University of Chicago Law School

Chicago Unbound

Journal Articles

Faculty Scholarship

1996

Bad Faith Breach of Contract by First-Party Insurers

Alan O. Sykes

Follow this and additional works at: https://chicagounbound.uchicago.edu/journal_articles

Part of the [Law Commons](#)

Recommended Citation

Alan O. Sykes, "Bad Faith Breach of Contract by First-Party Insurers," 25 Journal of Legal Studies 405 (1996).

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

“BAD FAITH” BREACH OF CONTRACT BY FIRST-PARTY INSURERS

ALAN O. SYKES*

ABSTRACT

Insurers may at times exploit the delay inherent in the civil litigation process to induce needy insureds to settle for less than the amount that the contract promises. The prospect of extracontractual remedies for such “bad faith” at the end of the litigation process can make these tactics unprofitable and thus serve a potentially valuable function. But the remedy may be worse than the problem, as the courts seem to find bad faith on the part of insurers who have genuine and reasonable disputes with their policyholders over the terms of the policy or over factual issues essential to the insured’s right to recover. The ability of the courts to identify opportunistic behavior accurately is thus in doubt, and the possibility arises that bad faith doctrine in first-party cases does little to police misconduct while doing much to cause uneconomic increases in the premiums that policyholders must pay.

CONTRACTS usually include an implied covenant of good faith and fair dealing.¹ A failure by a party to act in good faith may thus lead to a breach of contract action regardless of the express terms of the contract. A failure to act in good faith may also constitute a tort.²

In many commercial settings, actions based on breach of the obligation to act in good faith are exceedingly rare. But in the insurance industry, the importance of the “bad faith” action has become profound. Depending on the jurisdiction, the type of insurance and the alleged failing of the insurer, an insurer who acts in “bad faith” may incur liability not only for damages owing under the policy but also for uncovered economic losses of the insured, the insured’s emotional distress damages and attor-

* Professor of Law, University of Chicago, The Law School. I have received useful comments from David Friedman, Louis Kaplow, and Steven Shavell, from the participants in law and economics workshops at Harvard University and the University of Michigan, and from a thoughtful referee. I thank the Scaife and the Bradley Foundations for financial support.

¹ See E. Allan Farnsworth, *Contracts* § 7.16 (1990); *Restatement (Second) of Contracts* § 205; U.C.C. § 1-203.

² See Dan Dobbs, Robert Keeton, David Owen, & W. Page Keeton, *Prosser and Keeton on Torts* § 92 (1988).

[*Journal of Legal Studies*, vol. XXV (June 1996)]

© 1996 by The University of Chicago. All rights reserved. 0047-2530/96/2502-0005\$01.50

ney's fees, substantial punitive damages, and various statutory penalties. This expansion of the remedies for breach of insurance contracts has spawned a wealth of bad faith litigation over the past few decades to the point that it has become an important subspecialty of insurance practice.³

The earliest allegations of bad faith by insurers arose under third-party liability insurance policies, usually after the insurer had rejected a settlement offer within the policy limits and the insured thereafter incurred liability in excess of those limits.⁴ More recently, actions seeking recovery for bad faith under first-party medical, disability, casualty, and life policies have become commonplace. In these cases, the allegation is that the insurer has refused without sufficient justification to pay benefits owing under the policy and forced the insured to litigate to obtain them. It is said that a majority of American jurisdictions now allow a tort-based bad faith action in these first-party cases.⁵ Such actions are the focus of this article.

The commentary on the wisdom of expanded remedies for breach of first-party insurance contracts is mixed. Some commentators argue that these contracts are no different from other contracts and that a breach should entitle the insured to nothing beyond the remedies available to other contract plaintiffs.⁶ Other commentators favor additional remedies in first-party cases on the premise that insurers have unusual capacity to take advantage of an insured who has suffered a substantial loss.⁷ Yet those who favor special remedies differ considerably in their views as to what the appropriate remedy should be and the appropriate circumstances in which to allow it.⁸

This article will not settle the debate, for the issues are complex and ultimately turn on controversial empirical questions. What can be said is

³ See, generally, Stephen Ashley, *Bad Faith Actions: Liability and Damages* (1994); William Shernoff et al., *Insurance Bad Faith Litigation* (1984); Symposium on the Law of Bad Faith in Contract and Insurance, 72 Tex. L. Rev. 1203 (1994).

⁴ The economic commentary on third-party cases includes Kenneth Abraham, *Distributing Risk* 188-203 (1986); Michael Meurer, *The Gains from Faith in an Unfaithful Agent: Settlement Conflicts between Defendants and Liability Insurers*, 8 J. L. Econ. & Org. 502 (1992); Kent Syverud, *The Duty to Settle*, 76 Va. L. Rev. 1113 (1990); Alan Sykes, *Bad Faith Refusal to Settle by Liability Insurers: Some Implications of the Judgment Proof Problem*, 23 J. Legal Stud. 77 (1994); Alan Sykes, *Judicial Limitations on the Discretion of Liability Insurers to Settle or Litigate: An Economic Critique*, 72 Tex. L. Rev. 1345 (1994).

⁵ See Ashley, *supra* note 3, § 5:01.

⁶ *Id.* § 2:14; William Powers, Jr., *Border Wars*, 72 Tex. L. Rev. 1209 (1994).

⁷ See Abraham, *supra* note 4, at 183-88.

⁸ See *id.* (suggesting the possible virtue of statutory reform to control jury discretion); Note, *Attorney's Fee Recovery in Bad Faith Cases: New Directions for Change*, 57 S. Cal. L. Rev. 503 (1984) (favoring attorney's fees); Note, *The Breach of First Party Insurance Contracts*, 56 S. Cal. L. Rev. 1345 (1983) (favoring emotional distress damages).

that the case for substantial extracontractual damages in bad faith cases, particularly large punitive awards, is an uneasy one. A theoretical argument for such damages can be made under proper assumptions, but it appears that the courts have extended bad faith remedies to circumstances in which the case for them cannot be made and that reining in the remedy more properly may be exceedingly difficult. Thus, the question arises whether the cure is worse than the disease.

To summarize the key points that follow, some theoretical justification exists for punitive damages when the insurer's refusal to pay a claim is utterly baseless and strategic—cases defined here as "intentional breach." Perhaps the strongest argument for stiff extracontractual penalties in these cases is that insurers might otherwise exploit the prospect of delay in litigation to extract favorable settlements from insureds with high implicit discount rates. But the importance of this possibility should not be exaggerated. Impatient insureds may have some viable counterstrategies against an opportunistic insurer, such as assignment and borrowing. Moreover, additional damage remedies for the wrongful denial of meritorious claims could be provided by contract, and it is important to consider what inference should be drawn from their failure to emerge in competitive insurance markets. Perhaps the proper inference is that error costs in the administration of such measures would swamp any benefits. And even if the absence of contractual remedies reflects market failure, it is possible that measures to reduce delay in litigation, such as mandatory arbitration, would better serve the parties joint interests than measures to increase the damages award.

The analysis is more complicated for cases that do not involve intentional breach. Not surprisingly, the case for extracontractual damages is weaker here, though it cannot be dismissed altogether as a theoretical matter. On the one hand, a complete contingent contract between an insurer and an insured would no doubt at times require the insurer to make prompt payment of a claim despite some factual or legal uncertainty as to its validity. And in the absence of complete contracts, there may be room for opportunism by the insurer if the only penalty for the denial of a claim is the possibility that it might have to be paid later. On the other hand, it may be impossible as a practical matter for courts to distinguish opportunistic behavior by an insurer in the face of uncertainty about the merits of a claim from behavior that has sound justification.

An examination of some of the leading cases on first-party bad faith will indeed suggest that courts award substantial damages for "bad faith" in circumstances where opportunism is by no means evident. Thus, although I do not mean to suggest that insurers never take advantage of their insureds in a manner that might properly be called "bad faith," the

hard question is whether the error costs associated with judicial attempts to police that behavior may generate more problems than they solve.

Section I sketches the evolution of remedies in first-party cases. Section II develops some theoretical points about intentional breach cases, while Section III considers cases where the dispute over the validity of the claim is bona fide. Section IV then takes a critical look at a number of prominent cases.

I. REMEDIES FOR WRONGFUL DENIAL OF FIRST-PARTY CLAIMS

States vary significantly in the extent to which they allow special remedies for “bad faith” breach of first-party insurance contracts and in the circumstances that they define as “bad faith.” At one end of the continuum lie states that refuse to depart from traditional principles of contract law, restricting the insured to the ordinary remedies for breach irrespective of the insurer’s basis for denying the claim. At the other end lie states that allow the insured to recover substantial tort damages for a negligent or “unreasonable” denial of a claim, including damages for emotional distress and often punitive damages. In between are the states that purport to allow tort remedies only on a showing that the insurer denied the claim without any basis.

Thorough surveys of the doctrine in various jurisdictions may be found elsewhere.⁹ The objective here is simply to provide the reader with a flavor for the range of possibilities that emerge. It is convenient to divide the discussion between traditional contract remedies for breach of the insurer’s promise to pay benefits and the newer tort cause of action for “bad faith” refusal to pay benefits.

A. *Traditional Remedies for Breach of First-Party Insurance Contracts*

A first-party insurance contract is at its heart a promise to pay a sum or sums of money to the insured (or heirs of the insured) when covered contingencies materialize and certain conditions have been met. The minimum remedy for breach of such a promise, available in all jurisdictions, is an action to recover the amounts owing under the terms of the insurance policy. At one time courts generally held that no further remedy was available.

Yet it has long been said that the usual objective of contract damages is to place the promisee in the same position that the promisee would

⁹ See, for example, Ashley, *supra* note 3, § 2:15, and Ch. 5; Robert Keeton & Alan Widiss, *Insurance Law* § 7.9 (1988).

have enjoyed had the contract been performed.¹⁰ An aggrieved insured can point to a number of ways in which an award of money damages at the conclusion of litigation, equal to what the insurer should have paid in the first place, fails to achieve this goal. Such an award deprives the insured of the opportunity to earn an ordinary return (interest) on the funds owing during the pendency of litigation. It also fails to compensate the insured for the costs of litigation. The insured may suffer yet additional economic losses where, for example, the insurer's breach delays repairs necessary to the operation of a business, delays the purchase of medical services necessary for the insured to return to work, and so on. Finally, an insured who is uncertain of ultimate payment or faced with a lengthy delay in payment may suffer emotional anxiety. A genuine effort to compensate the insured fully for the breach would require an award that covered all of these losses.

Most likely to be recoverable in a contract action is prejudgment interest. Although courts at one time regularly denied interest in breach of contract (and other civil) actions, it is now commonly awarded, usually by statute, where the amount in controversy is readily ascertainable by the party who owes it and thus may be deemed "liquidated."¹¹ Insurance claims typically meet this test.¹²

Attorney's fees, by contrast, are still routinely denied in ordinary breach of contract actions. A number of jurisdictions make an exception to this "American rule," however, in cases involving breach of insurance contracts. In general, these exceptions also arise by statute.¹³ Some apply the rule that insureds who must resort to the legal process to collect insurance benefits are entitled to fees as a matter of course if they prevail in their claims,¹⁴ while others award fees if the amount claimed by the insured at the outset of litigation is equal to or sufficiently close to the amount actually recovered.¹⁵ In addition, Rule 11 of the Federal Rules of Civil Procedure and certain similar state court rules allow a trial judge to impose sanctions, including fee shifting, when one-party asserts frivolous claims or defenses.

