2016

Remedies

Ariel Porat

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

Part of the Law Commons

Recommended Citation

This Working Paper is brought to you for free and open access by the Coase-Sandor Institute for Law and Economics at Chicago Unbound. It has been accepted for inclusion in Coase-Sandor Working Paper Series in Law and Economics by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
Remedies
Ariel Porat*


Abstract

This chapter emphasizes the common denominators of the remedies available to the victim in torts and contracts. Some remedies which are more typical of either contracts or torts are also discussed. The topics covered by the chapter are liability rules vs. property rules, specific performance vs. damages, the foreseeability requirement, tort liability for pure economic losses and non-pecuniary losses, caps on consequential damages, damages for bodily injury and lost income, liquidated damages, probabilistic recoveries and offsetting risks.

Introduction........................................................................................................................................... 2
I. Property Rules and Liability Rules................................................................................................ 3
   A. General Framework..................................................................................................................... 3
   B. Refinements.............................................................................................................................. 7
II. Specific Performance, Damages, and Efficient Breach .............................................................. 8
   A. General Framework..................................................................................................................... 8
   B. Refinements.............................................................................................................................. 10
   C. The Promisee’s Incentives......................................................................................................... 12
III. Scope of Liability.......................................................................................................................... 14
   A. Foreseeability ............................................................................................................................ 14
   B. Pure Economic Loss.................................................................................................................. 15
   C. Non-Pecuniary Loss .................................................................................................................. 18
   D. Caps on Consequential Damages............................................................................................. 19
IV. The Measure of Recovery........................................................................................................... 21
   A. Bodily Injury and Lost Income................................................................................................ 21
   B. Liquidated Damages.................................................................................................................. 23
V. Partial Recoveries ........................................................................................................................ 26
   A. Probabilistic Recoveries.......................................................................................................... 27
   B. Offsetting Risks....................................................................................................................... 29

* Alain Poher Professor of Law at Tel Aviv University and Fischel-Neil Distinguished Visiting Professor of Law at the University of Chicago Law School. I thank Hed Kovetz for excellent research assistance.
Introduction

Analyzing the substantive law without its remedial part is almost meaningless. If we knew that the law imposes liability for negligence or for breach of a contract, but knew nothing about the remedies to which the victim is entitled in case his rights are infringed, we would know very little about the impact of the law on the real world. Therefore, in all legal systems the remedies are interlinked with the substantive law. In civil legal systems the same code which allocates entitlements among the parties also sets the remedies for protecting those entitlements. In both civil and common law legal systems it is hard to imagine a court deciding a substantive law dispute without taking into account, explicitly or implicitly, the remedies which are available to victims. Indeed, it is hard to imagine the creation of the substantive law, either by legislatures or courts, without careful consideration of the remedial consequences of its breach.

The claim that substantive law and remedies are interlinked with one another might imply that each legal field must have its own unique remedies. If that were the case, remedies would not be an independent topic, but rather a subtopic in each and every substantive field of the law.¹ This implication, however, is wrong. Remedies in different legal fields have much in common, and the study of remedies as a topic can teach us a lot, especially when the goals of the substantive legal fields are similar (Cooter 1985). Consider tort law and contract law. Under its efficiency rationale, tort law should minimize social costs, thereby enhancing social welfare. In order to achieve this goal, tort law should provide incentives for both the injurer and the victim to take efficient precautions. Similarly, contract law should also provide the parties with efficient incentives, in order to enable them to maximize the contractual surplus. In both torts and contracts providing the injurer/promisor and the victim/promisee with efficient incentives is done through a combination of substantive and remedial law. It should therefore come as no surprise that the remedies in both legal fields share much in common.

¹ In most law schools in common law jurisdictions, an important part of the Contract Law or the Tort Law class is remedies (for breaching a contract or for wrongfully inflicting harm, respectively), and only few law schools offer a Remedies class.
This chapter emphasizes the common denominators of the remedies in torts and contracts. Some remedies which are more typical of either contracts or torts are also discussed. While the remedies in both fields are similar, they are not identical, and often are adapted to the legal context to which they apply.

