyer’s driveway if the snow is over five inches deep,” and that the contract mentions no other conditions. This contract does not provide *explicitly* for the contingency that the seller’s snow-clearing equipment would be stolen. But the contract does *implicitly* cover that contingency, for in a formal sense the contract covers *all* contingencies: it divides them into two general categories, those in which the snow is up to five inches deep (whatever else happens), and those in which the snow is over five inches deep (whatever else happens).

For our purposes, then, it needs to be kept in mind that what is meant by contractual incompleteness in regard to a contingency is not that a contract fails to cover the contingency in a formal, implicit sense through the use of clauses depending on

¹⁶ These are all examples of contingencies affecting the cost of performance for the seller. Also of possible importance are contingencies in which the seller might be offered a high price to clear snow for someone else.

¹⁷ These are examples of contingencies lowering the value of performance to the buyer and thus that might lead him not to want performance.

¹⁸ See, e.g., the cases mentioned in FARNSWORTH, *supra* note 4, at § 17.6.

general events, but rather that the contract fails to address the contingency in an explicit sense. The reason that I stress this point, and that it could be said to be crucial, is that in reality, many (and, I would say, on reflection, most) conditions mentioned in contracts are not single contingencies, they are events, agglomerations of individual contingencies. Hence, when a particular contingency does occur, it may well not have been explicitly addressed in the contract, meaning that we will not truly know what the parties would have specified in a provision for the actual contingency at issue.

Explanation for the incompleteness of contracts. A number of explanations have been offered for the observed incompleteness of contracts.¹⁹ Most obviously, it takes time to discuss and to include contingent provisions in contracts. If a contingency is unlikely, this effort may not be worthwhile in view of the low probability that a provision for it would prove useful.²⁰

Second, many contingencies would be hard for a court to verify. Consider a provision excusing the seller from having to clear snow if he had back pain. If it would be infeasible or expensive for a court to verify that an individual experienced back pain, then the provision would be problematic, as it could allow the seller to claim the excuse of back pain opportunistically. Hence, provisions for contingencies might not be desirable to

¹⁹ On why contracts are incomplete, *see, e.g.*, Oliver D. Hart, *Incomplete Contracts*, 2 NEW PALGRAVE DICT. ECON. 752 (John Eatwell, Murray Milgate, & Peter Newman eds., 1987), Alan Schwartz, *Incomplete Contracts*, 2 NEW PALGRAVE DICT. ECON. & LAW. 277 (Peter Newman ed. 1998), and STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 299-301 (2004).

²⁰ Suppose that including a contingent provision would take up fifteen minutes of time for each of two lawyers negotiating a contract and that each charges \$250 an hour. The legal cost of the provision would then be \$125. Suppose too that including the provision would save the parties \$1,000 if the contingency were to occur. Then if the likelihood of the contingency is below 12.5 per cent, the expected value of the provision would be less than \$125, so it would not be worthwhile including. More generally, if c is the cost of a contingent provision and b is the benefit of including it, the provision will not be worthwhile including if its likelihood is below c/b .

include in contracts when the courts would encounter difficulty in confirming their occurrence.

A third reason why we would expect only limited use of contingent provisions is that our legal regime, under which parties usually are able to commit breach and pay damages, serves as an implicit substitute for contingent provisions. As will shortly be discussed, under this regime, a party will be motivated to perform if the cost of so doing is not high, in order to avoid paying damages; whereas he will be led to commit breach if the cost of performing is high, because paying damages will be less expensive than performing. This behavior – performing when the cost is below a threshold and not performing when the cost would exceed a threshold – is in at least qualitative alignment with the performance obligations that would be set out in a contract that is completely detailed about cost contingencies, reducing the need to have such a contract.

Still another reason why it may be rational for parties not to take pains to include many contingent provisions in a contract concerns the general possibility of renegotiation of their contract. The parties can anticipate that if they do not provide for a troublesome contingency and it occurs, they will often be able to renegotiate and resolve their problem. If, for instance, the seller finds that it would be unexpectedly costly to perform when the contract requires that, he might be able to obtain a release from his obligation by paying the buyer some bargained-for sum. Of course, the outcome of such renegotiation may be uncertain and it may introduce an added risk into a contract, but the possibility of successful renegotiation lessens the consequences of failure to make contingent contractual provisions.