Additional economic losses caused by breach are potentially recover-

¹⁰ For example, Farnsworth, *supra* note 1, at 840.

¹¹ Comment, Prejudgment Interest: Survey and Suggestion, 77 Nw. U. L. Rev. 192, 199 (1982); Ashley, *supra* note 3, § 8:09.

¹² See Ashley, *supra* note 3, § 1:01.

¹³ *Id.* § 8:10.

¹⁴ See, for example, *Polito v. Continental Casualty Co.*, 689 F.2d 457 (3d Cir. 1982); Ashley, *supra* note 3, § 8:09.

¹⁵ See, for example, *First Marine Ins. Co. v. Booth*, 317 Ark. 91, 876 S.W.2d 255 (1994).

able as "consequential damages" under usual principles of contract law, but only if they were "foreseeable" by the breaching party under the rule of *Hadley v. Baxendale*. Depending on the facts, therefore, insurers may escape liability on the grounds that any economic losses were beyond their contemplation at the time of contract formation.¹⁶

Finally, damages for mental anguish are usually not recoverable in contract actions, unless a central purpose of the contract was to provide for emotional needs.¹⁷ Some commentators have suggested that insurance contracts fall within this exception, but a number of jurisdictions have rejected that assertion or otherwise denied damages for mental anguish on grounds of unforeseeability.¹⁸

B. The "Bad Faith" Cause of Action

Even before the "bad faith" action was conceived, insurance plaintiffs could occasionally invoke alternative theories and recover more generous damages. An insurer who sells a policy without any intention of honoring its promises commits a fraud;¹⁹ an insurer who deliberately causes an insured severe emotional distress to extract a favorable claims settlement, or who recklessly causes severe distress, may have engaged in the "extreme and outrageous" behavior necessary to support an action for intentional infliction of emotional distress.²⁰ Either theory, which converts the insurer's breach of contract into an intentional tort, will support the award of a full panoply of economic and noneconomic compensatory damages and punitive damages.

Both fraud and intentional infliction of emotional distress are difficult to establish, however, in most cases. Fraud requires proof of a false statement of present intention to honor the promise or some other deliberately false statement of fact that induces detrimental reliance. Intentional infliction of emotional distress requires extreme behavior calculated or likely to result in severe emotional distress, and either or both is often absent.

The implied covenant of good faith and fair dealing in contract is the starting point for the creation of further remedies. An insurer who denies

¹⁶ See 16A John Appleman, *Insurance Law and Practice*, § 8878.15 (rev. ed. 1981).

¹⁷ See Farnsworth, *supra* note 1, at 934.

¹⁸ See, for example, *Kewin v. Mass. Mutual Life Ins. Co.*, 409 Mich. 401, 295 N.W.2d 50 (1980).

¹⁹ See Prosser and Keeton on Torts, *supra* note 2, § 109.

²⁰ See Restatement (Second) of Torts § 46 (1966); *Fletcher v. Western National Life Ins. Co.*, 10 Cal. App. 3d 376, 89 Cal. Rptr. 78 (1970).

a claim without reasonable foundation may be said to have violated this covenant and, thus, to have committed a second breach of contract, along with breach of the express promise to pay benefits. But an additional contract theory of recovery, by itself, does little for the plaintiff. Unless the damages recoverable are affected, the second contract theory is redundant.

The important doctrinal move, therefore, is to hold that breach of the implied covenant of good faith is a tort (or, in the alternative, to hold that usual limitations on contract damages do not apply to a breach of this covenant). In first-party cases, the California Supreme Court took the lead with *Gruenberg v. Aetna Insurance Co.*²¹ The court had previously held that an insurer was subject to liability in tort for failure to accept reasonable settlement offers under liability policies. "[I]n the case before us we consider the duty of an insurer to act in good faith and fairly in handling the claim of an insured, namely a duty not to withhold unreasonably payments due under a policy. . . . Where [the insurer] fails to deal fairly and in good faith with its insured by refusing, without proper cause, to compensate its insured for a loss covered by the policy, such conduct may give rise to a cause of action in tort for breach of an implied covenant of good faith and fair dealing."²² The reference in *Gruenberg* to "unreasonably" withholding payments has been interpreted as a kind of negligence standard, supporting liability where the insurer may have honestly believed that the denial of the claim was justifiable yet the insurer was unreasonable in that belief.²³

The majority of jurisdictions that have accepted the tort cause of action, however, have preferred a purportedly narrower conception of "bad faith."²⁴ Many follow the Wisconsin Supreme Court in *Anderson v. Continental Insurance Co.*²⁵

To show a claim for bad faith, a plaintiff must show the absence of a reasonable basis for denying benefits of the policy and the defendant's knowledge or reckless disregard of the lack of a reasonable basis for denying the claim. It is apparent, then, that the tort of bad faith is an intentional one. . . . The tort of bad faith can be alleged only if the facts pleaded would, on the basis of an objective standard, show the absence of a reasonable basis for denying the claim, i.e., would a reasonable insurer under the circumstances have denied or delayed payment of

²¹ 9 Cal. 3d 566, 108 Cal. Rptr. 480, 510 P.2d 1032 (1973).

²² 9 Cal. 3d 573-74, 108 Cal. Rptr. 484-85, 510 P.2d 1036-37.

²³ See Ashley, *supra* note 3, § 5:02.

²⁴ *Id.*

²⁵ 85 Wis. 2d 675, 271 N.W.2d 368 (1978).

the claim under the facts and circumstances. . . . Under these tests of the tort of bad faith, an insurance company, however, may challenge claims which are fairly debatable and will be found liable only where it intentionally denied (or failed to process or pay) a claim without a reasonable basis.²⁶

Despite the efforts of the *Anderson* court to define bad faith breach as an "intentional" tort, even a superficial reading of the opinion reveals elements of a negligence standard. The insurer is held to an "objective standard" in deciding whether it had a "reasonable basis for denying benefits." It may only refuse to pay claims that are "fairly debatable" by this standard. True, the defendant must have "knowledge" that it lacks a reasonable basis for denying the claim or act in "reckless disregard" of its lack of a reasonable basis, but *Anderson* stops well short of requiring proof that the insurer believed the claim to be meritorious but denied it anyway. Instead, it is often said that an insurer subject to the *Anderson* standard must satisfy the court or the jury that it took a position with respect to the claim that a reasonable insurer *could* take.²⁷ By contrast, the *Gruenberg* standard may be interpreted to require the insurer to show that its position was one that a reasonable insurer *would* take. Under either standard, the court and perhaps the jury must make some judgment about the objective "reasonableness" of the insurer's actions.²⁸

Once "bad faith" is established under the applicable standard, the full array of tort remedies generally becomes available when the action proceeds at common law.²⁹ Not only may the plaintiff recover for all economic harms (save attorney's fees in the absence of statute) but also for noneconomic harms. Because "bad faith" can often be characterized as "reckless," "malicious," "wanton," and the like, many plaintiffs will be able to collect punitive damages.³⁰

II. TORT REMEDIES IN INTENTIONAL BREACH CASES

An insurer commits an "intentional breach" when it refuses to pay a claim knowing with subjective certainty that the claim is meritorious and

²⁶ 85 Wis. 2d 691-93, 271 N.W.2d 376-77.

²⁷ Ashley, *supra* note 3, § 5:04.

²⁸ This blending of concepts from intentional torts and negligence is common even in jurisdictions that do not purport to follow *Anderson*. For example, the statutory remedy for "bad faith" in Illinois allows the court to award attorney's fees plus punitive damages of as much as \$25,000 when the denial of a claim is "vexatious and unreasonable." 3 IL. Rev. Stat. Ch. 73, Para. 767 (1991).

²⁹ Some states rein in the remedy by statute. See, for example, *id.*

³⁰ See, for example, Restatement (Second) of Torts § 908 (punitive damages are appropriate for "outrageous" conduct, "evil motive," or "reckless indifference to the rights of others").

that litigation will result in a judgment for the insured for the full amount claimed (putting aside the possibility of judicial error). It is conceivable that insurers may occasionally (but only occasionally, given reputational concerns) be tempted to deny claims under these circumstances. This section considers the possible incentives for such opportunism and their implications for the appropriate remedy.

A. The Ex Post Gains from Denying a Valid Claim

Consider first the question whether insurers can gain from the denial of a valid claim, measuring the gains from the time that the claim is filed. Insurers will not be tempted to deny valid claims unless they (or their agents) can somehow profit from it. Thus, if the denial of a valid claim simply results in litigation that imposes additional costs on the insurer and invariably results in a judgment that the claim must be paid in full anyway, no temptation to cheat on the contract will arise.³¹

Of course, delay in the payment of a fixed nominal sum will have some time value. If the reduction in the present value of the payment exceeds the insurer's costs of litigation, forcing the insured to litigate may appear profitable. But as noted in Section I above, prejudgment interest is now common in first-party insurance litigation. Unless the insurer's discount rate is higher than the applicable rate for computing prejudgment interest, therefore, delay has no value for this purpose. I assume for the remainder of the discussion that prejudgment interest will be computed at a market rate equal to the insurer's discount rate.

Any gains from the denial of unquestionably valid claims must then arise because (a) the insured will drop the matter and decline to file suit, or (b) the insured will file suit (or threaten to file) but settle for less than the present value of the amount that the insured can eventually collect.

The first possibility might arise because the insured's litigation costs exceed the expected award.³² If a claim is indisputably valid, however, litigation costs should not be terribly high, and insureds with limited resources should not have great difficulty in persuading a contingency fee attorney to take the case if the stakes are significant. Thus, the possibility that the insured will decline to file suit seems important mainly with respect to small claims. Yet, market discipline may serve reasonably well

³¹ In this discussion, I put to one side the possibility that insurers will somehow fail to align their agents' interests with their own.

³² Even then, the insured might threaten to file suit or take some initial steps toward filing an action that might persuade the insurer to pay the claim. The literature on negative expected value suits is pertinent here. See, for example, Lucian Bebchuk, *Suits That Are Known to Be Made Solely to Extract a Settlement Offer* (Harvard Univ., photocopy, 1993).

to discourage insurers from refusing to pay minor claims. The gains to the insurer from avoiding payment are small (because the claim is small), yet the reputational penalties discussed in the next section will still attach. Further, minor claims tend to arise during an ongoing insurance relationship, and there is a greater likelihood that the damage to the insurer's long-term relationship with the insured will outweigh any short-term savings. Finally, even if market penalties are not enough, it is clear that fee-shifting rules, now found in a number of state insurance statutes and implicitly in general procedural rules that authorize sanctions for frivolous claims and defenses (such as Federal Rule 11), can eliminate the problem. Hence, it does not appear that the wide panoply of tort remedies available in modern bad faith litigation can be justified by cases where the claim would otherwise prove smaller than the insured's litigation costs.³³

The possibility that the insured may file suit but settle for less than the present value of the expected award thus seems the more interesting one, yet it may seem peculiar at first to suggest that an insured might behave in that fashion. To be sure, settlement will save the insured some litigation costs absent a fee-shifting rule, and a hard-bargaining insurer might be able to extract most or all of the savings. Yet, for an indisputably valid claim, the insured's costs of trial seem modest, and indeed the case can probably be resolved on summary judgment. Further, it is not clear why an insurer would expect to extract most or all of the surplus from settlement even when it is sizable.

