INSURABILITY AND PUNITIVE DAMAGES

George L. Priest*

INTRODUCTION

This Article addresses issues relating to the insurability of punitive damages awards. There is no clear consensus today whether the law should allow insurance coverage of punitive damages. Our courts conflict sharply: some deny coverage on grounds of public policy; the majority allow coverage. Courts on both sides have regarded the insurability issue as implicating the most central goals of modern tort law, but they have failed to generate a consensus regarding whether coverage or the refusal of coverage best achieves these goals.

Indeed, there is substantial confusion in each of the alternative judicial approaches to the insurability of punitive damages. Many courts, for example, deny insurance coverage on grounds that it violates public policy to allow insurance to diminish the deterrent effect of punitive awards, which is the awards' chief function. Yet, even courts adopting this approach have crafted exceptions to the rule, allowing insurance coverage where the award serves some compensatory purpose or where other legal principles, such as strict vicarious liability, are implicated. The rationale for these exceptions, however, is not well worked out. Courts have not explained why there should be less concern over diminishing the deterrent effect of vicarious liability or why and to what extent the compensation goal should trump the deterrence goal.

Other courts have ruled, in contrast, that the deterrent effect of punitive damages can be achieved even if insurance coverage is allowed because insurers can raise premiums after an award or can

* Copyright 1989 George L. Priest

* John M. Olin Professor of Law and Economics and Director, Program in Civil Liability, Yale Law School. I am grateful to Judyth Pendell and to participants at the Conference on Punitive Damages, University of Alabama School of Law, for helpful advice and criticism, J. Hoult Verkerke and Daniel Fegan for research assistance and the Program in Civil Liability, Yale Law School, for support.
build the expected cost of punitives *ex ante* into the premium. Yet many of these courts refuse insurability where the underlying act is criminal or clearly intentional, though the opportunity for premium adjustment would appear equivalent. Still other courts have allowed insurance coverage of punitives even where the act is criminal on the ground that the separate criminal penalty will achieve the deterrent effect.

Confusion over punitives is not confined to the courts. The insurance industry itself appears confused on the issue. Insurability litigation involves the assertion by insurers that the general terms of a liability insurance policy should be read to deny coverage of punitives. But this argument of the insurer is necessary only because the insurer for some reason has failed to incorporate a more specific exclusion of punitive coverage into the basic policy itself. The failure of insurers to commit for or against punitive coverage has itself generated judicial schizophrenia on the issue. Many courts are led to resolve the insurability issue on simple grounds of insurance contract interpretation, while other courts—those that insist on maintaining the integrity of the punitive deterrent—commonly view the terms of the insurance contract as irrelevant. Some courts in fact have denied coverage on grounds of public policy, though the insurer concedes that the contract must be read to cover the punitive award.

This Article seeks to clarify these various issues through application of the modern understanding of the economics of insurance. Much of the contradiction in the judicial treatment of the insurability of punitives directly implicates the operation of insurance markets, in particular the disagreement over whether allowing coverage impairs the deterrent function of punitive damages. But other issues relating to insurability and, more generally, to the appropriate role of punitive awards become clearer when the economics of insurance are better understood. For example, simple principles of insurance can explain how allowing coverage of punitive awards affects the broader set of insureds and affects basic liability insurance availability. As we shall see, these principles show that the modern expansion of punitive damages liability, perhaps ironically, has increased the ability of insurers to provide effective insurance coverage of punitives, though it simultaneously has diminished the availability of basic liability insurance for an important segment of the insured population.
These insurance principles also demonstrate the very close relationship between the economic grounds for excluding punitive coverage and the moral grounds for awarding punitive damages in the first instance. The economic principles that define when the insured population benefits from the exclusion of punitive coverage parallel the moral grounds for punitive awards. The Article argues that there is an inverse relationship between feasible insurance of punitives and the moral or instrumental justification for punitive awards. To the extent that the expansion of punitive damages awards in modern times increases the feasibility of insuring punitive liability, it diminishes the underlying moral justification for punitive damages.

Part I describes briefly the current legal treatment of the insurability of punitive damages awards, documenting the disparate judicial approaches to the question. Part II, then, demonstrates the close relationship between the exclusion of insurance coverage and the definition of punitive liability. Section A sets forth the simple economics of insurance, explaining generally how insurance reduces risks and which risks can be effectively insured. Section B discusses the similarity between the economic grounds for insuring risks and the moral or instrumental grounds for punitive damages awards. Though this similarity has not been noticed in the legal or judicial commentary on the issue, Section B implies that there exists a broad complementarity between the economic approach to the insurability issue and the moral approach to punitive liability.

Part III applies the analysis developed in Part II to address the effects of the modern expansion of punitive liability in terms of both moral justification and insurance. It shows why the expansion of liability increases demand for punitive coverage and why insurers are placed in the increasingly difficult position of seeming to provide general punitive coverage, yet at the same time litigating coverage issues with increased vigor. Finally, Part III also addresses the implications of punitive insurability for the moral reform of punitive liability.

I. THE CONFUSION IN CURRENT APPROACHES TOWARD PUNITIVE INSURABILITY

Litigation over the insurability of punitive damages typically raises one or both of two issues, though the two issues are closely
related. First, many courts have ruled that the terms of the underlying insurance contract are irrelevant because the purpose of punitive damages is to deter. Allowing a tortfeasor to purchase insurance for potential punitive damages would diminish the deterrent effect of punitive liability in violation of public policy. Indeed, according to the strongest version of this principle, even if the insurance policy were expressly to provide punitive coverage, the court would refuse enforcement.\textsuperscript{1} Other courts, however, have contested one or both of the underlying premises of this approach. Some courts have held that punitive liability has a function beyond deterrence that would not be impaired by allowing insurance coverage. Other courts have denied that insurance coverage of punitives would significantly diminish the deterrent effect.