In all, then, we see that there are important and plausible reasons explaining the observation that contracts are substantially incomplete.

The question of the morality of breach when contracts are incomplete. To ascertain whether a breach in a contingency that was not explicitly provided for is moral or immoral under our definition, one needs to determine whether performance would or would not have been required had the contingency been expressly addressed, and whether the parties to the incomplete contract know this. For an onlooker to characterize a breach as immoral when the contract does not address the particular contingency that occurred would be for the onlooker to fail to recognize that, but for the practical reasons for incompleteness that have been discussed, the contract would have included an explicit provision and that it might have allowed nonperformance. I now consider the question of what the content of a contractual provision that was not actually made would have been if it had been made and whether the parties have reason to know it. This question is answered, first under the assumption that damages for breach equal the expectation measure and then that the damages do not equal the expectation measure.

The morality of breach of incomplete contracts when damages equal the expectation measure. When sellers have to pay damages for breach, they will be motivated to obey the contract if the cost of performance is less than the damages they would have to pay for a breach. If, though, the cost of performance exceeds the damages they would owe for a breach, they will have a financial reason to commit breach. Hence, they will tend to commit breach if and only if the cost of performance exceeds the measure of damages.²¹

²¹ For simplicity, I set aside litigation expenses, reputational concerns, and other factors apart from damage payments that may affect the decision to commit breach.

It follows that when the measure of damages equals the expectation measure, sellers will be led to commit breach if and only if the cost of performance exceeds the value of performance to buyers. But this is exactly when a seller would have been excused from performing in an explicit complete contract, as was explained above. In other words, *under the expectation measure of damages for breach, the seller will fail to perform in the same contingencies as the seller would be permitted not to perform in a complete contract. Accordingly, breach should not be characterized as immoral* under our assumptions.²² Note too that, at least in principle, the seller knows this, that is, in what contingencies breach would and would not be required in a complete contract, presuming that the seller knows the expectation measure of damages, for the seller compares his costs of performance to the expectation measure of damages.²³ Furthermore, we as onlookers know that when breach occurs, it must be moral, for we can infer that the cost of performance must have been higher than the value of performance from the willingness of the seller to commit breach.

²² This conclusion raises the question mentioned in note 6 *supra*. Namely, what if a party breaches an explicit contingent provision of a contract and pays expectation damages? Such a breach should be immoral since the contingent provision is explicit, yet the conclusion that I have reached is that payment of expectation damages makes the breach moral. The resolution of the apparent conflict is that it cannot happen that a party would be willing to pay expectation damages to breach a truly explicit contingent provision. Recall that I showed that, if an explicit provision calls for performance, it must be that the cost of performance is less than the expectation. Hence, in such a contingency, the seller will not find it worthwhile to commit breach. The reasons that in reality we do see breach of explicit contingent provisions lie mainly in two factors: that damages are really less than the expectation (see the discussion in section 3 below); or that an apparently explicit provision is not really explicit (see the discussion above in the subsection “*The observed incompleteness of contracts*”).

²³ That the seller knows the measure of damages is a natural analytical assumption. If the seller does not know the measure of damages, then the conclusions obviously change; the seller would commit breach based on his probabilistic views of damages, and whether breach would be immoral would be, in his opinion, a matter of probability.

In the contract concerning snow removal, if the seller commits breach after his equipment is stolen and pays expectation damages of \$200, the cost to him of performance must have exceeded \$200. This means that, had the parties discussed the theft contingency, the seller knows (and we know) that he and the buyer would have agreed that there would be no duty to perform in the contingency, and thus that the failure to clear snow is not immoral.

The morality of breach of incomplete contracts when damages fall short of the expectation measure – when damages fail to make the victim whole. Now assume that damages are less than the amount necessary to make the victim of a breach whole. Then since breach will tend to occur whenever the cost of performance exceeds the level of damages, *breach will occur more often than nonperformance would have been permitted in a completely specified contract – and thus breach might be immoral.* In our example, if the measure of damages is \$125 instead of the expectation of \$200, breach will occur whenever the cost of performance exceeds \$125. Consequently, if breach occurs when the cost is between \$125 and \$200, for instance when it is \$150, the complete contract would have insisted on performance. Such breach would be immoral, if the seller realizes that the true expectation is \$200.