What this analysis neglects, however, is the possibility that the insurer and the insured differ in their assessment of the "present value" of the judgment. While it is plausible to assume that the insurer's discount rate is approximately equal to the market rate of interest and to the rate used for computing prejudgment interest, the insured who has suffered a substantial loss may have an urgent need for funds. The bad faith cases are replete with references to insureds who lose their homes because a claim is denied, who are unable to secure needed medical care, even to

³³ The possibility that some individuals are nonlitigious and trepidatious about contacting an attorney in a small case, or that they are poorly informed about the possibility of fee-shifting or Rule 11 sanctions and thus do not bother to explore the possibility of suit, does not in my view justify greater remedies. Both concerns, if taken seriously, would argue for supracompensatory remedies in any type of litigation involving modest stakes and seem unconnected to the problem of misconduct by insurers in particular. Thus, they do not seem to justify special sanctions here, and the fact that the courts have chosen not to inflate the remedies available in small stakes cases generally suggests an adverse inference about their wisdom.

a disabled fellow whose "wheelchair was repossessed."³⁴ Individuals in dire financial straits might well agree to relinquish their contractual rights for an amount well below the present value of the amount to which they are entitled at the conclusion of litigation, discounted back to the present at a market rate.

Formal bargaining theory, for what it is worth, lends support to this intuition. Both the Nash bargaining model³⁵ and the alternating offers bargaining model of Rubinstein³⁶ suggest that, when an insured (or an

³⁴ See *Silberg v. California Life Ins. Co.*, 11 Cal. 3d 452, 113 Cal. Rptr. 711, 521 P.2d 1103 (1974).

³⁵ For purposes of this example, let the insurer's discount rate equal the rate at which prejudgment interest will be computed, call it r . The insured is entitled to receive an amount D under the policy, and the award at the conclusion of litigation will equal $(1 + r)D$. Because the insured's subjective discount rate exceeds r , however, $\$(1 + r)$ at the conclusion of the period is worth less than $\$1$ to the insured at the beginning—let the value of $\$(1 + r)$ discounted to the beginning of the period at the insured's subjective rate be $\$ \delta < \1 . Under these assumptions, the present value of the insurer's payment to the insured at the conclusion of litigation is equal to D from the insurer's perspective and δD from the insured's perspective.

Without more assumptions, we can predict only that the settlement amount will fall between D and δD —the insurer will not pay more than D , and the insured will not accept less than δD . The Nash cooperative bargaining model pins down the solution within this range by assuming that the parties will choose S in accordance with certain plausible axioms. See Eric Rasmusen, *Games and Information* 229–31 (1989). Let $u(\cdot)$ denote the insured's utility of wealth, increasing and concave, and assume the insurer is risk neutral so that its utility of wealth can be set equal to its wealth. Let w denote the insured's wealth after the occurrence of the loss in question. The Nash solution will maximize $[u(w + S) - u(w + \delta D)] \times (-S + D)$, subject to the constraint that $\delta D \leq S \leq D$. See *id.* Differentiation with respect to S and some rearranging yields the first-order condition $u'(w + S) = u'(w + \delta D)(D - S) + u'(w + \delta D)$. On inspection, this cannot hold at either corner solution, so we know that $\delta D < S < D$. Further, differentiating with respect to δ establishes that $\partial S / \partial \delta = u'(w + \delta D)D / [2u'(w + S) - u''(w + S)(D - S)] > 0$. That is, a fall in δ (equivalent to an increase in the insured's discount rate) reduces the settlement with the insurer. Further, as $\delta \rightarrow 1$ and the insured's discount rate approaches the market rate, the constraint ensures that $S \rightarrow D$.

³⁶ See Martin Osborne & Ariel Rubinstein, *Bargaining and Markets* (1990), Ch. 3.

Drawing on the Rubinstein framework, assume that litigation will conclude at the end of one "period," the length of which is arbitrary. For simplicity, ignore litigation costs, which would complicate matters but would not alter the basic point. Let r be the continuous time discount rate for the insurer and interest rate for the computation of prejudgment interest, and let $r + \rho$ be the insured's continuous time discount rate. Damages awarded at the conclusion of litigation will equal $e^r D$, with a discounted value to the insured at the beginning of the period equal to $e^{-(r+\rho)} e^r D = e^{-\rho} D$. The insurer is indifferent between paying D at the beginning of the period or $e^r D$ at the end.

Rubinstein imagines that the parties know that they will play an alternating offers game between the commencement of bargaining and the time that a deal is struck (or the game ends)—here, the time between the filing of a claim and the conclusion of litigation. If we assume that the number of offers is even, and that the time between offers is always the same, then the unique equilibrium settlement, S , is $S = e^{-\rho/2} D$. This solution is independent of the number of offers that we allow to be exchanged or the order of play, as long as each

heir) has an urgent need for funds due to the occurrence of a substantial loss, the insured may well be willing to surrender a future right for an amount considerably below its present value computed at market rates. Insurers aware of the situation can exploit it to obtain a more favorable settlement of the claim, and the extent to which they can do so will depend importantly on the amount of time that the insured must wait to collect a judgment at the conclusion of litigation as well as the insured's subjective sense of urgency.

Before proceeding, note that an insured usually cannot guard against such opportunism by purchasing more coverage up front and thereby raising the expected award at the conclusion of litigation. For casualty losses and medical expenses, insurance contracts that promise to pay more than the insured's actual loss are unavailable because of the moral hazard problem.³⁷ And, where coverage is more readily adjustable upward (as in the case of life insurance), it seems unlikely that an insured will often wish to purchase otherwise undesired coverage simply to counter a small probability of insurer opportunism that may never materi-

party is able to make an equal number of offers and the interval of time between offers is always the same.

To see why, suppose that there will be two offers, one by the defendant halfway through the period ($t = .5$) and one by the plaintiff at the beginning of the period ($t = 0$). Assume, as is standard, that an offer will be accepted if the offeree is weakly indifferent between acceptance and rejection. The plaintiff is indifferent between $e'D$ at time $t = 1$ and $e^{-.5(r+\rho)} e'D = e^{.5(r-\rho)} D$ at time $t = .5$. So the defendant will offer the plaintiff no more than this amount at $t = .5$, and the plaintiff will not accept less. The plaintiff then knows that he will settle for this amount if the game reaches $t = .5$. With this knowledge, the plaintiff will make an offer at $t = 0$. The defendant will be indifferent between paying $e^{.5(r-\rho)} D$ at time $t = .5$ and paying $e^{-.5r} e^{.5(r-\rho)} D = e^{-.5\rho} D$ at $t = 0$. The defendant will pay no more than this in settlement, at $t = 0$, and the plaintiff has no reason to offer to take less. So the parties will settle for this amount at $t = 0$, as reported above.

The reader can readily verify that reversing the order of play does not affect the solution—the plaintiff's offer will be $e^{.5r} D$ at $t = .5$, and the defendant will then offer $e^{-.5(r+\rho)} e^{.5r} D = e^{-.5\rho} D$ at $t = 0$. Likewise, with four offers at intervals $t = 0$, $t = .25$, $t = .5$, and $t = .75$, or any arbitrarily large, even number of offers at equally spaced intervals, the solution again remains the same, independent of the order of play.

As the difference between the insured's subjective discount rate and the rate of prejudgment interest goes to zero ($\rho \rightarrow 0$), the settlement amount approaches D , the amount to which the insured is entitled under the contract. As the insured's subjective discount rate rises ($\rho \rightarrow \infty$), the amount that the insured will receive in settlement falls steadily toward zero.

Although the notation suppresses the length of the "period" between the filing of a claim and litigation, the parameters must be understood to depend on it. In particular, as the length of the period increases, ρ increases as well, so that the insured's settlement amount falls.

³⁷ Indeed, even if insurance companies were willing to sell such coverage, the courts might well strike down such policies as "against public policy." See, for example, *Snethen v. Oklahoma State Union of the Farmers Educational and Cooperative Union of America*, 664 P.2d 377 (OK 1983).

alize. Any increase in coverage out of fear of opportunism is also plainly inefficient relative to the ideal insurance contract.

B. The Adequacy of Ex Ante Constraints on Strategic Behavior by Insurers

This section inquires whether market penalties for the denial of meritorious claims will suffice to discourage such behavior, granting *arguendo* that it is sometimes profitable *ex post*. For purposes of this discussion, assume that insurance contracts as written are jointly optimal from the perspective of the parties to the contract. They reflect a proper accommodation between risk sharing and the incentives for loss avoidance, contain appropriate terms regarding conditions of payment and coordination of coverage, and so on. Assume further that the refusal of an insurer to honor a contractual promise will force the insured to incur some cost in seeking third-party enforcement (whether this cost is ultimately shifted to the insurer is unimportant to the analysis that immediately follows). On these assumptions, the insurer’s refusal to honor the contract is plainly inefficient.

Unlike other kinds of contracts in which the possibility of “efficient breach” arises, the insurer’s refusal to pay will not avert production costs that exceed the value of performance, free resources that are more highly valued elsewhere, or otherwise increase joint wealth.³⁸ Instead, the insurance contract is simply a promise to convey money from one party to another; a breach of the promise to do so is a wash *ex post*,³⁹ and the possibility of breach creates joint losses *ex ante*.

In a competitive insurance market with perfect information, insurers who refused to pay meritorious claims could not survive in equilibrium *if* insurers had the ability to commit credibly to pay all covered claims *ex ante*. The logic is obvious—from any supposed equilibrium in which insurers would decline to pay covered claims with some positive probability, an insurer could commit to eschew such behavior and make a profit by raising price to capture part of the joint welfare gains while still leaving insureds better off than they would be dealing with another insurer.

³⁸ Breach of a contract to manufacture and convey a good might be efficient, for example, if costs have risen to the point that they exceed the value of the good to the buyer, or if a third party values the good more highly than the buyer and resale by the buyer is more expensive. See, generally, Steven Shavell, *Damage Measures for Breach of Contract*, 11 *Bell J. Econ.* 466 (1980).

³⁹ It is a wash measuring the parties’ joint wealth in dollars, although the familiar utilitarian argument might be made for transferring dollars to the insured whose marginal utility for dollars is high due to the covered loss.

Hence, for opportunistic refusal to pay covered claims to arise in insurance markets, some mixture of information failure and inability to commit to eschew opportunism must be present.

With respect to information failure, any insurer who frequently refused to pay covered claims would likely soon develop a reputation for behaving in this fashion and lose customers. But it is plausible that insurers might occasionally behave opportunistically without suffering a prohibitive reputational penalty. Customers who learn about the claims of others being denied may have difficulty distinguishing justifiable from unjustifiable denials, and the costs of obtaining accurate information about the claims practices of various insurers may simply exceed the gains. Of course, competing insurers would have an incentive to cure the information problem by disclosing opportunism by other insurers to attract their customers, but the costs of gathering information about such behavior, and the difficulties in persuading customers of its accuracy, might get in the way.

The analysis cannot stop here, however, for even if customers will not know of the incidence of opportunism by various insurers before choosing a policy, they could adopt a strategy of refusing to deal with any insurer who did not credibly commit to pay legitimate claims promptly. If such a commitment device existed, if its costs were sufficiently low, and if its presence in an insurance contract could be signaled sufficiently cheaply, one might expect customers to insist on it and to infer from its absence that a prospective insurer would indeed behave opportunistically with some probability.