The second issue in punitive insurability cases arises only after courts have gone beyond the purist deterrent obstacle. Once a court is willing to entertain punitive insurability, it still must examine the underlying insurance policy to determine whether its terms provide for coverage. That is, the next issue is whether the policy extended or excluded coverage of losses from the acts generating punitive liability. Of course, the interpretation of insurance policies often is closely related to courts' views of the deterrence function. Courts commonly will attempt to construe insurance contracts in a manner that does least harm to the deterrent objective of punitive liability.

\textbf{A. Insurance Coverage and the Deterrent Objective of Punitive Liability}

The judicial approach toward the insurability of punitive damages has been tremendously influenced by Judge Wisdom's powerful 1962 opinion in \textit{Northwestern National Casualty Co. v. McNulty},\textsuperscript{2} denying insurance coverage. According to Judge Wis-


\textsuperscript{2} 307 F.2d 432 (5th Cir. 1962), \textit{superseded in part by Va. Code Ann. § 38.2-2271 (1986)}. Because the federal court was interpreting Florida and Virginia law, the decision still accurately states Florida's public policy of prohibiting insurance coverage of punitive damages. However, as explained by the Supreme Court of Virginia in \textit{United Servs. Auto. Ass'n v. Webb, 235 Va. 655, 369 S.E.2d 196 (1988)}, the Virginia legislature enacted legislation providing that it is not against the public policy of the state to provide insurance coverage.
dom, the clear purpose of a punitive damages award is to punish and deter. This objective mandates that "damages rest ultimately as well [as] nominally on the party actually responsible for the wrong." The court further reasoned that "[i]f that person were permitted to shift the burden to an insurance company, punitive damages would serve no useful purpose." The court explained that allowing coverage, more realistically, would place the burden of the punitive levy not on the insurer, but on the insured public as a whole, since the punitive amount would be passed along to insureds in larger premiums. But this result should not occur, stated the court, because, in effect, society would be punishing itself for the wrong committed by the insured.

Judge Wisdom's opinion remains in force. It has been followed by many jurisdictions and has defined the debate for all others. The more modern approach to the insurability issue, however, has been extension of insurance coverage to punitive damages by de-emphasizing concerns over the deterrent effect of punitives that were central to McNulty. Some courts have bluntly denied that disallowing insurance coverage of punitive damages would deter particularly egregious acts. One court ruled that coverage of punitive damages so long as they do not arise from intentional conduct, thus superseding McNulty in Virginia. Id. at 657, 369 S.E.2d at 197.

3. McNulty, 307 F.2d at 440. The court emphasized that these reasons were especially strong in the context, as in this case, of drunken drivers: "[S]ocially irresponsible automobile drivers ought not escape the element of personal punishment in punitive damages when they are guilty of reckless slaughter or maiming on the highway." Id. at 441.

4. Id. at 440.

5. Id. at 442.

6. Id. The court also believed that allowing coverage of punitive damages would create a conflict of interest between the insurer and insured in settlement negotiations, and that there would be difficulties in assessing punitive damages with reference to the financial standing of the defendant without referring to the defendant's insurance coverage. Id. at 441.

7. Judge Wisdom was interpreting the law of Florida and Virginia. For a more recent Florida decision following McNulty, see U.S. Concrete Pipe Co. v. Bould, 437 So. 2d 1061 (Fla. 1983). But see United Servs. Auto. Ass'n v. Webb, 235 Va. 655, 369 S.E.2d 196 (1988) (in Virginia, McNulty has been superseded by Va. Code Ann. § 38.2-227 (1986), which allows insurance coverage of punitive damages assessed as a result of negligence, including willful and wanton conduct, but not of punitive damages assessed for intentional conduct).


should be allowed because there is no firm evidence of greater deterrence in states disallowing coverage.10

Still other courts have rationalized that allowing insurance coverage of punitive liability does not remove the deterrent effect of punitive damages. In one case, the court emphasized that the insurer had probably taken the prospect of punitive liability into account in the initial determination of the premium.11 Slightly differently, in Brown v. Maxey,12 the Supreme Court of Wisconsin asserted that an insured must still fear that the insurer will charge a higher premium after paying a punitive award or that the punitive award may exceed the insurance policy limits.13 In a similar recent decision, Baker v. Armstrong,14 the New Mexico Supreme Court emphasized the prospect of higher premiums and policy cancellation as strong deterrents to bad driving notwithstanding coverage of punitive liability.15 In still another case involving punitive liability for a criminal act, the court regarded the separate criminal sanctions as a sufficient deterrent, thus allowing punitive coverage.16

Other courts have suppressed the deterrence concern on other grounds. Some courts, for example, have stated that because insurance in some contexts is compulsory—for example, auto insurance—it would be inconsistent to deny insurance coverage for any auto-related liability.17 Other courts have interpreted the purpose of punitive awards to be, at least in part, compensatory, trumping the deterrence concern.18

12. 124 Wis. 2d 426, 369 N.W.2d 677 (1985).
13. Brown, 124 Wis. at 446-47, 369 N.W.2d at 668.
B. Punitives and the Meaning of the Exclusion of Intentional Acts

Once a court has rejected the summary refusal of punitive coverage on public policy grounds, it must interpret the terms of the underlying liability insurance policy to determine whether it provides coverage of punitive liability. In every case that I have seen, the underlying policy provides in language either identical or closely similar, that the insurer agrees to pay on behalf of the insured the ultimate net loss ... which the insured shall become legally obligated to pay as damages by reason of the liability imposed upon the insured by law ... because of (a) Personal Injury [sometimes “Bodily Injury”] ... caused by or arising out of an occurrence [sometimes “occurrence or accident”].

The policy defines “ultimate net loss” as the total sum which the insured ... become[s] obligated to pay by reason of Personal Injury.