3. The Morality of Breach: In Practice

Given the conclusions reached in the prior section, what can be said about whether the breach that we see in practice is moral or immoral? If damages tend to be fully compensatory, we could say that breach tends to be moral, as breach should occur if and only if contracting parties would have allowed nonperformance had they addressed in

their contracts the contingencies that engendered breach. But if damages are not really compensatory, breach might be immoral.

Are damages fully compensatory in reality? The expectation measure is, of course, the general damages remedy employed for breach of contract, and the expectation measure is defined as the amount that would restore the victim of a breach to the position that he would have enjoyed had there been performance.²⁴

Yet, as many observers have noted, the expectation measure as it is actually applied tends not to be fully compensatory and may leave the victim of a breach substantially worse off than he would have been had there been performance.²⁵ The reasons given for believing the expectation measure often to be undercompensatory include the following. First, courts are reluctant to credit hard-to-measure components of loss as damages. Hence, lost profits and idiosyncratic losses due to breach are likely be inadequately compensated or neglected. Second, courts are inclined to limit damages to those that could have been reasonably foreseen at the time the contract was made; consequential losses and other hard-to-anticipate losses frequently are not redressed. Third, damages tend not to reflect the often considerable delays that victims of breach suffer. Fourth, legal costs are not compensated.

Not only do expectation damages appear to be significantly undercompensatory in a general sense, damages for breach are probably effectively nonexistent unless the breach victim's losses exceed a threshold of at minimum several thousand dollars – for

²⁴ RESTATEMENT (SECOND) OF CONTRACTS §§ 346-247 (1981) and, for example, CALAMARI & PERILLO, *supra* note 3, at § 14.4.

²⁵ A good account of why damages are undercompensatory is given in Melvin A. Eisenberg, *Actual and Virtual Specific Performance*, 93 CAL. L. REV. 975, 989-996 (2005); *see also* Alan Schwartz, *The Case for Specific Performance*, 89 YALE L. J. 271, 276 (1979). What follows is a summary of reasons given by these authors.

losses must be greater than the cost of bringing suit for the breach victim to have a credible threat to litigate, and the minimum costs of bringing suit are arguably at least a few thousand dollars.²⁶

In view of the foregoing, the practical reality seems to be that breach could be immoral or moral, that we have to inspect the reasons for breach and the knowledge of the party committing breach to know which is the case. In a rough sense, the likelihood that breach is immoral is higher the lower are damages in relation to the true expectancy. Hence, the chances are greatest that breach is immoral when damages are effectively zero – that is, when breach-caused losses are below the threshold that would make suit worthwhile. If damages surpass this threshold, but are significantly below the value of performance, the breaching party must obtain a relatively large benefit from breach, exceeding by a substantial margin the damages he would have to pay, for the breach to be moral.²⁷ If, though, expectation damages are near the value of performance, breach is quite likely to be moral, and the reason for breach does not need to be examined to infer that that is so.

²⁶ If, for example, a lawyer's hourly fee is \$250, and only twenty hours of his or her time is required to litigate, the legal costs of litigation would be \$5,000, implying that the expected gain from suit would have to exceed this amount for a threat to litigate to be credible. I abstract here from many complicating issues, such as the reputational concerns of the party in breach (making the threat of suit more likely) and the value of the breach victim's time (making the threat of suit less likely).

²⁷ Consider a contract to paint a house that is to be put up for sale and thus that will be likely to command a higher price if freshly painted. Assume too that the owner would find delay costly (say the owner has signed a purchase and sale agreement on another home, set a closing date, and must sell his present house soon in order to secure funds to close on the new home). In particular, assume that the owner's expected cost from a delay in painting would be \$25,000 due to having to sell his home when it is in real need of a paint job and that the painter has an appreciation of this fact. But suppose that the damages the owner would obtain for delay would be only \$3,000, owing to the speculative nature of the price the owner could have obtained had the painting been done. Then the painter would be led to breach if he could obtain more than \$3,000 more in profit for doing a different job, yet if the painter's profit from a breach is less than \$25,000, the breach would be immoral; only if his reason for breach makes more than a \$25,000 difference would the breach be moral.