Various efforts at commitment can be imagined. For example, an insurer might announce a policy of punishing severely any claims adjuster who is found to have denied a claim without any justification after litigation. But without more, insureds might understandably be skeptical about the likelihood of its enforcement. Carrying the example a bit further, the insurer might then post a bond or add a contractual penalty clause to make its announced policy credible. But claims adjusters might respond to all of this by routinely paying invalid claims to eliminate any risk of punishment. Likewise, such a system would invite litigation over whether the bond should be forfeit or the contractual penalty paid over to the insured. Other devices to motivate claims adjusters to pay all valid claims promptly, such as declining to reward adjusters who save the company money, could also cause far more problems than they solve.

An alternative and perhaps more promising commitment device might strive to eliminate the *delay* between the denial of a claim and an impartial adjudication of its validity. The analysis in the preceding section suggests that this delay lies at the heart of the insurer's ability to take strategic

advantage of an insured. Although the parties to an insurance contract can hardly bind the courts to move more quickly, they can opt out of the court system altogether by contracting for arbitration.

While arbitration agreements may prove valuable, there are several reasons why they may not afford a complete solution to the problem. First, even if arbitration will reach a conclusion considerably more quickly than civil litigation, the time required for it may nonetheless be quite significant for an insured in need of funds. Second, it may be difficult to persuade insureds that arbitrators will proceed in an unbiased and neutral fashion. Third, and related, arbitral decisions are usually subject to challenge in court on grounds of bias, procedural irregularity, and the like, and an insured may fear that arbitration will be followed by a court challenge that would extend the delay until the insurer makes payment. Contractual devices to disable the insurer from engaging in strategic challenges, such as promises not to challenge the arbitral award or promises to give the funds to the insured during the pendency of any challenge, are themselves subject to opportunistic breach that can only be policed by the courts. Finally, and also related, a promise to proceed to arbitration promptly can itself be breached.

A third type of commitment device would simply promise the insured that, in the event a court finds the denial of a claim to have been baseless, the insured will receive some contractual penalty over and above what the policy otherwise promises to pay. This device is analytically similar to an increase in the damages awarded by the courts.

C. Increasing the Damages Award

Putting aside for now the question whether an increase in damages is to be provided contractually or by the courts, what increase in the award would eliminate the incentive for opportunism that might arise when the insured has an urgent need for funds? An award of prejudgment interest is plainly not enough to protect the insured when the insured's subjective discount rate exceeds the rate at which prejudgment interest is computed. Likewise, fee shifting will not suffice, although it will surely help.

Instead, the award must rise appropriately with the insured's subjective discount rate to ensure that opportunism is unprofitable. Of course, any effort to calibrate the penalty closely to the insured's subjective discount rate runs up against the obvious problem that this rate is unobservable, and the further problem that even if it were known its exact effect on the bargaining game *ex ante* is not.⁴⁰ But one can argue that there is no need

⁴⁰ I leave it to the reader to calculate the amount that would, in a formal bargaining model, just exactly eliminate the temptation to renege on the promise to pay the claim. Note

to calibrate the penalty carefully because the insurer can always opt to pay valid claims promptly—"overdeterrence" is not a concern putting aside errors. Thus, if cases in which the refusal by an insurer to pay an indisputably meritorious claim can be identified reliably, one might argue that nothing is lost by having a fairly stiff penalty that appears certain to wipe out any gains from opportunistic behavior.

This last observation perhaps suggests that letting the jury assess damages for the insured's emotional distress or for "punishment" may have some appeal. Indeed, such damages may tend to rise very roughly with the insured's subjective discount rate—the more the insured was desperate for funds when the claim was filed, the greater the degree of emotional distress from the denial of a claim, and the more a jury will likely "punish" the insurer. Statutory penalties that provide a recovery in excess of the ordinary compensatory amount, in accordance with the court's sense of "justice," are another option.⁴¹

But penalties of this sort could also emerge by contract. Certainly it would not be difficult for the insurance contract to provide that, in the event a court finds that a claim is denied without any justification, some fixed and substantial penalty will apply or some penalty set by the court in the interests of "justice" will apply up to some contractual cap.

Given that these types of contractual penalties have never been observed as best I can determine, particularly in the days before the advent of the first-party bad faith action or in the states that do not allow it today, two inferences are possible. One is that they are suboptimal. The possible reasons why are obvious. Courts may have difficulty ascertaining whether an allegation of opportunism by the insurer is meritorious, and juries may tend to err on the side of the sympathetic plaintiff rather than the deep-pocket insurer. Recognizing these problems *ex ante*, insurers may conclude that an increase in the damages remedy will create more problems than it solves, much like some of the other possible commitment devices discussed earlier.

But it also conceivable that the costs of informing insurance customers about the opportunism problem and its possible contractual solutions are too high, that insurance customers are incapable of optimizing with respect to low probability events, and so on. And if the absence of contractual remedies is a market failure, perhaps bad faith doctrine in the courts can usefully respond to it. It thus seems most unlikely that theorizing by itself can resolve whether bad faith doctrine is useful for addressing the

that computing prejudgment interest at the insured's subjective discount rate is excessive for this purpose in the Rubinstein model.

⁴¹ See the reference to the statutory scheme in Illinois, note 28 *supra*.

problem of opportunistic insurers bargaining with impatient insureds. What will become clear from the discussion of the cases in Section IV below, however, is that plaintiffs often win substantial bad faith recoveries on facts that do not reflect this sort of opportunism.

D. A Note on Assignment and Borrowing

Another reason to be somewhat skeptical of the "impatience" justification for tort remedies relates to the possibility that the insured's claim against the insurer can be assigned. The historical hostility to assignment of contract rights at common law has ebbed greatly in modern times, to the point that contract claims are typically assignable, at least where the assignment does not jeopardize the legitimate interests of the other party to the contract and the interests being assigned are not too inchoate.⁴² Although efforts to assign insurance coverage prior to a loss raise a number of problems and are subject to a number of contractual restrictions, the assignment of the right to collect payments for a loss that has already occurred usually presents few problems.⁴³

Thus, an insured with an urgent need for funds can usually sell for cash the right to collect insurance proceeds at the conclusion of litigation. If the assignee is more "patient" and has a subjective discount rate approaching that of the insurer, the ability of the insurer to extract a favorable settlement will be extinguished. The mere possibility that assignment might occur should to some extent discourage insurers from behaving strategically.

Closely related to assignment is the possibility that the insured may be able to borrow against the claim. A well-informed market lender should be willing to advance funds equal to the present value of the payment to the insured at the conclusion of litigation, discounted at a market rate, on the strength of a security interest in that payment.

Two objections to assignment or borrowing as a solution to the problem must be acknowledged. First, if the "market" for the insured's claim is thin, the assignee or lender may be able to bargain for much of the surplus just as the insurer might, thus giving the insured little advantage from the opportunity to assign. But if the insured can shop the claim around and induce potential assignees or lenders to compete for it, or if the insured has a friend or family member who is willing to purchase it for its expected value, this problem will not arise.

⁴² See, generally, Farnsworth, *supra* note 1, Ch. 11.

⁴³ See Keeton & Widiss, *supra* note 9, at 303-4 (noting that insurance contracts usually do not restrict such assignments and that clauses restricting them are often held invalid by the courts anyway).

The second objection is that the insured may have a difficult time persuading potential assignees or lenders that the claim is meritorious. They may be tempted to infer from the insurer's unwillingness to pay that some defense to the claim exists, and that assertions by the insured to the contrary are false. Investigation into the merits of the claim may overcome this problem, but of course investigation is costly.

In the end, readers who see justification for bad faith remedies in market failure may well conclude that the market for assignment of claims and borrowing against them fails as well. But it is important not to neglect it, and its existence provides some further basis for doubt about the need for the move from contract to tort in these first-party actions.

E. Mandatory Arbitration

Even assuming that market failures justify some form of intervention, "bad faith" damages are by no means the only option. The preceding analysis suggests that appropriate prejudgment interest and fee-shifting statutes (or Rule 11 fee shifting) can make opportunism unprofitable in many cases. For what remains, the difficulty lies mainly in the delay associated with civil litigation and the attendant ability of the insurer to exploit the insured's high implicit discount rate. Arbitration can reduce this delay, and if contractual arbitration agreements are not forthcoming, legislatures or insurance regulators might consider mandatory arbitration as an alternative. To ensure speed in the arbitration process, appeals could be limited, or insurers who lose after initial arbitration might be required to pay over the arbitral award to the insured pending any appeal.

A speedy arbitration process would have considerable advantages. As the succeeding sections will make clear, the great danger of existing bad faith remedies is that they will be applied in inappropriate cases, discouraging economically productive behavior by insurers and raising insurance rates unnecessarily. An expeditious arbitration system, by contrast, can eliminate the temptation for opportunism relating to delay without the potential distortions associated with errors in the award of bad faith damages. These potential advantages of arbitration suggest some considerable doubt whether the legal system has selected the appropriate policy instrument to address problems of intentional breach.

III. CASES OF UNINTENTIONAL OR "NEGLIGENT" BREACH

Insurance contracts are complex instruments, typically containing a wide array of exclusions and conditions. These limitations on coverage exist for a variety of reasons—to control moral hazard, to control adverse selection, to avoid wasteful duplication of coverage across multiple insur-

ers, to avert fraud, to exclude from coverage risks that the insurer cannot readily diversify or that exceed what the insurer may permissibly insure under applicable state regulations, and no doubt for other reasons as well.

Although most exclusions and conditions may be presumed to arise for good reason and to serve the contracting parties' joint interests, these terms inevitably lead to disputes. Many of these disputes are factual, such as disputes over the valuation of a casualty loss or the honesty of the insured's representations on a life insurance application. Others go to the meaning of a condition or exclusion (is a loss due to a fire that followed an earthquake a covered loss due to "fire" or an excluded loss due to "earth movement?"). For both types of cases, the dispute may ultimately proceed to litigation, and the insured may in the end prevail. The insured's right to standard contract remedies at that point is a given. But as indicated in Section I above, even the more restrictive formulations of the modern "bad faith" action allow the insured to recover non-economic damages and often punitive damages in cases where the insurer's behavior was sufficiently "unreasonable" by an objective standard. This section inquires whether these additional remedies are justified in situations where the insurer's refusal to pay is not an "intentional" breach of contract as defined earlier. That is, by hypothesis, the insurer's refusal to pay here rests on an honest subjective belief by the insurer that the claim *may* lack merit.

These cases are harder analytically because the characterization of the "first-best" behavior by the insurer is quite difficult and depends importantly on assumptions and on the type of policy provision in question. I cannot hope to provide an exhaustive analysis of all possibilities, and in the end I offer little more than a brief introduction to the kinds of questions that arise. Because of the complexity of the issues and the challenge of formulating a standard for identifying insurer opportunism in theory, much less in practice, the case for judicial intervention is weaker in my view than in the intentional breach setting. This theme will be developed further in the analysis of the cases in Section IV below, but it is first necessary to develop some theoretical background.

A. Factual Disputes: Optimal Claims Denial When Claims Investigation Is Costly for the Insurer and Proof of Claims Is Costly for the Insured

Under any insurance contract, payment is required only if certain facts are true: the insured has suffered a property loss of ascertainable value; the insured has died; the insured has incurred medical expenses associated with a covered illness or injury; the insured has suffered an injury

that is not compensable under another insurance arrangement such as workers' compensation; and on and on. Often such facts can be established cheaply and without much controversy. But the parties may at times confront circumstances in which an important fact simply cannot be proven to the insurer beyond a (perhaps considerable) doubt. In other instances, the insurer will entertain doubts as to the correctness of the factual claims made by the insured that can only be reduced or eliminated after an expensive further investigation. In still other instances, the factual uncertainty will resolve definitively at a later moment in time (such as at the conclusion of litigation against a third-party), and the issue is what to do with the insured's claim in the interim.