I. Property Rules and Liability Rules

A. General Framework
In a seminal article, Calabresi and Melamed distinguished between the allocation of entitlements and the remedies for protecting them, as two distinct stages in promoting efficiency (Calabresi and Melamed 1972). In particular, they argued that once entitlements are allocated, they can be protected by either property or liability rules. Under a property rule, no one is allowed to deprive the owner of his entitlement without his consent; under a liability rule, other people are allowed to do so, but must compensate the owner for his losses. Thus, suppose Polluter inflicts harm on Resident. The law should allocate an entitlement, either to Resident to live without the pollution or toPolluter to pollute without interference. Assume first that the law made the former choice, so that the entitlement is allocated to Resident. Now a second choice must be made: to protect the entitlement with either a property rule or a liability rule. Under a property rule, Resident can sue Polluter in court and get an injunction, prohibiting further pollution (rule 1, in Calabresi & Melamed’s terms); under a liability rule, Resident is entitled to compensation only, so it is Polluter’s choice whether to stop polluting, or instead pollute and compensate Resident for his losses (rule 2, in Calabresi & Melamed’s terms).

Assume next that the law allocated the entitlement to Polluter rather than to Resident. Here too Polluter’s entitlement can be protected with either a property or a liability rule. Under a property rule, no one can stop Polluter from polluting without his consent (rule 3); under a liability rule, Resident can stop Polluter from polluting even without his consent, but if she chooses to do so she should compensate Polluter for the harm he suffers due to stopping the pollution (rule 4). The four rules are summarized in the table below:
<table>
<thead>
<tr>
<th>Rule 1</th>
<th>Entitlement</th>
<th>Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Resident</td>
<td>property rule</td>
</tr>
<tr>
<td>Rule 2</td>
<td>Resident</td>
<td>liability rule</td>
</tr>
<tr>
<td>Rule 3</td>
<td>Polluter</td>
<td>Property rule</td>
</tr>
<tr>
<td>Rule 4</td>
<td>Polluter</td>
<td>Liability rule</td>
</tr>
</tbody>
</table>

Calabresi & Melamed analyzed the efficiency considerations in making the choices with regard to allocating and protecting the entitlements. First they argued that when transaction costs are low it does not really matter, from an efficiency (rather than a distributional) perspective, whether the entitlement is allocated to Polluter or Resident: as the Coase Theorem teaches us (Coase 1960), with low transaction costs the parties would reach the efficient solution regardless of the initial allocation of entitlements. Thus, if Polluter is the cheapest cost-avoider and rule 1 applies he would take measures to prevent the harm (otherwise Resident would get an injunction in court, prohibiting pollution); if rule 3 applies instead, Resident would offer Polluter a payment to stop polluting and Polluter would accept the offer. The same reasoning applies to the reverse case when Resident is the cheapest cost-avoider: Resident would either take measures to avoid the harm (under rule 3) or be paid by Polluter to do the same thing (under rule 1).

Things become more complicated when transaction costs are high, which makes contracting between the parties either hard or impossible. Here it is necessary to distinguish between two scenarios: first, when the cheapest cost-avoider can be identified; and second, when it is unknown who the cheapest cost-avoider is or if there is one at all. In the first scenario, the entitlement should be allocated to the party who is *not* the cheapest cost-avoider. Such an allocation would provide incentives to the cheapest cost-avoider to take measures to prevent the harm. Thus, if Polluter is the cheapest cost-avoider, allocating the entitlement to Resident and protecting her entitlement with a property rule (rule 1) would incentivize Polluter to prevent the harm. The same logic applies to the reverse case when Resident is the cheapest cost-avoider: here rule 3 would do the work.
The second scenario is the more interesting, since it calls for liability rules. Assume that it is impossible for either courts or legislatures to know whether stopping pollution is efficient or not. Under those circumstances a property rule would not be an adequate solution, since once such a rule is applied the parties might be stuck in an inefficient situation. For example, if rule 1 is applied pollution is prohibited, so even if pollution is efficient it will be prevented. Indeed, if the court or legislature knew that pollution is efficient, they would apply rule 3 and restore efficiency. But once courts or legislatures cannot know whether pollution is efficient or not, rule 2 could solve the problem. With rule 2, Polluter must decide whether to pollute and bear the resulting harm or stop polluting. Polluter will do whatever is cheaper for him, and that would also be cheaper for society. Thus, if the harm is 100 and prevention costs are 50, Polluter will stop polluting, while if prevention costs are 150, Polluter will continue to pollute. In both cases Polluter’s interest and the societal interest align.