4. The Morality of Breach: In The Opinion of Individuals

To this point, we have made use of the definition of breach that I assumed, the salient feature of which is that nonperformance in a contingency is morally permissible if the parties know that they would have agreed to nonperformance had they discussed the contingency. Is this definition of the morality of breach of contract consistent with the actual expression of moral opinions about contract breach? Or do individuals tend to disregard the question about hypothetical contractual terms and instead describe a breach as immoral because it is facially a violation of a promise? If it is made clear that breach would be accompanied by payment of compensatory damages, do attitudes toward breach change? To gain an understanding of these issues, I conducted a small-scale survey.²⁸

The number of respondents was 41.

The survey consisted of four questions, each of which asked about the morality of breach and could be answered as follows: (1) definitely unethical; (2) somewhat unethical; (3) neither ethical nor unethical; (4) somewhat ethical; (5) definitely ethical. I assigned a score of 1 to definitely unethical, a score of 2 to somewhat unethical, and so forth. Hence, the lower score, the less ethical a respondent felt breach would be.

The first question was designed to ascertain whether respondents believe that breach in general is unethical. It was as follows:

“Suppose that a Renovator has made a contract to do a kitchen renovation for a Homeowner. The Renovator then discovers that the job would cost him a lot more than he had anticipated because the price of kitchen equipment has risen

²⁸ The survey was designed jointly with Michael Simkovic.

sharply – so the Renovator would lose money on the job. Is it unethical for the Renovator to break his contract with the Homeowner?”

Note that the question does not mention whether damages would be paid. The average answer score was 2.41, meaning about midway between somewhat unethical and neither ethical nor unethical. Also, 38 of the 41 respondents found breach unethical or ethically neutral; only 3 of respondents answered that the breach would be somewhat ethical (none as definitely ethical).

The second question differed from the first in that respondents were told that, had the parties to the contract discussed the rise in prices, performance would not have been required:

“Suppose that a kitchen Renovator has made a contract with a Homeowner and finds that his costs have risen sharply due to an increase in the price of kitchen equipment. Suppose too that *the Renovator and the Homeowner did not discuss this unlikely possibility when they made their contract*. However, the Renovator knows what they *would have* agreed to if they had discussed this possible large cost increase in advance: *They would have agreed that the contract would be canceled if there was a large cost increase – the Renovator would be excused from the contract*. Under these assumptions, is it unethical for the Renovator to break his contract?”²⁹

The average answer score was 3.0, meaning ethically neutral. Also, 17 of the respondents found breach more ethical in this question than they had in the first question; none of the respondents found breach less ethical than in the first question.

²⁹ The emphasis in the question is as was presented to respondents. The same is true of questions 3 and 4 below.

In the third question, the assumption that the respondents were told to make was changed. They were told that, had the contracting parties discussed the contingency concerning the rise in prices, the parties would still have wanted performance:

“Suppose that a kitchen Renovator has made a contract with a Homeowner and finds that his costs have risen sharply due to an increase in the price of kitchen equipment. Suppose too that *the Renovator and the Homeowner did not discuss this unlikely possibility when they made their contract*. However, the Renovator knows what they *would have* agreed to if they had discussed this possible large cost increase in advance: *They would have agreed that the contract would remain in force despite a large price increase – the Renovator would still have to do the job.*”

The average answer score for this question was 1.56, which is to say, midway between definitely unethical and somewhat ethical. All but three of the respondents found breach to be somewhat or definitely unethical, and the other three considered it ethically neutral. Further, 31 of the respondents found breach to be less ethical than in the prior question where the hypothesis was that performance would not have been agreed to.

The fourth question was like the first, except that respondents were told that compensatory damages would be paid for breach:

“Suppose that a kitchen Renovator has made a contract with a Homeowner and finds that his costs have risen sharply due to an increase in the price of kitchen equipment. If the kitchen Renovator breaks his contract with the Homeowner, suppose that (as contract law says is required) *he compensates the Homeowner for his losses* – for delay, inconvenience, having to hire

another renovator, and so forth. Is it unethical for the Renovator to breach his contract?”

The average answer score for this question was 3.56, which is about midway between ethically neutral and somewhat ethical.

In summary, the individuals participating in the survey found the simple, unqualified fact of breach to be unethical on average (2.41 was the average for question 1). In other words, the felt reaction to the fact of breach is that it is an unethical act.³⁰ However, when individuals were *prompted* by being told what contracting parties would have agreed to had they discussed the particular contingency that arose, individuals tended to change their evaluation of the morality of breach, finding it better or worse in the expected way. When informed that if the problematic contingency had been discussed, the contracting parties would have said no duty to perform, individuals found breach ethically neutral (3.0 was the average for question 2). When apprised that if the problematic contingency had been discussed, the contracting parties would have said there was still a duty to perform, individuals found breach to be quite unethical (1.56 was the average score for question 3). And when told that breach would be accompanied by full damages payments, individuals again changed their opinion of breach, finding it to be somewhat ethical (3.56 was the average score for question 4).