And the policy defines “occurrence” (or “accident”) as an event ... which results in Personal Injury ... neither expected nor intended from the standpoint of the insured.19

The legal issue is whether these provisions of the insurance contract compel the insurer to compensate the insured for punitive liability.

Two issues have arisen most commonly in insurance contract construction. Some cases have involved the legal question of whether punitive liability implies something other than liability for personal injury. This issue, however, has been resolved most commonly in favor of coverage on the basis of the reference in the definition of “ultimate net loss” to coverage of “the total sum which the insured ... become[s] obligated to pay.”20 The far more common context of the insurability issue is litigation over the meaning of the term “occurrence” or “accident.” Typically, the insurer maintains that the act of the insured generating punitive liability should be regarded as, in some sense, intentional, thus ex-

cluded from insurance coverage because the term "occurrence" or "accident" is defined in the policy as an act not intended by the insured.

The source of confusion in this area of law derives from the lack of correspondence between these terms of the underlying insurance policy and the legal standards for the award of punitive damages. Of course, there is less controversy where punitive damages are awarded expressly because the tortious act is "intentional." The confusion arises instead because courts impose punitive liability in many other contexts as well, such as where the tortious act is "willful and wanton," "reckless," "grossly negligent," "outrageous," or "shocking to the conscience." The legal issue in these cases is whether acts defined according to these terms are or are not excluded from coverage by the "neither expected nor intended" policy clause.

There seems very wide agreement that criminal fines are not insurable, both on grounds of the deterrence policy and because criminal fines commonly require evidence of intent. Many courts also have held that, where punitive civil liability derives from an underlying criminal act, insurance coverage will be denied, largely for the same reasons. As described above, however, growing modern disbelief in the deterrent effect of prohibiting insurance coverage of punitive liability has undermined this approach. Indeed, in one case, a court held punitive liability to be insurable because the underlying act was criminal. According to the court, allowing the tortfeasor to pass on punitive liability to the insurer would not diminish deterrence of similar acts causing harm because of the prospect of separate criminal punishment.

It is a well-known feature of modern insurance coverage jurisprudence that courts are increasingly interpreting insurance

26. See supra text accompanying notes 9-16.
policies in ways that expand insurance coverage. These developments have affected the punitive liability issue as well. Some courts in punitive cases have adopted as an interpretive technique the policy of construing the contract against the interests of the insurer because the insurer drafted the policy. Other courts have adopted the more modest technique of interpreting only contractual ambiguities against the interests of the insurer. Still other courts have explicitly interpreted the contract with the purpose of allowing rather than denying coverage. Each of these interpretive techniques expands the insurability of punitive damages.

Though many courts have held that acts which are clearly intentional are excluded by the "neither expected nor intended" clause, other courts have given greater consideration to the meaning of the term "intent." One court, for example, recently held that coverage was allowed for punitive damages resulting from liability to a woman who had intentionally cut the brake cable on her automobile. According to the court, the woman may have intended suicide, but she did not intend necessarily to cause damage to others; thus, the harm was accidental rather than intentional. Other courts have distinguished between intentional acts and intentional injuries, allowing coverage though the act was intentional, because the specific injury was accidental. Another modern court denied liability coverage of a woman who intentionally collided with another vehicle, but allowed collision coverage under the policy of the woman's husband who, apparently, had not shared his wife's intent.

There has been similar disagreement among courts over the applicability of the "neither expected nor intended" clause to ac-
tions that fall short of clear intention. Thus, some courts have denied coverage of punitive liability stemming from willful and wanton behavior;\textsuperscript{37} other courts, however, have emphasized that the willful and wanton standard falls short of clear intent, and these courts have allowed punitive coverage.\textsuperscript{38} Similarly, some courts have denied coverage of punitives deriving from grossly negligent acts;\textsuperscript{39} yet, other courts have allowed coverage of the grossly negligent.\textsuperscript{40} Some courts have denied coverage of punitive liability from reckless or grossly reckless behavior;\textsuperscript{41} however, most have allowed coverage on the ground that there is often small difference between reckless behavior and negligent behavior, which is clearly insurable.\textsuperscript{42} These different interpretations stem largely from the modern trend toward expansive coverage interpretation as well as the modern undermining of the deterrence policy.\textsuperscript{43} As a general matter, decisions prior to the last decade are much more likely to deny insurance coverage of punitive damages; modern decisions, to allow coverage.

Indeed, the tendency to allow coverage of punitive damages characteristic of the modern approach has accelerated in recent years; courts have emphasized, as an independent ground for allowing coverage in the cases before them, the empirical fact that increasing numbers of rulings allow coverage. Thus, several courts have adopted the bootstrapping logic that coverage of punitives should be allowed because insureds increasingly expect that courts


\textsuperscript{38} \textit{E.g.}, Scott v. Instant Parking, Inc., 105 Ill. App. 2d 133, 245 N.E.2d 124 (1969). One court, in fact, argued that the willful and wanton character of an act is a factor that directly gravitates in favor of coverage. Pennsylvania Threshermen & Farmers’ Mut. Casualty Ins. Co. v. Thornton, 244 F.2d 823 (4th Cir. 1957). The insurer had asserted that, because the underlying action was characterized as “willful and wanton” to merit punitive damages, it should be regarded as intentional, rather than accidental. The court responded, however, that to accept this argument “would lead to the illogical and indefensible result, contrary to the purpose and spirit of liability insurance policies[...]: the more extreme the recklessness, the more likely the insurer would be to escape liability.” \textit{Id.} at 827.

\textsuperscript{39} \textit{E.g.}, American Sur. Co. v. Gold, 375 F.2d 523 (10th Cir. 1966).