Rule 2 is not the only alternative for solving the problem; rule 4 could be equally effective. Under rule 4, it is Resident rather than Polluter who decides whether to stop the pollution or not. If she decides in the affirmative she will order Polluter to stop polluting and reimburse him for prevention costs, but if she decides in the negative she will do nothing and bear the harm. Thus, if the harm is 100 and prevention costs are 50, Resident will order Polluter to stop polluting, but if the costs are 150, Resident will do nothing and bear the harm. It is easy to see that the comparisons between harm and prevention costs under rules 2 and 4 are exactly the same, the only difference being the identity of the party conducting the comparison and making the decision which follows. As with Polluter under rule 2, so too with Resident under rule 4, the interest of the party making the decision aligns with the societal interest.

The choice between rules 2 and 4 has redistributive consequences (exactly like the choice between rules 1 and 3): rule 2 (and 1) favors victims (Residents), while rule 4 (and 3) favors injurers (Polluters). But even if efficiency was the only consideration which matters for the law, the choice between the two rules could make a difference. The main efficiency consideration for choosing between the two rules is the availability of information to the courts applying the rules and the risk of errors which follow. Thus, if courts have better information about the harms to victims than about prevention costs,
rule 2 will be more efficient than rule 4, and if the reverse is true, rule 4 will be the more efficient rule. To understand why, consider the preceding numerical example when harm is 100 and prevention costs are 50, and assume that rule 2 applies. Efficiency-wise, the harm should be prevented since the harm is higher than prevention costs. But suppose now that the courts underestimate Resident’s harm—say, because they are unaware of the high value Resident ascribes to her property—and set damages at 40 instead of 100. Under those circumstances, Polluter will inefficiently pollute, since prevention costs are higher than damages. Applying rule 4 could solve the problem if, but only if, the courts estimate Polluter’s prevention costs accurately enough. Thus, if under rule 4, ordering Polluter to stop polluting would trigger Resident’s liability of 50, Resident will compare his harm of 100 (the realistic assumption here is that Resident accurately estimates her own harm) with expected liability of 50, and decide to stop the pollution. In this way efficiency would be restored. In our last example, rule 2 leads to under-deterrence; with different numbers and overestimation of Resident’s harm by courts, rule 2 might lead to over-deterrence. In both cases rule 4 might sometimes—but not always—solve the problem. In other cases the reverse might be true: under- or overestimation of prevention costs by courts may result in inefficiencies under rule 4, which rule 2 might sometimes—but not always—ameliorate. In cases where courts’ errors are not solvable under either rule 2 or 4 in a satisfactory manner, the case for a liability rule becomes weaker and the case for a property rule becomes stronger. With high transaction costs, uncertainty as to whether prevention by Polluter is efficient or not, and a high risk of courts’ errors with respect to both harm and prevention costs, it is hard to choose between a property or liability rule.

In the real world, rule 4 is very rarely applied (Chang 2014). The main reason seems to be that in cases when transaction costs are high, typically victims are numerous and cooperation between them—a prerequisite for rule 4’s implementation—is often implausible due to a free-riding problem. Thus, when there are many Residents who should decide whether to stop Polluter from polluting in return for monetary payment—as rule 4 requires—each Resident might refuse to share the costs, knowing that if other

---

2 For a case where it was applied, see Spur Industries, Inc. v. Dell E. Webb Development Co., 494 P. 2d 700 (Ariz. 1972).
Residents pay, pollution will stop anyway and she will be able to reap the benefit for free. The same problem typically does not arise with rule 2: the injurer needs no one’s cooperation in order to inflict harm on the victim and afterwards compensate him for it (Porat 2009).³

B. Refinements

Calabresi & Melamed’s article has inspired many commentators who developed the original arguments made by Calabresi & Melamed, offered new applications, and also criticized some aspects of them. Due to limitations of space, only a few contributions to the literature will be mentioned below.