5. The Morality of Breach: In the View of Commentators

³⁰ Note that this was so even though in the question put, the circumstance of the Renovator might have evoked some sympathy since the reason for breach was that costs rose (rather than that he decided to take another job).

As I noted at the outset, most sources on contract law and legal commentators who have addressed the issue of the morality of breach have considered it to have a generally unethical dimension, or at least have described it as having that character in the opinion of others. This includes the Restatement of Contracts, as I said, authors of hornbooks such as Farnsworth,³¹ and prominent contracts scholars such as Atiyah, Barnett, Craswell, Eisenberg, and Fried.³² Eisenberg's statement in a recent article is representative:³³ "The moral meaning of making a [contractual] promise is to commit yourself to take a given action in the future even if, when the action is due to be taken, all things considered you no longer wish to take it." That is, he believes that a contractual promise has moral valence, so that a breach might be immoral.

These writings do not devote attention to the main point developed here, that contracts frequently do not provide for the specific contingencies that lead to breach, so that one cannot truly say that an explicit promissory duty was violated when a breach occurred. It is evident that the view of commentators about the morality of breach is similar to the view of individuals in general, as represented by the survey results reviewed above. Namely, the breach of a contract is seen as what it appears to be, as a violation of a promise and thus as having a morally inappropriate aspect. Perhaps, though, commentators would alter their opinions if they considered the incomplete nature

³¹ In discussing economic analysis of remedies, FARNSWORTH, *supra* note 4, comments at 764 that it "leaves no place for notions of the sanctity of contract and the moral obligation to honor ones' promises."

³² P. S. ATIYAH, PROMISES, MORALS, AND LAW (1981), Randy E. Barnett, *A Consent Theory of Contract*, 86 COLUM. L. REV. 269 (1986), Richard Craswell, *Contract Law, Default Rules, and the Philosophy of Promising*, 88 MICH. L. REV. 489 (1989), Eisenberg, *supra* note 25, at 1011-1013, and FRIED, *supra* note 2, at 17.

³³ Eisenberg, *supra* note 25, at 1012.

of contracts and that breach tends to occur when an express contractual promise would have allowed nonperformance, just as the surveyed individuals changed their opinions when they were asked about the morality of breach against this background.

In any event, let me now comment on the views about breach of Holmes and of law and economics scholars concerned with “efficient breach.” Holmes’ ideas about breach are often characterized as adumbrating those of economic analysts, because he emphasized the point that when a person without morals, a “bad man,” is contemplating breach, his decision would be determined primarily by the prospect of having to pay damages for that act.³⁴ Yet Holmes hardly believed that individuals generally act without morals, and, although he did not squarely address whether breach might be immoral, we have no reason to think that he would not consider it so.³⁵ That Holmes asserted that the law gives individuals an option to perform or to pay damages does not imply that he thought breach was an ethically neutral act; he sought to describe the law,³⁶ not to evaluate behavior.

³⁴ Holmes, *supra* note 5, at 459-460, 462. In particular, at 459, he says, “If you want to know the law and nothing else, you must look at it as a bad man, who cares only for the material consequences which such knowledge enables him to predict, not as a good one, who finds his reasons for conduct, whether inside the law or outside of it, in the vaguer sanctions of conscience.”

³⁵ It is clear that Holmes discusses the motivations of the bad man to draw out the direct effect of legal sanctions on behavior, not because he thinks that good men fail to exist, as the ending phrase in the quotation of the last note shows. He even states that he does not want a wrong meaning to be taken from his address: “I take it for granted that no hearer of mine will misinterpret what I have to say as the language of cynicism. The law is the witness and external deposit of our moral life....The practice of it... tends to make good citizens and good men.” Holmes, *supra* note 5 at 459.