\textsuperscript{40} \textit{See, e.g.}, Harrell v. Travelers Indem. Co., 279 Or. 199, 567 P.2d 1013 (1977).


\textsuperscript{42} \textit{E.g.}, Thornton, 244 F.2d at 827.

\textsuperscript{43} Discussed \textit{supra} text accompanying notes 9-16.
will allow insurance coverage of punitive liability. Other courts, taking only a more general version of the same approach, have concluded that, since punitive damages themselves are being awarded so frequently in modern times, it is increasingly difficult to distinguish between negligent and punitive liability for purposes of insurance coverage.

Finally, still other courts have emphasized the apparent inconsistency of insurers themselves on the insurability issue as a reason to allow coverage of punitives. Presumably, insurers are highly interested in deterring harmful activities for which they may have to pay. But many courts in recent years have emphasized that, despite the long history of litigation over the meaning of the "neither expected nor intended" clause, insurers have not acted to clarify the issue by expressly excluding punitive liability from coverage. The failure of insurers to exclude punitive coverage, thus, has served importantly to expand punitive coverage.

II. THE OPERATION OF INSURANCE, THE DETERMINATION OF INSURABLE RISKS, AND THE DEFINITION OF PUNITIVE LIABILITY

Understanding how insurance markets operate is central to the evaluation of the effects of allowing or disallowing insurance coverage of punitive damages. Section A explains how insurance reduces risks. It shows that there are clear constraints on the nature of risks that can be effectively reduced through the insurance mechanism. Section A describes why some risks are uninsurable and, thus, why insurers exclude coverage of certain risks in the basic liability policy. As we shall see, the exclusion of coverage of such risks expands insurance availability by allowing the offering of broader basic insurance coverage or basic coverage at a lower premium, making insurance available to individuals or firms that might not purchase insurance otherwise. Since all humane societies aspire to maximize compensation of the injured, there is substan-

tial social benefit derived from increasing the prospects for broad compensation insurance for victims of tortious activity.

Section B, then, discusses the definition of punitive liability. It shows the close similarity between the economic definition of risks excluded from insurance coverage and the moral definition of acts subjected to punitive damages.

A. How Insurance Can Reduce (Certain) Risks

Insurance operates where losses have some stochastic or probabilistic character. A loss that is certain to occur in some particular period cannot be insured against; one can only accumulate savings before the loss occurs or after the loss is suffered to restore the previous economic position. In contrast, for insurance to reduce the risk level, the insured losses must be probabilistic, either as to whether the losses will occur at all (for example, whether a product will prove defective) or as to when losses certain to occur actually will occur (for example, whether one will die before or after full life expectancy).

For a loss or a set of losses to be probabilistic means that the occurrence of the loss or set can be described by a probability distribution. The mean of the distribution represents the most likely probability of occurrence of the loss; the distribution or error term surrounding the mean represents the greater or lesser likelihood that the loss or set of losses will occur. The expected cost of the loss is determined by summing the amount of the loss weighted by these probabilities. Obviously, whatever the mean expected magnitude of the loss, the broader the probability distribution around the mean, the greater the total expected cost. More precisely, given some mean, expected cost is determined by the variance of the distribution around the mean. The variance of the distribution measures the risk associated with the loss.47

Insurance can reduce the risk of losses by aggregating uncorrelated losses. To the degree that losses are uncorrelated (that is,

47. The variance of a distribution is the sum of the squares of the differences between the mean of the distribution and each random variable:

\[ \text{variance} = \sigma^2 = \sum (X - \bar{x})^2, \text{ where } \bar{x} = \text{the mean of the distribution and } X, \text{ each random variable.} \]
statistically independent), aggregation will reduce variance by leading the error terms of the risks to cancel out. Aggregation does not change the extent of underlying loss. But the cumulative risk of loss, measured by the variance of the distribution, can be reduced by aggregation, again, to the extent the individual risks are statistically independent. For statistically independent risks, the sum of the aggregated risks is less than the sum of the risks taken individually.

This risk-reducing function derives from operation of the law of large numbers—the empirical phenomenon according to which the probability density function of a loss tends to become concentrated around the mean as the sample number increases. The law of large numbers implies that as one increases the number of insured persons possessing independent and identically valued risks, one increases the accuracy of prediction of the risk generated by each individual. The increase in predictive accuracy derives from the reduction in the variance of risk of expected outcomes.

It is important here to distinguish between employing a large population of insureds to shift losses and employing a large population of insureds to reduce the risk level by cancelling out risk terms. Many scholars refer to insurance as loss spreading, but there is an important difference between simply spreading losses and reducing risks. The difference is that, to the extent the losses and accompanying risks are truly independent, their aggregation not only spreads them, diminishing the impact of a loss on an individual insured, it also reduces the total risk level of the pool below the preaggregated sum of individual risks. This reduction of the total risk level is the result of the operation of the law of large numbers on statistically independent risks: it increases the ability to predict the risk level (which is what is meant by cancelling out risk terms).

Loss spreading, in contrast, serves only a distributional end. Spreading does not change the risk level; it merely distributes existing risks across a set of the population different from the set that suffered the risks in the first instance. The law of large numbers will not apply if the risks faced by members of the pool are

48. Losses are statistically independent if the occurrence of one loss does not affect the probability of occurrence of the other.
not statistically independent to some degree. Aggregating such risks would be unproductive because the reserves an insurer would have to maintain would equal or, perhaps, exceed the reserves individuals would have to maintain if they were uninsured.50

By reducing the risk level, effective risk aggregation reduces the premium necessary to insure a given risk. This can mean that fuller insurance coverage (larger insurance benefits) can be offered for the same dollar premium. Or it can mean that insurance can be made more broadly available for risks that would otherwise be uninsured. In contrast, loss spreading, because it does not change the risk level, cannot directly achieve any of these benefits.