In the absence of transaction costs, an ideal insurance arrangement would address these contingencies, specifying the exact proof required to "establish" essential facts, the circumstances under which the insurer must respond to doubt about the facts by incurring additional costs of investigation, and the disposition of the claim pending the outcome of additional investigation or litigation. It would be attentive to the relative burden of establishing the facts, placing the costs of establishing them on the party that can bear them most cheaply. It would also be attentive to the joint costs of uncertainty about the underlying facts and to the joint gains from efforts to reduce uncertainty. These factors would depend in turn on the reasons why the fact is important to the contract in the first place. When the elimination of uncertainty about the facts is uneconomical or will require a considerable period of time, we might expect that insurers would contract to bear the risk of uncertainty in many instances simply because insurers are better risk bearers. But in other instances such an arrangement might create serious incentive problems by reducing the penalty for unproductive behavior by insureds (fraud and moral hazard are two obvious examples).⁴⁴

Given the range of contractual provisions in play and the way in which the optimal bargain may depend on particular facts and circumstances, however, we might anticipate that the transaction costs of addressing these matters expressly will often exceed the benefits *ex ante* and that insurance contracts will then fail to provide much guidance as to the appropriate treatment of factual uncertainties. Indeed, express attention

⁴⁴ The trade-off between optimal risk sharing and optimal incentives to avoid loss (the moral hazard problem) has received the most attention. See, for example, Steven Shavell, *On Moral Hazard and Insurance*, 93 Q. J. Econ. 541 (1979); Ariel Rubinstein & Menahem Yaari, *Repeated Insurance Contracts and Moral Hazard*, 30 J. Econ. Theory 74 (1983). Such issues also lie at the heart of much of the literature on optimal agency arrangements.

to the treatment of factual uncertainties in insurance agreements appears to be fairly rare.⁴⁵

This absence of contractual constraints on the insurer’s ability to reject claims in the face of factual uncertainties no doubt creates possibilities for opportunistic behavior measured against the first-best contingent contract. Just as in the intentional breach setting, an insurer might take advantage of factual uncertainty to try and extract a favorable settlement from an “impatient” insured, even when the insurer suspects that a fact in dispute is probably true and an ideal contract would require payment under those circumstances. Alternatively, the insurer may believe that it has superior resources in litigation and, thus, has a reasonable chance of prevailing, notwithstanding a subjective belief that the claim should be paid despite the associated factual uncertainty. Or the insurer may refuse to pay unless the insured undertakes additional costly efforts to prove the requisite facts under circumstances when the insurer can investigate more cheaply. As with intentional breach, such conduct may be subject to a market penalty *ex ante*, but for the reasons given previously it is not clear that market penalties will suffice.

Yet once we allow that the insurer has some basis for suspecting that a claim may lack merit, it becomes far more difficult to know whether a decision to deny the claim is inconsistent with what an ideal contract would require. The discussion of cases in Section IV below will develop this idea at some length, but in advance of that discussion I wish to make three simple and more general points. First, it is *likely* that optimal insurance arrangements under conditions of imperfect information will allow insurers to deny some claims because of a dispute over the facts even when the insured’s assertions about the facts are true (but not verifiable by the insurer). Second, it is *possible* that an ideal insurance arrangement under conditions of imperfect information would allow the insurer to deny a claim even when an objective observer would conclude that the insured’s assertions about the facts are probably true. Third, it is *possible* that litigation may serve as an efficient screening device for sorting meritorious claims—that is, in the face of factual uncertainty as to which insureds have meritorious claims and which do not, the cheapest device for sorting the good from the bad may be the dispute resolution process that insureds may invoke after a claim has been denied.

To illustrate these points, consider the class of disputes between prop-

⁴⁵ Exceptions arise. I recently received an advertisement for a travel accident policy that contained elaborate provisions regarding the conditions under which a presumption of death will arise if the insured disappears while traveling.

erty insurers and their insureds involving a suspicion that the insured has filed a fraudulent claim. For purposes of the illustration, I will make a number of strong assumptions that I recognize will be inaccurate in many real cases. Suppose that, during the course of routine claims processing, claims adjusters sometimes develop a suspicion of fraud. Perhaps the adjuster will find no evidence of forced entry despite a claim by the insured that the property was burglarized, for example, or the adjuster will detect a suspicious degree of nervousness or uneasiness on the part of the insured in response to questions. Assume also that additional efforts to identify fraudulent claims are extremely expensive (the use of private investigators and the like) and even then unlikely to develop much useful information. Consequently, it is uneconomical to investigate the possibility of fraud further after the suspicion of fraud arises (save perhaps through litigation, discussed later). Further, assume that a suspicion of fraud *invariably* arises in cases of actual fraud, and *rarely but occasionally* in the absence of any fraud. The insurer and insured both know this fact, although the insurer lacks the ability to distinguish genuine cases of fraud on an individual basis. For simplicity, assume in addition that claims adjusters will either pay claims in full or deny them in full—partial denial is not an option. Last, assume that an insured who commits fraud and whose claim is denied is strictly worse off than if the fraud had not been attempted. What is the equilibrium claims denial policy under these assumptions?

Note first that fraud is a source of joint losses. The fraudfeasor/insured may expend resources to hide property, may sell it in an underground market for less than its true value, and so on. Such costs are unproductive and arise solely in the hope of making a profit by extracting a transfer from the insurer. Furthermore, insurers must recover the costs of claims that they pay, and fraudulent claims will cause insurance premiums to rise. The result will be inefficient risk bearing by all insureds, who respond to premium increases by purchasing less coverage. In an ideal world, therefore, the level of fraud would be nil, and the parties to insurance contracts can benefit from measures to control it.

By assumption here, it is prohibitively costly to identify fraud with certainty, and the insurer must decide how to act on the basis of the only information that it is economical to obtain—the suspicion of fraud by a claims adjuster. The insurer has two pure strategies available. It can ignore the suspicion of fraud and pay all claims, or refuse to pay any claim when a suspicion of fraud arises. The insurer might also deny claims with some probability when a suspicion of fraud arises (a “mixed” strategy).

The pure strategy of ignoring the suspicion of fraud in every case is unlikely to be optimal, or to be an equilibrium, because fraud would then

become so rampant that the premiums for theft insurance would make it uneconomical for anyone to purchase it. The pure strategy of denying all claims when fraud is suspected will eliminate fraud under the assumptions above (true fraud always gives rise to suspicion) but at the cost of denying some claims that are valid, thus causing insureds to bear an inefficient level of risk. This could not be a (Nash) equilibrium if it indeed eliminated all fraud, for some insurer would have an incentive at that point to offer a policy under which the suspicion of fraud would be ignored. A policy of denying all suspicious claims is also likely to be suboptimal because under modest additional assumptions a small move in the direction of a mixed strategy could reduce the inefficiency of risk bearing without causing much offsetting cost attributable to increased fraud. Thus, the best feasible insurance arrangement, and the eventual equilibrium, likely involves the denial of suspicious claims with some probability. We cannot say much more without additional assumptions that would allow us to work out the model in greater detail.

We need not go to all that trouble, however, to be confident of two conclusions: first, the parties cannot avoid operating under circumstances where valid claims are denied with some probability—the only way to avoid that problem would be to eliminate any penalty for fraud, at which point the theft insurance market would no longer be viable. Second, depending on parameters, it is *possible* that the best feasible insurance arrangement would result in the denial of a considerable number of meritorious claims as a percentage of the total number of claims denied.⁴⁶ The greater the extent to which fraud causes a large joint loss because of the expenditures necessary to perpetrate it, and the less the insured's risk premium associated with the occasional denial of a meritorious claim, the more it will pay to concentrate on reducing fraud while tolerating a substantial percentage of false positives in the claims denial process.

Of course, the denial of a claim may lead to a lawsuit, and unless the parties provide otherwise the insured can then recover whenever a necessary fact can be established by a preponderance of the evidence. This contractual "default rule" (the right to sue) will typically act as a check on the number of false positives.

⁴⁶ A somewhat related point is made by Kaplow, in a model in which he shows that the private incentive to invest in proof of loss is excessive. After a covered loss occurs of, say, \$1,000, an insured will spend up to \$999 to establish his right to recover that amount. Ex ante, however, such behavior dissipates the gains from being insured against the loss, and an optimal insurance agreement might well restrict expenditures on proof of loss even if the result is that insureds are not fully reimbursed for all insurable losses. See Louis Kaplow, *Optimal Insurance Contracts When Establishing the Amount of Losses Is Costly*, 19 Geneva Paps. on Risk and Ins. Theory 139 (1994).

The possibility that the denial of a claim will result in a lawsuit raises the third point to be made with this illustration. In particular, suppose that insureds with valid claims that have been denied because of a suspicion of fraud will typically file suit. Insureds whose claims are fraudulent, by contrast, will not file suit or will drop the suit in the early stages if it does not result in a quick settlement. Then the willingness of the insured to press the claim to conclusion serves as a screening device and *may* under the right conditions be an efficient screening device.

To make persuasive the claim that the litigation process may help to screen valid claims from fraudulent ones, it is necessary to explain why the costs of litigation are higher for fraudulent claims than for valid claims so that it does not pay fraudfeors to litigate (or perhaps does not pay them to press their claims to conclusion). To make persuasive the further claim that screening through litigation may be efficient, it is necessary to make the additional argument that litigation is a cheaper way of sorting claims than any available alternative. Consider each point in turn.

Insureds whose claims are fraudulent may have good reason to fear that evidence will somehow develop to impeach their factual claims. If such evidence develops during the course of litigation, they may be subject to a range of penalties, from contempt of court to an action for malicious prosecution to criminal penalties for perjury or criminal fraud. Their attorneys may also be subject to sanctions if they have reason to know of the fraud, and thus insureds hoping to perpetrate fraud may run the risk that their attorneys may discover something to suggest fraud and, in turn, signal their discovery by withdrawing from the case. Insureds with valid claims, of course, need not fear such outcomes. Their willingness to testify under oath and to incur the risk of sanctions thus serves as a kind of bonding device that may be much less costly for them than for fraudfeors.

These observations go to the second issue as well—why litigation may be a cheaper way to sort cases than investigation by the insurer outside of the litigation process. In the course of informal dealings with the insured, the threat of civil or criminal sanctions for false statements may be quite remote and unlikely to materialize, whereas the likelihood of serious sanction for demonstrably false testimony given under oath may be much greater. An insured involved in formal litigation may thus be more likely to tell the truth. Furthermore, the insurer's ability to investigate the insured's claims effectively may be considerably increased in litigation by the availability of discovery and the prospect of sanctions for plaintiffs who do not comply with discovery requests in good faith.

For these reasons, it is not implausible that insureds who file fraudulent claims will not file suit after they are denied or will drop the case quickly

if faced with resistance from the insurer. It is also plausible that sorting cases in this manner, despite the need for the parties to incur significant litigation costs, is nevertheless cheaper and more accurate than the alternatives (the use of private investigators and the like).