³⁶ Although Holmes’ characterization of the law may seem basically sound, some have criticized it. An interesting example is Sir Frederick Pollock, who suggested in a letter to Holmes that the law was not indifferent between breach and performance, for why else would it sanction an individual for inducing another to commit breach? Pollock also pointed out that in the German Civil Code specific performance, not damages, is generally the favored remedy for breach (as is true today – *see infra* note 44). *See* letter of Sept. 17, 1897, in 1 HOLMES-POLLOCK LETTERS 79-80 (Howe ed. 1941).

Many economically-oriented writers on breach of contract have focused on what is called the theory of efficient breach.³⁷ An efficient breach is a breach that fosters a utilitarian, aggregate measure of social welfare, because the breach would lead to avoidance of unduly costly performance or would allow sale of a good to a third party willing to pay more than the promisee. Writers on efficient breach have observed that breach will tend to be efficient under the expectation measure (since a party contemplating breach will commit it if and only if his benefit would exceed the value of performance to the other side). These writers often further recommend that, as a matter of actual policy, breach should be encouraged when it is efficient, even though the role of morality is not analyzed by them.³⁸

How does the theory of efficient breach relate to what has been said in this article? Two points should be made. First, because damages are likely to be less than the true expectation, as stressed here, it is not clear that breach will in fact tend to be efficient. Thus, it is not evident that contracting parties should be encouraged to commit

³⁷ See originally Robert L. Birmingham, *Breach of Contract, Damage Measures, and Economic Efficiency*, 24 RUTGERS L. REV. 273, 284-285 (1970). The theory has been developed and expounded by, among others, Charles J. Goetz and Robert E. Scott in *Liquidated Damages, Penalties, and the Just Compensation Principle: Some Notes On an Enforcement Model and a Theory of Efficient Breach*, 77 COLUM. L. REV. 554 (1977) and by Richard Posner in the various editions of his textbook, the latest being RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 118-126 (6th ed., 2003). The theory is now widely recognized (but usually at most partially endorsed) within the legal community, as is evidenced, for example, by its mention in RESTATEMENT (SECOND) OF CONTRACTS ch. 16, Introductory Note (1981). One should note, however, that most economic analysis of contracts published in economics journals asks not about the social desirability of damage measures but rather about the desirability of damage measures for the contracting parties themselves. See, for example, my own article, Steven Shavell, *supra* note 8, and much of the literature discussed in PATRICK BOLTON AND MATHIAS DEWATRIPONT, *CONTRACT THEORY* chs. 11-13 (2005).

³⁸ See, e.g., Robert L. Birmingham, *supra* note 37, who states at 284 “Repudiation of obligations should be encouraged where the promisor is able to profit from his default after placing the promisee in as good a position as he would have occupied had performance been rendered,” and RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 57 (1st ed., 1972), who writes, “If his profit from breach would also exceed the expected profit to the other party from completion of the contract, and if damages are limited to loss of expected profit, there will be an incentive to commit a breach. There should be.”

breach when they are willing to pay damages, according to the social welfare criterion of efficient breach theory.

Second, and this point aside, when efficient breach occurs, it coincides with the terms of completely detailed contractual promises and thus should not be seen as immoral (given my definition of moral behavior), as I have emphasized. This conclusion has nothing directly to do with promoting a utilitarian measure of social welfare, though, and thus might be regarded a kind of coincidence. As the reader recalls, the conclusion flows from the logic explaining the nature of the contract that the parties themselves would want (they might not give two straws for social welfare).³⁹ I will remark in the next section, however, on the efficient breach theorists' normative recommendations about breach behavior, assuming that the damage measure is equal to the expectation. My analysis so far does not allow me to do this, for I have not discussed social welfare.⁴⁰

6. Conclusion

In conclusion, I would like to comment on the implications of this article for the general normative thrust of the writing of traditional commentators concerned with the morality of breach and also of the efficient breach theorists. I will assume that the social objective is to promote an individualistic measure of social welfare, one based on individuals' utilities. In particular, then, our moral feelings will have a direct effect on

³⁹ The argument in section 2 leading to the conclusion that performance would be specified in a contingency if and only if the cost of performance is less than the value of performance rested on how much the buyer would pay and on how much the seller would demand for a provision requiring performance.

⁴⁰ Although the subject of this article has been the morality of breach, which concerns normative statements (such as "you ought not to commit breach because that would be wrong"), and I have defined and considered a definition of moral behavior, I have not said how the morality of behavior enters into social welfare, which is by definition the criterion to be employed for social evaluation.

social welfare because they are themselves components of individuals' utilities⁴¹ and they will also exert an indirect influence on social welfare because they provide incentives toward socially desirable behavior.