A closely related method of reducing risks is by risk segregation: the insurer attempts to distinguish relatively high-risk from low-risk insureds and then to assign them to narrowly defined risk pools. In the insurance industry, risk pool definition is referred to as insurance underwriting.

Precise risk pool definition extends insurance availability by controlling adverse selection. Adverse selection is a problem central to every insurance context. An insurer must collect into a risk pool individuals with a sufficiently narrow range of exposure to risk for the insurance to remain financially attractive to each member of the pool. Since insurance premiums must be set according to the average level of risk brought to the pool, the wider the range between high-risk and low-risk pool members, the greater the difference between the average risk and the risk of the low-risk members. If the disparity between the premium and the risks added by low-risk members becomes too substantial, low-risk members will drop out of the pool because they find alternative means of protection cheaper than market insurance. At the extreme, as low-risk members drop out, the insurance pool will unravel.51 One of the most important reasons that some risks are uninsurable is that insurers are unable to narrow the assortment of risks within a risk pool. Those insurers who are better at identification and segregation can offer lower premiums to low-risk insureds and can thus extend insurance availability.

It is well established that the more precisely insurers can seg­
regate risks by insurance discrimination of this nature, the more
broadly insurance can be offered in the society. A court that
wanted to maximize insurance availability in the society would
adopt policies that encouraged maximally effective discrimination
in order to segregate risks into the narrowest possible pools. The
difference between the view of insurance as loss-spreading and in­
surance as risk-reducing becomes quite sharp at this point. Loss
spreading is often defended solely on distributional grounds:
Spreading shifts the costs of bearing losses away from parties that
generate them to parties viewed as better able to bear them—for
example, assigned risk pools in auto or medical malpractice insur­
ance that shift the costs generated by more risky drivers or doctors
to the less risky. Loss shifting of this nature can only be achieved
by compulsion and can only be defended on grounds of moral pref­
erence. But the implications of loss shifting of this nature are
often neglected. By defeating the risk-reduction benefits of segre­
gation, this form of loss shifting increases the risk level, increases
the underlying injury level, and reduces insurance availability. The
gains to the subsidized high-risk insureds are paid for in increased
risk, increased injury, and less available insurance, losses often ig­
nored in the emphasis on simple loss spreading.

The aggregation and segregation functions of insurance, then,
are similar in both method and effect: Both serve to increase pre­
dictive accuracy in order to reduce the risk level and the effective
costs of injuries. Reducing injury costs, however, necessarily gener­
ates offsetting effects. Where expected injury costs are lower, the
underlying level of activity and the underlying injury rate will in­
crease, a phenomenon known as moral hazard. All insurance
regimes generate this effect. Moral hazard increases the costs of

52. See Rothschild & Stiglitz, *Equilibrium in Competitive Insurance Markets: An Es­
say on the Economics of Imperfect Information*, 90 Q.J. Econ. 629 (1976); Crocker & Snow,
Econ. 321 (1986).

53. See G. Calabresi, *Costs of Accidents: A Legal and Economic Analysis* 36-54
(1970), for many other examples.

54. Dean Calabresi openly admits this point. Id. at 278-85.

55. The costs of the increase in injuries because of insurance will be less than the
reduction in effective costs achieved by insurance; otherwise, there will be no effective de­
mand for insurance.
injuries and, thus, increases the risk level. Insurers will attempt to control moral hazard by the definition of insurance coverage.

Insurers will constrain or, at the limit, exclude coverage of losses particularly susceptible to insured moral hazard. The omnipresent exclusion in life insurance policies of coverage for death by suicide is an obvious example. The exclusion serves to control moral hazard by removing the incentive that providing large monetary amounts to beneficiaries would add to other forces compelling the act. 56 Less dramatically, the exclusion in consumer product warranties of coverage of easily broken glass parts or the easily marred product finish, or the exclusion in auto warranties of coverage of engine damage from racing or towing heavy loads serves a similar function. 57 These exclusions place the burden of loss on the insured itself, increasing insured preventive efforts and, at the same time, culling out (segregating) high-risk insureds relatively more susceptible to such losses. 59

The beneficial effects of coverage exclusions as a method of controlling moral hazard have been largely neglected. Because these insurance provisions directly allocate losses to the insured, their existence might seem antithetical to broad loss spreading. But here again the limitations of the loss spreading metaphor become clear. A coverage exclusion admittedly places all of some identified loss on the insured. To this extent, it constrains loss spreading. But as such provisions lower total insurance costs, they allow the extension of basic insurance benefits more broadly to the society. Thus, for example, the exclusion of life insurance coverage

56. Today, life insurance policies typically exclude coverage for death by suicide only for the first two policy years. It is not clear that the two-year limitation on the exclusion derives from insurer judgment that two years is a sufficient margin to control moral hazard. More probably, the limitation derives from direct regulatory pressure or indirect judicial pressure manifesting the desire to allow coverage whenever possible and the consequent refusal to enforce exclusions deemed unreasonable. In 1987, only 1.4% of deaths of life insurance policyholders were by suicide. AMERICAN COUNCIL OF LIFE INSURANCE, 1988 LIFE INSURANCE FACT BOOK 100. It is implausible that 98.6 percent of life insureds would voluntarily wish to purchase coverage for the purpose of covering death by suicide after the second policy year.


58. Direct risk monitoring by insurers (such as requiring the installation of specific safety devices and inspecting for compliance) is very similar to the control of moral hazard by exclusions.

to suicides allows the premium for basic life insurance to be lower and makes it possible for some individuals who would not or could not purchase life insurance at a higher premium to obtain basic life insurance protection. Thus, paradoxically, denying insurance coverage to particular high risks through coverage exclusions can maximize the availability of basic insurance coverage to the society.