Lucian Bebchuk pointed out that Calabresi & Melamed offered an “ex-post analysis”; namely, an analysis that takes as a given the costs and benefits that would be generated for the parties with and without externality-producing actions (Bebchuk 2001). This analysis, claimed Bebchuk, does not capture the entire picture. The allocation of entitlements and the way in which they are protected divide values between the parties differently, and this ex-post division has a considerable impact on the parties’ ex-ante decisions. Ex-ante decisions take place before the decisions are made whether to undertake externality-producing actions and influence the parties’ potential payoffs with or without these externality-producing actions. A full account of the efficiency of any given allocation of entitlements and how they are protected—Bebchuk’s argument goes—must consider not only the ex-post analysis but also the ex-ante analysis.

A couple of other articles, one authored by Ronen Avraham and the other by Ian Ayres, dealt with the situation discussed in subsection A, in which the court applying a liability rule lacks information about the values the parties ascribe to their entitlements. Avraham and Ayres suggested sets of rules, constructed on a combination of Calabresi and Melamed’s rules and some additional rules that can potentially encourage parties to reveal their true valuations of their entitlements, thereby facilitating more efficient outcomes (Avraham 2004; Ayres 1996).

³ Sometimes, however, injurers have an interest in directing their injurious activities toward the same victim/s, since marginal harm decreases when more injurers join the existing ones. In such cases cooperation problems among injurers would emerge (Dillibary 2013).
Lastly, Barbara Luppi & Francesco Parisi addressed the question regarding how remedies should be chosen when there are asymmetric transaction costs (Luppi and Parisi 2011). They defined asymmetric transaction costs as situations in which different alternatives for reallocating resources entail different costs, such as when it is less costly to transfer an entitlement from one use to another than in the reverse direction. Luppi & Parisi consider the possibility of using mixed remedies in such cases; for example, applying a property rule when A is the owner of the entitlement and B is the potential infringer, and a liability rule if the positions of A and B are reversed.

Calabresi & Melamed’s framework is very useful in many contexts. In contract law, for example, specific performance could be characterized as a property rule, while damages as a liability rule. The next section further elaborates on this point.4

II. Specific Performance, Damages, and Efficient Breach

A. General Framework
There has been extensive debate in the legal literature over which remedy—specific performance or damages—should be the primary remedy and which the exception. For many, this debate represents a much broader dispute between law and economics and deontological scholars over the nature and goals of the law (Shiffrin 2009; Shavell 2009; Posner 2009). In the beginning, law and economics scholars argued that damages should be the primary remedy since specific performance, but not damages, discourages efficient, and therefore desirable, breaches of contract. Later, law and economics scholars developed more nuanced arguments, showing that specific performance is often the more efficient remedy.

The following two examples illustrate scenarios of efficient breach, and the discussion which follows clarifies the conditions under which a damages remedy would encourage such breaches and at the same time discourage inefficient breaches.

Example 1. Gain-Seeking Breach. Seller undertakes to manufacture a machine for Buyer. Expected costs of production are 80, the price which is paid upfront is

4 The typical application of Calabresi & Melamed’s framework is intentional infliction of harm, as in nuisance (pollution) cases. Is that framework suitable for accidental harms? Should a negligence rule be characterized as a property or a liability rule? For the argument that negligence law should be characterized as liability rules, see Porat, 2009, at pp. 199-200; for the counterargument, see Coleman & Kraus 1986; Zipursky 1998, pp. 55-70.
90, and the value of the machine to Buyer is 100. After the contract is concluded, a second buyer shows up offering Seller 110 for the same machine. Seller can produce only one machine at a time, so he breaches the contract with Buyer and sells the machine to the second buyer. The only loss Buyer suffers is the machine’s lost value.

Example 2. Loss-Avoiding Breach. Seller undertakes to manufacture a machine for Buyer. Expected costs of production are 80, the price which is paid upfront is 90, and the value of the machine to Buyer is 100. After the contract is concluded, due to a shortage of manpower and materials, costs of production rise to 110. Seller breaches the contract. The only loss Buyer suffers is the machine’s lost value.

In both Examples 1 and 2, the breach of the contract is efficient. Under the assumption of perfect compensation with expectation damages as the measure of recovery, Seller pays damages of 100 to Buyer and reaps a benefit of 10 from the breach. An implicit assumption in the argument that the breach in Examples 1 and 2 is efficient is that due to high transaction costs, renegotiation between Seller and Buyer after contracting is too costly, and with Example 1, high transaction costs also preclude the sale of the machine (or assigning the right to the machine) by Buyer to the second buyer. Indeed, damages remedy, like any liability rule (see Part I), could be justified only with high transaction costs, since otherwise specific performance, like any property rule, would be preferable.