An important normative aspect of many commentators' writing on breach is their moral advice. Their writing often suggests that individuals ought to feel a general ethical duty to obey contracts, that is, a desire to obey contracts above and beyond that due only to having to pay damages for breach. If we could shape individuals' moral feelings, would we want individuals to put a thumb on the scale in favor of contract performance? The answer suggested by the analysis here is that we would not want individuals *always* to do so, but we would want that done *when necessary*, to make up for the slack due to inadequate damages. In other words, according to a perfectly calibrated and flexible moral system, the moral sentiments would come into play if and only if they are needed to correct the too-great incentive of a "bad man" to commit breach, when the personal benefit from breach would exceed damages but not the true value of the expectancy. This ideal moral system, note, is consistent with the spirit of traditional commentators' advice, and is inconsistent with the spirit of efficient breach theory, in that morality has a useful role to play. Still, the ideal moral system is inconsistent with the advice of traditional commentators, and is consonant with efficient breach theory, in the sense that moral forces would not discourage breach when breach would be efficient.

The actual moral system, however, is not as flexible as the ideal one. The moral impulses probably cannot be freely tailored to turn on for this kind of contract breach and

⁴¹ For example, an individual's utility will decrease if he or she experiences guilt for violating a moral duty.

to turn off for that one.⁴² If so, the price that we have to pay for our having the fairly general and socially valuable moral sense that breaking contracts is undesirable is that that view will be applied by us in some circumstances where breach is efficient. This in turn will mean that some desirable breaches will not be committed owing to the moral discomfort that that would entail and that other desirable breaches will be committed despite the discomfort that that will entail (this discomfort will, unfortunately, lower social welfare). Further, the implications for moral advice about breach become complicated, for when giving moral advice, we have to consider the degree to which the advice will be understood as special to the circumstances of the breach, or as having a more general effect, and thus entailing the implicit disadvantages just mentioned.

Closely related to their moral advice to contracting parties is commentators' advice about the law, namely, that the moral desirability of satisfying contractual promises argues for legal policy that fosters performance. Some have said, for example, that the scope of specific performance should be expanded for moral reasons among others.⁴³ A striking instance of such thinking is that underlying the approach of the German Civil Code to contract performance, according to which the general remedy for breach is supposed to be specific performance.⁴⁴ Of course, the theme of this article has

⁴² Many have argued thus, as well as that moral views ought not to be too qualified (lest they be subverted by opportunistic interpretation). See, e.g., R. M. Hare, *Principles*, 73 PROC. ARISTOTELIAN SOC'Y 1, 10 (1973); RICHARD B. BRANDT, A THEORY OF THE GOOD AND THE RIGHT chs. 9, 11 (1979), and Jonathan Baron, *Normative, Descriptive and Prescriptive Responses*, 17 BEHAV. & BRAIN SCI. 32, 32-33 (1994).

⁴³ See, e.g., Randy E. Barnett, *Contract Remedies and Inalienable Rights*, 4 SOC. PHIL. & POL. 179 (1986).

⁴⁴ See, for example, KONRAD ZWEIGERT & HEIN KÖTZ, INTRODUCTION TO COMPARATIVE LAW 472-474, 483-485 (Tony Weir, trans., 3rd rev. ed., 1998). As this reference makes clear, however, specific performance is more aspirational than real in Germany, since in fact victims of breach can elect damages and usually do so. In a sense, the fact that in reality damages are usually

been that contracts are to an important extent incomplete promises and thus on reflection that the morality of promise-keeping does not imply that performance should always occur. Yet I have noted that damages seem to be systematically less than the true expectation measure. The import of these observations for legal policy is not clear. Damages are inadequate because it is time-consuming and expensive for the legal system to resolve what would often be contentious proceedings about subjective elements of loss from breach. It may be that our legal system works better avoiding the costs of ascertaining these problematic elements of loss, relying on moral forces, such as they are, to fill the gap in inducing appropriate performance. In any event, the belief that there is a clear and overarching moral reason to alter contract law to enhance the keeping of contracts appears to me to be the product of an oversimple view of the moral sentiments and of a related failure to take into account the importance of the incompleteness of contracts.

employed yet the official view is that specific performance is best shows how strongly felt the official view must be.