Although the illustration in this section focuses on fraudulent claims, its lessons are more general. Whenever it is costly for the parties to an insurance contract to establish the facts on which a claim rests beyond a doubt, it may be optimal for them to tolerate erroneous denial of claims, perhaps even when the evidence on which the insurer relies does not permit a terribly powerful inference about the underlying facts. And where the dispute between the parties turns on the truthfulness of factual assertions made by the insured, it is possible that litigation may serve a useful screening function.

B. Disputes over the Interpretation of the Insurance Contract

A glance at any insurance treatise or casebook will confirm that a substantial proportion of the disputes between insurers and insureds involves differences over the interpretation of language in the contract, claims that the justifiable expectations of the insured are frustrated by a strict reading of the language, or disputes over how the insurer should behave in the face of contractual silence regarding a contingency that has materialized. Through the years, the courts have evolved two bodies of doctrine to deal with such disputes. The first involves strict application of the *contra proferentem* principle (construe contracts against the drafter). The second involves decisions that protect the "reasonable expectations" of the insured despite possibly contrary language in the contract.⁴⁷ Some of the cases to be discussed in the next section involve these types of disputes. They are analytically distinct from those discussed previously because the insurer cannot be said to have breached the contract intentionally (unless the claim is that the insurer breached an implicit bargain) and the pertinent facts (other than the meaning of the contract) are not uncertain.

The reader may find it odd that an insurer could behave in "bad faith" by advancing a self-interested interpretation of an ambiguous provision, urging that the clear language of the contract should trump the other party's "expectations," or behaving in a self-interested fashion in the face of contractual incompleteness. Such arguments and behavior are the stuff of many contract disputes, and outside the insurance field no suggestion is made that tort remedies are needed as a response.

⁴⁷ See Keeton & Widiss, *supra* note 9, § 6.3.

One might argue, however, that insurance disputes are different because the probability of success in litigation for the plaintiff/insured is significantly greater in an insurance case. *Contra proferentem* is applied ruthlessly against insurers, even to the point that courts strain to find ambiguity where most observers might conclude that none exists.⁴⁸ And decisions protecting “reasonable expectations,” sometimes in the face of clear language to the contrary, are also much more common in insurance than elsewhere. Thus, it might be argued that in many of these disputes the insurer should forecast a certainty or near certainty of defeat at the conclusion of litigation, even when the insured cannot point to clear language in the contract that establishes a breach. A refusal to pay a claim in the face of this forecast, one might argue, is no different from a refusal to pay when the insurer is in breach of the contract’s plain language. It forces the insured to incur unnecessary litigation expenditures when the outcome of litigation is not in doubt and is only rational because of the insurer’s hope that it can take advantage of impatient insureds by extracting a more favorable settlement.

This argument is only persuasive, however, if one further assumes that the construction of the bargain implicit in the eventual outcome of litigation is consistent with what an ideal insurance contract would require or that the insured’s victory in litigation otherwise penalizes the insurer for unproductive conduct. Thus, for example, if the vindication of “reasonable expectations” is equivalent to forcing an efficient bargain onto an insurer who has duped an insured into agreeing to inefficient terms, it would indeed be better if the insurer simply agreed to honor the jointly optimal bargain when the claim was filed rather than trying to force the insured to settle for less by threatening to litigate. Bad faith remedies discourage this unproductive practice by insurers. Likewise, if contractual ambiguity reflects a calculated effort by the insurer to mislead the insured, the insurer’s conduct borders on fraud and the prospect of tort penalties for it is unobjectionable.

But there are other possibilities. Ambiguity, like contractual silence, may result simply from the transaction costs of addressing every possible contingency in writing. Furthermore, ambiguity is to some degree in the eye of the beholder, and what an insurer thought to be clear at the outset may be confusing to others. Likewise, if the contract is unambiguous but contrary to “reasonable expectations,” the insurer may be ignorant of the contrary expectations, and in any case those expectations may not

⁴⁸ *Id.*

reflect the jointly optimal bargain. The application of contra proferentem or reasonable expectations doctrine may thus yield a result that is quite inconsistent with what is in fact jointly optimal.

If so, it is hardly clear that a full recovery by the insured according to the insured's construction of the bargain is desirable. A fortiori, it is not so troubling that the insured may, in the absence of bad faith remedies, settle for less than the present value of the amount that a court may eventually hold to have been due under the policy.

IV. BAD FAITH DOCTRINE IN PRACTICE: "UNREASONABLE" DENIAL OF THE INSURED'S CLAIM

Many of the most widely cited cases on first-party bad faith, such as *Gruenberg v. Aetna*⁴⁹ and *Anderson v. Continental Insurance*,⁵⁰ were decided on the pleadings. The allegations of misconduct by the insurer in those cases, if true, would establish an "intentional breach" of contract by the insurer as that term was used in Section II above, and the availability of tort remedies is potentially attractive.

Among the cases that are litigated to conclusion, however, the basis for the finding of "bad faith" is often much less comfortable. This section reviews a number of these cases, selected because of the range of interesting issues that they raise and, in some instances, because of the size of the judgment against the insurer.

I do not suggest that this sample of cases is in any sense random or that it provides the basis for a statistically meaningful inference about the mine run of bad faith cases. Likewise, by selecting cases that in my judgment reflect erroneous findings of "bad faith," I do not mean to claim that genuine bad faith never arises in practice. Rather, the purpose of this section is to suggest how the courts can carry the doctrine too far, especially where evidence of "intentional breach" as defined above is lacking. The conclusion will offer some tentative thoughts about the implications of this observation for reform.

⁴⁹ Cited in note 21 *supra*. In *Gruenberg*, the plaintiff filed a claim under a fire insurance policy. The plaintiff alleged, in effect, that the insurer's agent suggested to the police, without any basis, that the insured had committed arson and then denied the claim when the insured, acting on the advice of counsel, refused to appear at a civil deposition while the criminal charges were pending. Those charges were later dismissed for lack of probable cause.

⁵⁰ Cited in note 25 *supra*. *Anderson* involved an allegation that a property insurer refused to make payment on a casualty loss despite the fact that coverage was clear and the insured had presented proof of loss to which the insurer had raised no objection.

A. Suspicious Claims

In *Frommoethelydo v. Fire Insurance Exchange*,⁵¹ the insured submitted a claim for property that he claimed to have been stolen in a burglary. It was his second claim for burglary within a year. As part of his claim, he asserted that some \$3,000 worth of audio/video equipment had been stolen, and he produced a receipt for it. The date on the receipt appeared to have been altered, which was confirmed by an expert on documents alteration, and a private investigator for the insurer found a sales clerk at the store that issued the receipt who indicated that the insured had asked him to backdate it. The true date of sale, as best the insurer could determine, was *after* the date of the alleged burglary. The insurer made a report to the authorities to this effect, and the insured was arrested for insurance fraud. On the morning of trial, however, the prosecutor dismissed the charges, apparently because the insured had found some witnesses who would testify that they had seen substantial quantities of audio/video equipment in the insured's home prior to the alleged burglary. The insurer thereafter continued to refuse to pay the claim, and the insured filed suit. The jury found bad faith, awarding the full amount of the property claim (less than \$9,000), emotional distress damages of \$250,000 and punitive damages of \$1.25 million. The California Supreme Court affirmed the jury's finding of bad faith but reversed the damages award except for the amount of the claimed loss. It reasoned that the bulk of the damages sustained by the insured related to the aborted criminal prosecution, instituted as a result of a privileged communication from the insurer to the authorities that was not made with any malice. It thus ordered a new trial on the damages question.

The case is interesting mainly for the court's determination that the insurer had acted in bad faith despite the fact that its report to the authorities was not malicious and in any case was privileged. It found bad faith, as a matter of law, in the insurer's failure to investigate the claim further once it knew that the criminal charges had been dismissed and that the insured claimed to have witnesses who had seen audiovisual equipment in his home prior to the burglary.⁵² The facts correspond nicely to the illustration in Section III above, where the insurer has good reason to suspect fraud and where further investigation is both costly and unlikely to resolve the matter definitively. It is possible that the salesperson lied about backdating the receipt despite the absence of any apparent incentive to do so, and it is also possible that the insured wanted a backdated

⁵¹ 42 Cal. 3d 208; 228 Cal. Rptr. 160; 721 P.2d 41 (1986).

⁵² 42 Cal. 3d 220.

receipt because he had lost the receipt for equipment that he had in fact purchased prior to the second burglary and he needed to fabricate proof to support his claim of a genuine loss. But it is also possible that the insured was engaged in fraudulent behavior and that any witnesses he had found to testify at his trial were prepared to lie themselves, were mistaken in their recollection of when they saw the equipment, or the like. Under these circumstances, it is imminently plausible that an optimal policy would permit the insurer to deny the claim or to decline to incur further costs of investigation.⁵³ In the face of contractual silence on such matters, any judgment to the contrary seems pure conjecture.⁵⁴

Another interesting case in this genre is *T.D.S. Inc. v. Shelby Mutual Insurance Co.*⁵⁵ A fire at a restaurant was of suspicious origin, and the evidence of arson was strong upon investigation. It was unclear who started the fire, but one investigator found evidence of incendiary activity inside a locked office to which only the insured had the keys.⁵⁶ The

⁵³ Much the same criticism might be leveled at *Capstick v. Allstate Insurance Co.*, 998 F.2d 810 (10th Cir. 1993). The insured claimed that his used automobile, worth about \$1,500, caught on fire while he was trying to repair it. The insurer suspected that the fire had been deliberately set and found some evidence of a flammable substance in the interior of the vehicle after the fire in a location that was difficult to reconcile with an accidental gasoline spill during mechanical repairs. Both an investigator and a mechanic retained by the insurer doubted the insured's explanation for the fire, although they could not rule it out as a possibility. Nonetheless, the case was allowed to go to the jury on the theory that the insurer trumped up the expert testimony knowing that it had no reasonable basis for doubting the insured's version of the facts and failed to investigate the claim adequately by, *inter alia*, failing to talk to the insured's neighbors who had seen the vehicle on fire. The trial judge allowed that the jury could find "clear and convincing evidence" of "wanton or reckless disregard" for the rights of another. On appeal, the Tenth Circuit affirmed the jury's award of \$1,500 in ordinary contract damages, \$3,000 in additional consequential damages, and \$2 million in punitive damages.

⁵⁴ Another California case imposing bad faith liability for failure to investigate a claim adequately is *Egan v. Mutual of Omaha Ins. Co.*, 24 Cal. 3d 809; 157 Cal. Rptr. 482; 598 P.2d 452 (1979). In *Egan*, the insured claimed disability benefits on an ongoing basis, despite notations on his doctor's records that he would be expected to return to work much earlier. Later, after back surgery, he claimed to be disabled again, but information in his medical records at that point suggested that his perpetual back problems were due to a degenerative disc and thus to "illness" rather than "injury," for which disability benefits were more limited. Thus, the case involved an initial suspicion of fraud and a subsequent dispute over whether any bona fide disability was covered given its origin. The California Supreme Court affirmed a finding of bad faith against the insurer based on abusive statements to the insured by claims adjusters (they called him a "fraud"), their failure to talk personally with his physicians rather than basing their denial of the claim on the written record, and their failure to have him examined by their own physicians. It set aside the jury's \$5 million punitive award as excessive, however, and sent the case back for retrial on the issue of the proper level of punitive damages.

⁵⁵ 760 F.2d 1520 (11th Cir. 1985).