With a remedy of specific performance, and with the assumption of high transaction costs, in both examples Seller will perform the contract inefficiently. Similarly, with disgorgement damages (in the amount of the profits made by Seller in Example 1 or in the amount of the savings made by Seller in Example 2), Seller will lack any motivation to breach efficiently since he gains nothing from a breach. Note that if compensation is lower than expectation damages, Seller in both examples might breach the contract even if performance is efficient. Thus, if damages are 90 (reliance damages), Seller will breach even if the second buyer in Example 1 offers him 95, or if costs of production in Example 2 rise to 95. In both cases the breach is inefficient but creates a benefit of 5 to Seller, which might motivate him to breach.

The notion of efficient breach is justified not only by the general notion of promoting social welfare, but also by the more specific idea of incomplete contracts
(Shavell 2009; Markovits and Schwartz 2011 & 2012). According to the incomplete contracts idea, contract law provides the parties with default rules which apply to their contracts unless they opt out of those rules. The default rules save the parties transaction costs (in terms of negotiation and drafting costs), which they would have incurred but for the default rules. In order to achieve their goal, the default rules should be compatible with most parties’ interests, since otherwise most parties would opt out of the defaults and transaction costs would increase rather than decrease. The default rules will be compatible with most parties’ interests if they are efficient, viz., if they allocate risks and provide the parties with incentives in ways which increase the contractual surplus. The efficient breach idea—so the incomplete contracts argument goes—increases the contractual surplus, and is therefore compatible with most contractual parties’ interests. Therefore, it should be considered a desirable default rule. Note that increasing the contractual surplus serves not only the promisor’s interest, but also the promisee’s, since the expected benefit of the option to breach efficiently will be shared by the two parties through price adjustment when making their contract. Therefore, basing the efficient breach idea on the theory of incomplete contracts might be more effective in defending this idea from deontological (or other non-utilitarian) attacks than basing it on the general notion of promoting social welfare.

**B. Refinements**

The efficient breach argument is valid, as long as damages are fully compensatory. When damages are under-compensatory, specific performance often becomes the most efficient remedy. A notable category of cases where specific performance is the primary remedy since damages are typically under-compensatory is the sale of unique goods. In a now classical article, Anthony Kronman proposed a rationale for the willingness of courts to allow specific performance when the contract’s subject matter is a unique, rather than fungible, good. He suggested that with unique goods, far more so than with fungible goods, there is a substantial risk of under-compensation of buyers for two reasons: first, buyers often attach a subjective value to the unique good, and that value is not compensated for in the event of breach. Second, subsequent to breach of unique good contracts by sellers, buyers incur search costs in finding a substitute good, on top of the
search costs they incurred when finding the original good, and those additional search costs are also not compensated for (Kronman 1978). Since, according to Kronman, the parties would prefer specific performance over damages if the benefit to the seller from having the option to breach and pay damages is less than the costs of the breach to the buyer, and since the costs of the breach to the buyer with unique goods are typically higher than with fungible goods, Kronman concluded that specific performance would more often be preferred by the parties with unique goods than with fungible goods.

Another assumption underlying the efficient breach argument relates to transaction costs. In gain-seeking breaches such as in Example 1, the efficient breach argument assumes that transaction costs make it hard, even impossible, for Buyer to find the second buyer who values the subject matter of the contract more, and sell it to her. When this assumption is relaxed, specific performance might be the most efficient remedy.