With this introduction, it is possible to evaluate with greater clarity the purpose of the exclusion from insurance coverage of occurrences expected or intended by the insured, the exclusion that figures centrally in the punitive insurability cases. As described above, insurance is only possible where the underlying risks are probabilistic and where the risks are not susceptible to insured moral hazard.

The coverage exclusion of occurrences expected by the insured represents an obvious effort to constrain insurance to probabilistic risks. If the insured knows or expects that a particular occurrence will happen, the loss caused by the occurrence cannot be said to be probabilistic and, thus, cannot be effectively insured.

The exclusion of occurrences intended by the insured represents a direct effort to control insured moral hazard. The point is obvious. No one would contest that it is in the interest of an insurer (as well as of the society) for the insurer to encourage all feasible precautions to reduce the likelihood of unintended harm. Insurers will commonly compel insureds to install sprinkling systems, for example, or engage in periodic safety inspections to reduce the accident rate. If these insurer efforts reduce the risk level and increase insurance availability, insurer efforts to control intentional harm-causing activities, a fortiori, will have similar effect.

Put in economic terms, activities of insureds are differentially susceptible to marginal incentives from insurance. Where the existence of insurance significantly reduces marginal incentives to reduce harm, the harms become uninsurable. An act for which the occurrence is subject to the volition of the insured is highly susceptible to insurance incentives: the existence of insurance directly reduces the costs of such acts and makes them uninsurable. To give a concrete example, potential life insurance benefits obviously increase incentives for suicide. As a consequence, life insurers must
exclude coverage of suicide for some period. Similarly, to the extent that insurance provides coverage of losses intentionally caused to neighbors, the level of community tolerance is likely to decline.

Put more generally, the definition of insurance coverage can be described as an effort of the insurer to provide the broadest level of insurance to some dominant set of the insured population. In terms of risk proclivity, the population of insureds can be seen as arrayed along a continuum, from the lowest to the highest risk. The level of insurance coverage that can be made available will depend upon differences in risk proclivity and the costs associated with the underlying risks. Insurers will maximize insurance by defining coverage so that it best meets the insurance needs of the dominant set of insureds. It follows necessarily that risks which are not insurable either because they are not probabilistic in nature or because they are highly susceptible to insured moral hazard will be excluded from coverage. Such exclusions place the costs of excluded activities upon the individuals controlling the acts and thus reduce the extent of underlying harm.

The exclusion of coverage of losses intended by the insured thus benefits two classes of individuals. The first class is the broader set of insureds not intentionally engaging in acts causing harm. These insureds gain because the costs of intentional harm-causing activities will not be averaged into the premiums they pay. In this sense, it might be said that the dominant set of insureds expresses market demand for the exclusion of intentional acts. The second set of beneficiaries is those that might suffer loss if insurance for intentional acts were available. That is, the number of intentional harm-causing actions and the extent of harm intentionally caused would be higher if insurance coverage were available than if coverage were excluded. As a consequence, some set of potential victims benefits from the insurance exclusion.

60. See supra note 56.

B. Uninsurable Risks and the Normative Grounds for Punitive Awards

There is a close similarity between the definition of losses excluded from insurance coverage and the definition of actions subject to punitive liability. Our society characterizes those acts appropriate for punitive liability variously: acts that are reckless, grossly negligent, willful and wanton, outrageous, morally shocking. These various terms may be more or less descriptive in specific factual contexts. Each of these terms suggests, however, that those actions subject to punitive liability stand apart by a substantial distance from the range of activities normal for, or common to, the larger mass of citizens.

In each of these two processes of definition, the activities of the society may be seen as arrayed upon a continuum. In the insurance context, risks that are extreme, grossly different from the risks faced by the dominant set of insurance purchasers, must be excluded from coverage. The coverage exclusion compels those engaging in the harm-causing acts to pay their own way, rather than averaging the much higher costs they generate into the premiums paid by the larger population of insureds.

The definition of acts made subject to punitive liability is closely similar, though the reference in this context is to a continuum of moral justification rather than to a continuum of risks. Our society punishes by means of punitive liability actions that are morally extreme, grossly different on a moral plane from the actions of the dominant set of citizens. Punitive liability is a form of moral exclusion of behavior. Punitive liability compels those engaging in such acts to bear themselves the moral burden of their actions. Punitive liability distinguishes certain types of behavior from the behavior that is average or normal to the citizenry. Rather than embracing such behavior as within the range of the morally acceptable, punitive liability segregates and excludes such behavior from the acceptable moral pool.

The similarity between the continuum of risks and the continuum of moral justification suggests the potential for resolution of the issue of the insurability of punitive damages. It is equally in the public interest to exclude coverage of losses generated at the extreme of the risk continuum as it is to punish with punitive damages actions at the extreme of the moral continuum. The key
to a resolution derives from comparing the set of actions that cannot be insured and thus must be excluded from coverage with the set of actions that the society condemns with punitive liability.

Currently, neither the terms of existing insurance policies nor the legal definitions constituting the grounds for imposition of punitive liability provide much assistance in making this comparison. The exclusion of acts "intended by the insured" is vague and does not provide guidance regarding the point along the risk continuum at which the actions of an insured become uninsurable. Similarly, the legal terms describing behavior subject to punitive liability—"willful and wanton," "grossly negligent," "shocking to the conscience"—though evocative, do not define clearly the boundary along the moral continuum dividing the merely tortious from actions so morally extreme as to be punished punitively. The increasing judicial difficulty in distinguishing acts subject to punitive liability from the merely negligent is a confession of the point.62

The next Part attempts to employ the understanding of insurance to clarify this relationship. It describes how the reference to acts "intended by the insured" might be defined to focus more clearly on truly uninsurable risks. Current judicial interpretations of the clause have drawn largely from concepts of intent characteristic of other legal fields, such as the criminal law, rather than from the role of intent in insurance. As we shall see, there are many ways for courts to redefine the policy clause to enhance insurance availability.