⁵⁶ 760 F.2d 1524.

insurer thus suspected the insured all along but concealed this suspicion during the course of various contacts with the insured. When suit was brought, the jury rejected the insurer's claim that the insured was the arsonist and proceeded to award punitive damages in the amount of \$2,100,000, an amount that was affirmed on appeal.

The insurer evidently held an honest belief that the insured was the arsonist, and no exculpatory evidence ultimately came to light. Nonetheless, the jury evidently concluded that the affirmative defense to contractual liability (arson) asserted by the insurer was not established. The additional bad faith damages were awarded not on the premise that the arson defense was frivolous or asserted without basis but on the insurer's "fraudulent" concealment of its suspicion that the insured was the arsonist, which allegedly hampered the ability of the insured to exculpate himself. Yet it is not difficult to understand why an investigator would not in general desire to disclose his suspicions to the target of the investigation. If that party had indeed committed the suspected acts, placing him on notice of the suspicion might lead to additional unproductive efforts at concealment that would increase the investigation costs of the other party and perhaps undermine the ability of the investigation to reach the correct result. The punitive award here seems as likely to raise the costs of policing fraudulent claims uneconomically as to enable those suspected of fraud to prove their innocence.

Cases such as *Frommoethelydo* and *T.D.S., Inc.* raise the distinct possibility that bad faith doctrine may discourage jointly valuable efforts by insurers to police fraud. This danger is doubly troubling in light of recent fears that insurance fraud is rampant. Although its exact magnitude is difficult to ascertain for obvious reasons, various estimates suggest that excess medical claims alone may add over \$100 per year to the average cost of auto insurance for each consumer,⁵⁷ that about 10 percent of all auto insurance claims are fraudulent,⁵⁸ that property/casualty insurance fraud may amount to about \$20 billion annually,⁵⁹ that health care fraud may be as much as \$100 billion annually,⁶⁰ and that workers' compensation fraud may account for as much as 25 percent of all claims.⁶¹ It has

⁵⁷ See Stephen Carroll, Allan Abrahamse, & Mary Vaiana, *The Costs of Excess Medical Claims for Automobile Personal Injuries* (Rand Institute for Civil Justice 1995).

⁵⁸ See Insurance Information Institute Reports, *Insurance Fraud* (Ruth Gastel ed. June 1995) (citing study by National Insurance Crime Bureau).

⁵⁹ *Id.*

⁶⁰ *Id.* (citing Justice Department study).

⁶¹ *Id.* (citing statement by New Jersey workers' compensation official and Insurance Research Council poll).

even been suggested that insurance fraud is second only to illegal drug trafficking in the amount of illicit revenue that it generates for perpetrators.⁶² The potential joint returns to insurance companies and honest policyholders from reductions in fraud thus appear to be quite considerable across the board. If reasonable efforts to reduce fraud can become "bad faith" down the road, however, insurers may simply give up on them.

B. Factual Disputes Relating to Exclusions

*Aetna Life Insurance Co. v. Lavoie*⁶³ has a complex factual history, but in essence the facts were as follows. The plaintiff went to a physician complaining of arthritic pain, rectal bleeding, and a general sense of malaise. She was admitted to a hospital, where she spent 23 days as an inpatient, undergoing a wide variety of tests to rule out various problems. Ultimately, she was diagnosed as having arthritis, which had already been diagnosed some time earlier by the same physician, coupled with a mild gastrointestinal problem that was thought to have a psychological cause. The hospital bill was submitted to Aetna, plaintiff's first-party insurance carrier, and about half was paid in timely fashion. Aetna disputed the remainder of the bill on the grounds that the charges were not "necessary" or "reasonable," in accordance with an exclusion to that effect in the policy. In particular, Aetna disputed the need for certain tests (an electrocardiogram, an electroencephalogram, and a CAT scan) in light of the plaintiff's symptoms, and it disputed the need for the extended in-patient stay, arguing that the necessary tests should have been done on an out-patient basis. The plaintiff's physicians, not surprisingly, wrote letters to Aetna defending the necessity and reasonableness of the treatment.

Nothing in the policy stated precisely who would make the determination of reasonableness and necessity or how it would be made. According to the court, however, it was standard procedure at Aetna for complicated issues of this sort to be reviewed by internal physicians, and the evidence suggested that Aetna's initial denial of the claim was made without internal-physician review. An Aetna claims officer had nevertheless reported to the insured that the initial decision to deny the claim followed a review by the "medical department." The trial court, and the Alabama Supreme Court, seized on this fact as critical evidence of bad faith, along with evidence that certain portions of the plaintiff's medical record (such

⁶² See *Fighting Fraud*, 93 No. 9 Best's Rev.—Life-Health Ins. Ed., January 1993, at 56.

⁶³ 470 So.2d 1060 (Ala. 1984).

as the "nurses' notes") were never obtained or reviewed by certain claims officials at Aetna.⁶⁴ As the dispute dragged on, Aetna did have its internal physicians review the claim and continued to maintain that the denial was proper. At trial, Aetna also produced medical experts to support the proposition that the charges Aetna refused to pay were not reasonable or necessary. Nevertheless, the Alabama Supreme Court concluded that "[t]he evidence is overwhelming that Aetna acted in bad faith,"⁶⁵ and it affirmed the jury verdict for \$3.5 million in punitive damages.⁶⁶

Both the finding of bad faith in *Lavoie* and the size of the punitive award attached to it raise serious concerns. Alabama purports to adhere to the *Anderson* standard for a finding of bad faith, under which the insurer does not act in bad faith if it has a "legitimate or arguable reason" for failing to pay the claim—if the claim is fairly debatable.⁶⁷ Even assuming that the claims adjuster who initially denied the claim was lying about consultations with the "medical department," it does not follow that the claim itself was not fairly debatable. As noted, Aetna had no difficulty producing experts who would testify to that effect and, perhaps more important, it does not seem preposterous to question the need for a 23-day hospital stay, complete with CAT scan and a battery of other tests, for a patient whose only problem in the end was preexisting arthritis and a nervous stomach. To be sure, medical tests are often appropriate to rule out more serious problems, and perhaps the plaintiff's physicians did nothing that many others would not have done under the circumstances. But the magnitude of the medical undertaking in relation to the proven severity of the patient's problems raises at least a suspicion that the plaintiff's physicians may have gone overboard in their expenditure of resources on diagnosis. The tendency of physicians to engage in excessive treatment under indemnity plans is, of course, a serious problem and is often said to be one of the reasons why health care costs in the United States are "excessive."⁶⁸

⁶⁴ *Id.* at 1066–68.

⁶⁵ *Id.* at 1066.

⁶⁶ For another Alabama case with large punitive damages, this time in a dispute over whether the applicant for disability insurance had misrepresented his income or become disabled because of a preexisting condition, see *Nationwide Mutual Ins. Co. v. Clay*, 469 So.2d 533 (Ala. 1985) (\$1.25 million award affirmed).

⁶⁷ 470 So.2d at 1066, citing *Gulf Atlantic Life Ins. Co. v. Barnes*, 405 So.2d 916 (Ala. 1981).

⁶⁸ See, for example, "Managing Health Costs Requires Many Actions," *Bus. Ins.*, November 13, 1989, at 25 (referring to studies suggesting that as much as 25 percent of all health care expenditures are "inappropriate and unnecessary").

The Alabama Supreme Court did not directly disagree but instead focused on the departures from Aetna's internal procedures and the apparent false statement by a claims adjuster to establish bad faith. Even granting that Aetna's claims adjuster departed from established claims processing procedures, however, it is hardly clear that massive penalties should attach. Such a rule discourages insurance companies from formulating internal procedures that may reduce the transaction costs of claims processing. Likewise, where those procedures do exist, such a rule may induce them to engage in wasteful monitoring expenditures to prevent agents from departing from established procedures even where, as may have been the case here, the departure is harmless and the denial of the claim was in fact justified.

Finally, and perhaps most importantly, the dispute in *Lavoie* is best viewed as a dispute between physicians and an insurer over the physician's right to payment for questionable charges, not as a dispute between an insurer and an insured hard-pressed for cash. The hospital had taken the position that the plaintiff was responsible for the hospital bill if the insurer refused to pay it, but there was no evidence that the insured had actually paid the hospital anything relating to the claims that were denied. The hospital had not filed suit against the insured, and it seems unlikely that it would do so given that the plaintiff was elderly and living on Social Security and that Aetna had determined the disputed charges were medically unnecessary. In all likelihood, therefore, the only party who stood to lose from the denial of the claim was the hospital, and the case simply does not fit the mold of the opportunistic insurer deliberately exploiting the insured's dire straits to extract a more favorable claim settlement. Rather, it seemingly represents an effort by an insurer to police excessive medical diagnostics under a sensible policy exclusion, in the face of inevitable uncertainty and disagreement about their necessity.

A different set of issues arose in *Eichenseer v. Reserve Life Insurance Co.*⁶⁹ The insured purchased medical insurance and, less than 3 weeks later, entered the hospital complaining of acute abdominal pain. Shortly afterward she underwent a hysterectomy. The policy excluded coverage for preexisting conditions, and the attending physician at the hospital made a note on the medical record that the patient had experienced pain in the lower abdomen for "2 or 3 years." In fact, this was an error, and he had apparently meant to write "2 or 3 days." The insurer denied the claim on the basis of the incorrect notation, and thereafter ensued a comedy of errors in which the physician initially refused to alter the

⁶⁹ 881 F.2d 1355 (5th Cir. 1989), eventually affirmed in 934 F.2d 1377 (5th Cir. 1991).

medical records, but finally agreed after 9 months to submit an affidavit acknowledging his mistake. For its part, the insurer apparently lost the insured's medical records for a time, which would have revealed an absence of prior complaints about abdominal pain, and later temporarily lost the physician's affidavit acknowledging his error. Eventually, the insurer assembled all the documentation and paid the claim, but not before the insured had filed an action that she refused to dismiss thereafter. The trial court awarded her \$1,000 in compensatory damages and \$500,000 in punitive damages, an amount ultimately affirmed after further haggling about the constitutionality of the punitive award.

Had it been shown that the insurer deliberately "lost" information in the claim file to avoid paying the claim, an intentional breach would exist for which tort remedies might be appropriate. But the court's own characterization of the facts suggests not an opportunistic or malicious insurer but a bumbling insurer, which initially perceived legitimate grounds to deny the claim and subsequently, through a series of bureaucratic mistakes, took an inordinate time to respond to the evidence that its initial information contained a key error, the initial error being someone else's fault. The court attributed the insurer's incompetence to its "severe organizational problems" owing to understaffing and a recent surge in claims.⁷⁰ The trial court found simply that the insurer failed to make "reasonable efforts" to investigate the claim as required by Mississippi law and that its failure to do so was in "reckless disregard" for the rights of the insured.⁷¹

It seems quite likely that the insurer in this instance failed to afford the insured the quality of claims processing services that an ideal insurance contract would require and, more precisely, failed to take jointly cost-effective measures to reduce the incidence of error. It may thus be seen as "negligent." The wisdom of a large punitive award in such a case raises issues familiar from the discussion of punitive damages in negligence cases generally.⁷² If the standard of "reasonableness" is clear, set optimally, and administered without error, damages that exceed full compensation (in expected value) are unnecessary to proper deterrence but do not cause overdeterrence problems because the actor subject to punitive liability can avoid it simply by behaving reasonably. But if the standard of care is unclear, sometimes set incorrectly with respect to a

⁷⁰ 881 F.2d at 1357.

⁷¹ 881 F.2d at 1362.