In an article published shortly after Kronman’s article, Alan Schwartz argued for a much broader application of specific performance than suggested by Kronman. Schwartz claimed that the risk of under-compensation is substantial not only with unique goods, as Kronman suggested, but also with many fungible goods (Schwartz 1979). Furthermore, in contrast to Kronman, Schwartz claimed that the benefit to the seller of having the option to breach and pay damages is often less with fungible goods than with unique goods. Therefore, with fungible goods, the parties’ ex-ante preferences would not necessarily warrant damages, rather than specific performance, as their preferred remedy. Furthermore, for Schwartz, the central consideration in the choice between damages and specific performance as a remedy is which one entails lower post-breach negotiation costs. This depends, according to Schwartz, mostly on whether (in our Example 1) it is easier for the seller or for the first buyer to find the second buyer who values the good more than the first buyer: only if it is the seller could damages be preferable to specific performance. Thomas Ulen took the argument for specific performance one step further, suggesting that specific performance should be the routine remedy for breach. Ulen, like Schwartz, regarded the post-breach negotiation costs as a central factor in the efficiency of specific performance (Ulen 1984).
While the first-generation writings on efficient breach focused on the promisor's decision to perform or breach, Richard Craswell analyzed the effects of the contractual remedies on various decisions made by the promisor and promisee (Craswell 1988). Craswell considered the effect of the remedies on the decisions as to whether to enter into the contract in the first place and what level of precautions to take in order to reduce the probability of breach. Craswell explained that even if post-breach negotiation costs are zero, the prevailing remedy will still affect the parties’ decisions made before the promisor’s decision whether to perform or breach.

In the economic analysis of contract law it is implicitly assumed that the efficient breach argument is equally valid with respect to both loss-avoiding and gain-seeking breaches (Posner 2009). By contrast, lay people’s intuition is different: experimental studies have indicated that people react more tolerantly to loss-avoiding breaches than to gain-seeking breaches (Baron and Wilkinson-Ryan 2009). Moral philosophers have also distinguished between the two types of breaches, arguing that a breach to pursue a gain is more reprehensible than a breach to avoid a loss (Zamir and Medina 2010, p. 265). Lastly, behavioral law and economics could explain people’s different reactions to the two types of breach as a reflection of people’s different attitudes to losses as opposed to gains (Cohen and Knetsch 1992; cf. Zamir and Ritov 2010). Recently Maria Bigoni, Stefania Bortolotti, Francesco Parisi & Ariel Porat suggested that also from an economics perspective, the case for allowing the promisor an option to breach is typically more vital in loss-avoiding breaches than in gain-seeking breaches (Bigoni et al. 2014).

C. The Promisee’s Incentives

While the effects of remedies on the incentives of the promisor have been thoroughly analyzed in the literature, only a few scholars have analyzed their effects on the promisee’s incentives. A notable exception is Robert Cooter who pointed out that with full compensation the victim’s incentives are eroded and over-reliance might result (Cooter 1985). For example, if the promisee knows that there is a high likelihood of a breach, he might rely as if the likelihood of a breach is zero, knowing that he can reap the

---

5 For an analysis of the mitigation of damages defense, see Goetz & Scott 1983. But the mitigation of damages defense applies only after a breach (or an anticipatory breach) occurs, and it is effective only when the victim's behavior is verifiable.
benefits of reliance if performance takes place but externalizes its costs to the promisee—
who will reimburse him for those costs—if a breach occurs. Cooter suggested that when
damages to the promisee are awarded at a fixed amount, the promisee relies efficiently
since he fully internalizes both the costs and benefits of his reliance. Thus, liquidated
damages, if they are set at the level of expected harm and remain invariant with respect to
actual harm, might solve the over-reliance problem.\(^6\)

Robert Cooter and Ariel Porat discussed the erosion of the victim’s incentives
not only with respect to over-reliance, but also with respect to noncooperation (Cooter &
Porat 2002). They noted that with fully compensatory damages the promisee who could
cooperate with the promisor and reduce the probability of a breach might be unwilling to
do so, especially if noncooperation is nonverifiable (otherwise a duty of cooperation or a
comparative fault defense could solve the problem (Porat 2009)). Cooter and Porat
suggested a novel theoretical solution which they called “anti-insurance.” According to
their solution, the promisee and promisor make a contract with a third party (“anti-
insurer”) according to which the promisee assigns his right to damages to the anti-insurer,
so that in case of a breach the promisee receives no compensation and the promisor pays
fully compensatory damages to the anti-insurer. Before a breach occurs the anti-insurer
pays to the promisor and promisee for the valuable right to collect damages in case of a
breach. Both the promisor and promisee now have efficient incentives since they fully
internalize the costs and benefits of the precautions they will take to reduce the
probability of a breach or reduce over-reliance.