Part III also addresses the relationship between actions that are uninsurable and actions so morally extreme as to justify punitive liability. Part III argues that, although the moral definition of acts deserving punitive liability is likely to extend beyond the economic definition of uninsurable acts, the insurance definition sets a lower bound for punitive liability. It is an implication of this approach that to the extent the modern expansion of punitive liability makes punitive damages insurable, it violates the moral justification for the punitive levy.

62. See supra text accompanying notes 37-43.
III. Defining “Actions Intended by the Insured” to Enhance Insurance Availability

This Part seeks to apply insurance principles to the interpretation of the policy exclusion of “actions intended by the insured.” The objective here is straightforward. Part II showed that the exclusion of risks substantially different from risks faced by the dominant population of insureds serves both to enhance insurance availability and to reduce the accident rate. Excluding coverage of extreme risks lowers the premium for basic insurance coverage or allows the extension of basic benefits. Both effects better achieve the broader social goal of providing greater compensation for the injured. In addition, however, the exclusion of coverage of extreme risks will affect the level of activities generating such risks, reducing the underlying accident rate. Thus, the exclusion of coverage of extreme risks simultaneously achieves two unambiguously important policy goals: extending basic compensation and reducing the extent of underlying injury.

As we shall see, current judicial interpretations of the clause are vastly different from interpretations that would achieve these effects. Because litigation involving the clause implicates the criminal-like imposition of punitive damages, courts have most typically defined “actions intended by the insured” by reference to criminal law concepts of intent. There is no reason to believe, however, that concepts of intent drawn from the criminal law will correspond to concepts of intent appropriate for enhancing insurance availability because the ambitions and concerns implicit in these legal fields are so different.

At the minimum, it is clear that insurance availability could be enhanced and the injury rate reduced if “actions intended by the insured” were defined with reference to the prerequisites of insurability. A minimally sufficient definition is that an action should be regarded as “intended by the insured” if the loss from the action is uninsurable either because the loss is not probabilistic or because it is highly susceptible to insured moral hazard. At base, this definition identifies those actions for which insurance cannot reduce risks. As shown in Part II, insurance is only effective as a risk-reducing mechanism where losses are probabilistic and not

63. Note that there is a substantial overlap between these definitions: Actions highly susceptible to moral hazard by the insured are not truly probabilistic.
vulnerable to insured moral hazard. It follows that to interpret such actions to be within the scope of insurance coverage would unambiguously reduce insurance availability because the addition to the premium for basic coverage would preclude insurance for some set of marginal insurance purchasers. It also follows that to allow coverage of such actions would increase insured moral hazard, thus increasing the underlying injury rate.

The current judicial approach to the interpretation of the "actions intended by the insured" clause contradicts these insurance principles. As described earlier, courts today struggle over the relationship between the standards for the award of punitive damages—"willful and wanton," "grossly negligent," "reckless," "shocking to the conscience"—and the concept of intent. The substantial intelligence of the judiciary has enabled many fine distinctions between these terms to be drawn.

According to the insurance interpretation of the "actions intended by the insured" clause, however, the inquiry is different. The question is not whether behavior that a jury has found reckless or willful and wanton is equivalent to an action intended by the insured, but whether the action justifying punitive liability has some probabilistic character or can be distinguished in some way from the moral hazard of the insured. Put differently, to achieve the dual objectives of enhancing insurance availability and reducing the injury rate, a court must ask whether the action is one that the dominant population of insureds would wish excluded from coverage because its risk of loss is so extreme. Is the risk of loss one that the dominant population of insureds faces with some probabilistic frequency? Thus, is the risk one for which insurance might reduce risks? Or is the risk uninsurable because it is not probabilistic and not invulnerable to insured moral hazard?

The insurance interpretation of the "actions intended by the insured" clause is substantially different from the current judicial approach to the problem. Take, for example, the various cases described in Part I involving suicides. There are very strong reasons to believe that the dominant population of insureds would wish to exclude coverage of losses from suicide or attempted suicide: First, such losses are not probabilistic. Were there insurance, suicide would be quintessentially an action involving insured moral haz-
ard. Second, the voluntary market almost universally excludes such risks. 64

There are many cases involving suicide or attempted suicide, however, in which victims injured in a suicide attempt have recovered punitive damages and in which courts have interpreted the underlying policy to provide coverage for punitive liability. For example, in *Farm & City Insurance Co. v. Potter*,65 a person cut the brake cable of her car and drove off, attempting suicide. Her subsequent multiple collisions injured others and generated punitive liability. The court held that her auto policy covered punitive damages because her intent only extended to her own suicide and because she had attempted, given the situation in which she had placed herself, to avoid injuries to others. 66

An insurance interpretation, however, would suggest the opposite outcome. The actions of the woman were not probabilistic, and they could directly be described as insured moral hazard. The availability of auto insurance could be expanded if the higher loss costs added to the risk pool by such individuals were culled out by means of a coverage exclusion. In addition, suicides, as a general matter, are not indifferent to the financial implications of their acts on dependents and beneficiaries, as the exclusion of coverage of suicide in life insurance policies attests. Denying insurance coverage of losses related to suicides, thus, might well reduce the accident rate, if only by shifting the choice of suicide method to a means less risky to others.