⁷² See William Landes & Richard Posner, *The Economic Structure of Tort Law* 184–85 (1987); Steven Shavell, *Economic Analysis of Accident Law* 146–51 (1987).

given actor, and administered with nontrivial risk of error, punitive damages may well cause actors to be excessively careful. Likewise, insurers are not unitary actors, and the task of motivating agents to behave appropriately is nontrivial. Punitive damages for agents' negligence can easily induce excessive monitoring to prevent it. After *Eichenseer*, insurers may be induced to overinvest in claims processing, in the monitoring of claims processors, and so on, for fear that their own claims processing errors can be characterized as "reckless disregard" for the insured and lead to large punitive liability. And even if uneconomic investments in claims processing do not result, the prospect of unavoidable punitive liability will still cause premiums to rise needlessly.⁷³

C. Contractual Silence and Ambiguity

In *Silberg v. California Life Insurance Co.*,⁷⁴ the plaintiff was injured while working at a cleaning plant and filed a workers' compensation claim. He also filed a claim under his first-party medical policy (issued by the defendant), which contained an exclusion for injuries that were covered by the workers' compensation system. The policy limit under the first-party medical policy was \$5,000, with a \$100 deductible. The workers' compensation insurer disputed coverage on the grounds that the plaintiff was the owner-operator of the plant where the accident occurred and thus was not a covered "employee." During the pendency of the dispute over workers' compensation coverage, which took over a year to resolve, the defendant refused to pay the plaintiff's medical claims under the first-party policy. A combination of the plaintiff's disabling injury and his inability to pay various creditors, including his doctors, eventually caused the plaintiff to lose his business and suffer two "nervous breakdowns."

The dispute with the workers' compensation insurer was finally settled for several thousand dollars. The plaintiff then sued the defendant for medical expenses in excess of the workers' compensation settlement and for "bad faith." The defendant took the position that it owed nothing because the injury was, in fact, covered by workers' compensation. The

⁷³ Perhaps it might be argued that punitive damages for "negligence" in claims processing are needed because some claims are not worth pursuing without them and insureds with small claims will not bother to file suit without some extracontractual kicker—an argument akin to the underdetection rationale for supracompensatory damages elsewhere. But the need for additional penalties is not nearly so obvious once the market is taken into account, as insurers with inept claims departments will surely lose customers. Further, even on the assumption that market penalties are inadequate, a combination of fee shifting and prejudgment interest ought suffice to address this problem.

⁷⁴ 11 Cal. 3d 452, 113 Cal. Rptr. 711, 521 P.2d 1103 (1974).

plaintiff argued that the policy should be construed so that the defendant was liable for any *amounts* not covered by workers' compensation, not simply for injuries or illnesses not covered. He further argued that, during the pendency of the dispute with the other carrier, the defendant should have paid the plaintiff's medical bills and filed a lien on any benefits later received by the plaintiff from the workers' compensation carrier to which the plaintiff was not entitled under the terms of the first-party policy.

The jury, affirmed on this issue by the California Supreme Court, found that the workers' compensation exclusion was ambiguous so that the defendant was obliged to cover medical expenses not paid by workers' compensation (*contra proferentem*). It awarded the plaintiff \$4,900 in compensatory damages (the policy limit less the deductible), \$75,000 in compensatory damages for "physical and mental distress," and \$500,000 in punitive damages. Although the trial judge had set aside the punitive award and was affirmed on that issue, the \$75,000 compensatory award was upheld by the California Supreme Court because the evidence showed a breach of the covenant of good faith and fair dealing "as a matter of law."⁷⁵

The case has a number of troubling features. Because the workers' compensation carrier settled and because the defendant had a solid argument from the language of the policy that it excluded coverage for injuries covered by workers' compensation,⁷⁶ the court found bad faith primarily in the failure of the defendant to advance funds to the plaintiff while he haggled with the workers' compensation carrier over its obligations. By implication, the defendant would have acted in "bad faith" even if the plaintiff's medical bills had eventually been covered in full by workers' compensation (as they might well have been had the plaintiff elected to take the workers' compensation carrier to court instead of settling with it and suing the first-party carrier).

Yet it is by no means clear that an ideal insurance arrangement would require the insurer to advance funds to the insured in the manner that the court found essential to "good faith." First, even granting *arguendo* that funds should be advanced to the insured by someone during the pendency of a coverage dispute involving two carriers, which carrier has that obligation? The court just as easily might have said that the workers' compensation carrier should have advanced funds, at least up to the

⁷⁵ 11 Cal. 3d 462.

⁷⁶ 11 Cal. 3d 461-62. A dissent argued that the injury was clearly covered by workers' compensation and that the policy unambiguously excluded coverage for it—thus, said the dissent, the defendant owed the plaintiff nothing. 11 Cal. 3d 466-68.

amount that it could subsequently recover by asserting a lien on the proceeds of the first-party policy.

Second, policy provisions that allocate the loss in cases of overlapping coverage have utility in allowing insurers to price their products appropriately and in putting each carrier on notice of the particular moral hazard and adverse selection problems that it may have to confront. A rule that reduces the likelihood that the loss will fall on the proper carrier is a source of potential joint costs. Plainly, the insured's incentive to press a claim against one of the carriers diminishes, or disappears altogether, if the beneficiary of the successful claim is another insurance company rather than the insured himself. True, one carrier might take an assignment of the insured's claim against the other, but unless the insured has an incentive to cooperate in pressing the claim fully an assignment may not be of much value. Indeed, the insured may have an incentive to compromise the claim in return for a secret payment that the insurer who has advanced funds may never discover.

Third, the procedure suggested by the court for advancing the funds and securing their repayment—by obtaining a lien on any proceeds from the other policy—is not costless. When the stakes are only a few thousand dollars, as in *Silberg*, requiring the first-party insurer to become involved in the legal proceedings against the workers' compensation carrier may impose costs that exceed any gains.

For these reasons, it is not obvious that an ideal insurance policy would require insurers to advance funds to the insured in the event of a coverage dispute such as that in *Silberg*. It also seems likely that these coordination of coverage issues come up sufficiently often that insurance contracts, or at least customary practice in the business, would have something to say about the appropriate treatment of the problem. But despite the absence of any contractual requirement to advance funds or any custom in the business of doing so, the conduct of California Life here was nevertheless bad faith "as a matter of law."

*Sparks v. Republic National Life Insurance Co.*⁷⁷ is also troubling. The insured and his family were involved in the crash of a private plane, leaving them with sizable and ongoing medical expenses. Because of injuries to the insured, his business eventually closed and he stopped paying premiums on his health insurance policy. The policy provided for cancellation in the event of nonpayment of premiums, and the insurer insisted that medical expenses incurred after the policy had been terminated would not be reimbursed. The insured, however, argued that he

⁷⁷ 647 P.2d 1127 (Ariz. 1982).

had "expected" the insurer to cover all medical expenses relating to any injury or illness commencing when the policy was in force, even if the policy was thereafter terminated.

The policy contained provisions stating that the insurer would pay benefits while the policy was "in force" and additional provisions expressly limiting the liability of the insurer for costs incurred after the insurance was no longer in force. The latter provisions provided that no such expenses would be covered beyond the date when the "policy is terminated."⁷⁸ The trial court, affirmed by the Arizona Supreme Court, nevertheless held that the policy was ambiguous on the question whether the insurer was liable for expenses incurred by the insured after the termination of the policy. The Arizona Supreme Court reasoned that the language of the policy was confusing and that the coverage limitations were hidden away in sections where the insured might not look for them. It also argued that the sales brochure had said nothing about the coverage limitations at issue, creating a reasonable expectation on the part of the insured that coverage would survive the termination of the policy.⁷⁹ Accordingly, the insurer was in breach of contract. Moreover, said the court, a jury could find that the insurer had no "fairly debatable" basis for denying the claim in the first place. It thus affirmed an award of \$1.5 million in compensatory damages (mostly consequential and far in excess of the policy limits) and \$3.0 million in punitive damages.

The insurer in *Sparks* can perhaps be faulted at some level for taking advantage of the insured's foolishness. Had the insured known that a failure to pay premiums would have terminated coverage, the insured might have made certain that premiums were paid to keep coverage in force. But there was no suggestion that the insurer affirmatively misled the insured on this issue, simply that the provisions terminating coverage for nonpayment of premiums were confusing to a lay person and hidden away in the fine print. Many provisions of insurance contracts have this quality about them, yet they are vital to the ability of insurers to price their products appropriately and to control the risks that they undertake to bear. For this reason, the readiness of courts to find ambiguity or confusion in technical language, and to protect "reasonable expectations" that run contrary to the fine print, may be questioned. But even if it is proper to hold insurers to promises that they never make on this basis, it strains credulity to suggest that claims are not "fairly debatable" when the language of the contract supports the insurer's position. Mas-

⁷⁸ 647 P.2d at 1133.

⁷⁹ 647 P.2d at 1133-35.

sive punitive awards in such cases can increase the costs of insurance uneconomically and reduce the efficiency of risk bearing in the insurance market at large.

V. CONCLUSION

Insurers may at times employ lamentable tactics to reduce their payments under first-party insurance policies, primarily by exploiting the delay inherent in the civil litigation process to induce a needy insured to settle for less than the amount that the contract promises. The prospect of tort remedies at the end of the litigation process can make such tactics unprofitable and thus serve a potentially valuable function. Yet we must be alert to the danger that the remedy may be worse than the problem—a fact that may explain the absence of greater market response to it. Nevertheless, one cannot rule out the possibility that the courts can step in constructively when an insurer knowingly refuses to make good on its clear obligations.

Unfortunately, however, the courts seem to find tortious conduct on the part of insurers who have bona fide disputes with their policyholders over the terms of the policy or over factual issues essential to the insured's right to recover. The ability of the courts to identify opportunistic behavior in such cases is very much in doubt, and the distinct possibility arises that bad faith doctrine here does little to police misconduct while doing much to cause uneconomic increases in the premiums that policyholders must pay.

One solution might be to modify the instructions in bad faith cases to make clear that the insured must prove that the insurer subjectively believed the claim to be valid but denied it anyway. Such a requirement was apparently intended by the Wisconsin Supreme Court in *Anderson*. The difficulty, of course, is that proof of state of mind often relies on evidence about what a reasonable person would have believed given the facts, and so courts (like the *Anderson* court and other courts that purport to follow *Anderson*) inevitably allow cases to go to the jury where the proof of bad faith rests on arguments about what "reasonable" people would have believed given the facts. This practice opens the door to findings of bad faith when the denial of the claim was at most an unintentional breach and perhaps even justifiable. As a practical matter, therefore, it may be quite difficult to limit tort remedies to cases of true intentional breach. Nevertheless, some progress might be made through a tightening of jury instructions to emphasize the need for a finding that the insurer's agent knowingly denied a meritorious claim.

Another solution, of course, is to abolish bad faith remedies altogether

and restrict the plaintiff to contract damages plus interest, along with Rule 11–style fee shifting if the court believes that the insurer’s defense was frivolous. Those of us who are skeptical of the market failure arguments for intervention in the first-party cases will assuredly favor this approach, and even those who are more willing to find market failure must concede that a serious question arises whether the error costs associated with modern first-party bad faith doctrine outweigh its benefits. It is also apparent that some sort of mandatory arbitration system, designed to resolve disputed claims quickly, may well be a superior policy response to the problem of opportunistic claims denial, if indeed that problem is serious enough to require any response at all.