Many other current interpretations of the “actions intended by the insured” clause would also compel revision under the insurance approach. As described in Part I, much of the dispute over insurance coverage of punitives derives from differing judicial estimates of the deterrent value of disallowing coverage of punitives. The current trend toward allowing punitive coverage builds on modern skepticism of any deterrent effect of punitives at all or of any independent deterrent effect. For example, in *Price v. Hartford Accident & Indemnity Co.*,67 a court allowed coverage of punitive damages stemming from a drag racing incident in which the son of an insured driver had been found grossly negligent, wan-

---

64. See supra note 56.
65. 330 N.W.2d 263 (Iowa 1983).
66. Potter, 330 N.W.2d at 266.
ton and reckless and, moreover, had had a history of reckless driving of which the parent was well aware. According to the court, there was no important public interest in denying coverage of the punitive liability, since the boy was still subject to criminal penalties including possible loss of license and compulsory attendance at traffic school, and because the boy’s auto insurance rates were likely to increase. 68

According to the insurance approach to the interpretation of the “actions intended” clause, however, all of these considerations are largely irrelevant. The description of the boy’s actions as grossly negligent, wanton, or reckless is largely irrelevant. Alternative criminal penalties are largely irrelevant. 69 The insurance issue is whether the actions of the boy were sufficiently different from the actions of the dominant set of insureds to justify exclusion or, in terms of insurance principles, whether his actions either had some probabilistic character or evidenced moral hazard.

The insurance approach, without difficulty, would deny insurance coverage of the punitive liability. The boy was not a driver typical of the dominant population of insureds who, by some probabilistic chance, found himself in a drag race. Drag racing is moral hazard. The dominant population of insureds does not want insurance coverage of liability from drag racing, but is likely to strongly prefer exclusion of such liability (if it were feasible in a standard auto policy) to reduce the premium for basic coverage. 70

Indeed, the parent’s knowledge of the boy’s previous recklessness dramatically demonstrates the potential beneficial effects of denying coverage. Where parents become aware of the financial risks of allowing uncontrolled children to drive their cars, the injury rate attributable to young drivers is likely to decline.

Of course, it is possible to give many other examples of how the “actions intended by the insured” clause would be defined differently according to principles of insurance than it is currently interpreted by the courts. In truth, the opinions in all modern cases would be different because, to date, no court has defined a coherent insurance approach to the issue. Even courts denying coverage of punitives on deterrence grounds are innocent of insurance

69. Of course, criminal liability is highly suggestive of insured moral hazard.
70. Indeed, the standard automobile warranty excludes coverage of engine losses from drag racing. See Priest, Product Warranty, supra note 57.
reasons for denial. The insurance reasons are stronger than the deterrence reasons, since we do not have supporting evidence of a deterrent effect.

The puzzle remains, however, why insurers have not modified policies to exclude coverage of punitive liability despite invitations to do so by the courts. The question defined more precisely, however, is why insurers have not modified policies within the last two decades as punitive liability has expanded. Prior to the 1970s, the imposition of punitive liability was infrequent and courts were much more likely to interpret the “actions intended” clause to exclude coverage.

The basic principles of insurance provide an answer. To the extent that the scope of punitive liability is extended to a broader set of underlying actions, the likelihood increases that actions generating punitive liability will have some probabilistic character or will be unrelated to insured moral hazard. If so, then the losses from these actions can be insured. To the extent actions subject to punitive liability are probabilistic, they are insurable. Perhaps curiously, the expansion of punitive liability enhances the insurability of punitive damages.

A different way of putting this point is that, as punitive liability is expanded to a broader set of underlying activities, the difference diminishes between actions typical of punitive liability and actions typical of the dominant set of insureds. As a consequence, insureds prefer an insurance policy defined to provide general coverage of punitive liability. The reluctance of insurers to amend basic policies to more precisely exclude coverage of punitive liability reflects responsiveness to consumer demand for liability insurance.

This approach explains why insurers have not responded to the judicial suggestion of a more precise exclusion, yet have continued to litigate individual coverage cases. The dominant set of insureds may prefer general coverage of punitive liability because of its probabilistic application. Yet the dominant set of insureds may still benefit by the exclusion of particularly extreme loss-causing behavior. That is, as the range of behavior generating punitive damages expands, the insurability of such damages also increases.


72. I am grateful to Judyth Pendell and John Backer of Aetna Life & Casualty for comments on this point.
liability increases, a general exclusion might reduce the level of insurance, though a more particular—perhaps, case-by-case—exclusion of especially extreme actions may still increase insurance availability. In this view, punitive insurability litigation derives from the difficulty of incorporating into the basic policy itself an exclusion more carefully aimed at extreme behavior.

This explanation, too, suggests that the current judicial approach to the issue should be reversed. Courts today commonly invoke the absence of a general exclusion of punitive liability as a ground for awarding coverage in individual cases. But the basic principles of insurance show that this frames the issue exactly backwards. The expansion of general punitive liability prevents a comprehensive exclusion of punitives. Courts, instead, could enhance insurance availability and reduce the accident rate by more vigilant exclusion of coverage in cases of especially extreme behavior.

Finally, the principles of insurance raise doubts that the expansion of punitive liability in recent years continues to achieve the moral function to which such liability aspires. We have seen that punitive liability has so expanded in modern times as to become probabilistic in nature and, thus, insurable. For it to be insurable, however, suggests that punitive liability is a realistic prospect for the dominant population of insureds.

Yet where the dominant population of insureds is subject in a systematic way to the prospect of punitive damages, the special moral force of punitive liability disappears. Punitive liability no longer represents the punishment of morally extreme behavior. Punitive liability becomes a cost, a mere input, added to other costs of productive activities. As no more than a normal cost of operation, punitive liability loses its economic justification, and it loses its moral justification.

The point can be made more strongly. It is not necessary to believe that the determinants of insurability are exactly congruent with the moral grounds for punitive punishment. But the economic grounds of insurability must surely define a lower bound of morally justifiable punishment. Losses suffered from actions that are insurable because probabilistic in character lack the requisite level of moral depravity to justify punitive liability. It follows that, to

---

73. See, e.g., cases cited supra note 46.
the extent the modern expansion has guaranteed the insurability of punitive liability, the moral justification for punitive liability has been lost.