Another more practical solution for improving the promisee’s incentives is to
under-compensate her. With the risk of under-compensation, the promisee will be more
willing to cooperate or avoid over-reliance than with fully compensatory damages.
Indeed, with imperfect compensation the promisor’s incentives are deficient, but once the
promisee’s incentives are taken into account allowing some level of under-compensation
could be optimal (Cooter and Porat 2004).

\(^6\) But may create other inefficiencies: see infra Section IV.B.
III. Scope of Liability

Not all harms are compensable and not all victims can recover. First, in both torts and contracts only foreseeable harms are recoverable. Second, in tort law, proximate cause is a limit to liability: if the harm is too remote liability will not be imposed. Third, in negligence law, liability will be imposed only if the negligent injurer owed a duty of care to the victim. Both proximate cause and duty of care are used by courts as a means to limit liability for policy considerations (Restatement 3d Torts: Liability for Physical and Emotional Harm, § 7, cmt. a, 2010; Dobbs 2000, p. 448). This section discusses several topics, all of which raise questions as to the appropriate scope of liability.

A. Foreseeability

In both torts and contracts foreseeability of losses is a precondition for the imposition of liability (Dobbs 2000, pp. 443-470; Farnsworth 2004, pp. 792-799). There are several efficiency justifications for this requirement.

First, if the losses are not foreseeable there is no sense in imposing liability for their materialization, since such liability will not affect the behavior of the injurer or the promisor. That is because if losses are unforeseeable the costs of taking them into account when deciding which precautions to take are prohibitively high, so the injurer and the promisor will ignore them anyway, with or without liability (Landes and Posner 1987, pp. 246-247).

Second, in negligence, when injury of any kind is unforeseeable—namely, the probability of an injury is very low—expected harm is low and the costs of precautions typically exceed expected harm. When the costs of precautions are higher than expected harm the injurer is non-negligent and liability should not be imposed. According to this justification for the foreseeability requirement, the unforeseeability of any injury is an indicator that there is no negligence on the injurer’s part to begin with. Note that this justification applies to a narrow set of cases where any injury—as opposed to only the one which resulted in the litigated losses—is unforeseeable.

Third, on many occasions the unforeseeable losses are foreseeable for the promisee and victim. In such cases the promisee and victim are typically the cheapest cost-avoiders of the unforeseeable losses, while the injurer and promisor are not. Leaving
the unforeseeable losses on the victim’s and promisee’s shoulders motivates them to take measures to avoid those losses, either before or after the wrong or breach takes place. This is especially essential when those measures are nonverifiable, and neither the comparative fault nor the mitigation of damages defense is applicable.

Fourth and last, in the contractual context the foreseeability requirement incentivizes the promisee to disclose private information to the promisor regarding the promisee’s unforeseeable losses, which is vital to the promisor for the purpose of deciding whether to breach or perform, and what level of precautions to take to avoid a breach. This justification for the foreseeability requirement is relevant, when conveying information from one party to another is possible and when the unforeseeable losses are foreseeable for the promisee (Ayres and Gertner 1989). To better understand this justification, assume that the promisee’s expected losses from a breach are 100 and foreseeable for her, but only a loss of 10 is foreseeable for the promisor. Without the foreseeability requirement, the promisee would not convey to the promisor any information regarding his high potential losses, since if the promisor knew about that, he would charge a higher price for his undertakings. As a result, with only partial information, the promisor would have deficient incentives to perform. In contrast, with the foreseeability requirement, the promisee would convey the information regarding her high potential losses to the promisor in order to make those losses foreseeable for him and therefore recoverable by the promisee. Conveying the information would secure the promisor’s efficient incentives, thereby increasing the contractual surplus.

B. Pure Economic Loss

Victims often suffer pure economic losses. In contracts, losses are often purely economic in the sense that the promisee lost profits and nothing else, but as has already been explained, liability is routinely imposed and for good economic reasons. In torts, in contrast, courts are commonly reluctant to impose liability for pure economic losses. There are several efficiency considerations which could justify this reluctance.

---

7 Supra Part II.
First, pure economic losses are often a private rather than a social cost\(^8\) (Bishop 1982). Imagine an injurer who created a nuisance to restaurateur 1, who shut down his restaurant for a week and lost profits of 100. Assume that the restaurant’s patrons fully mitigated their losses by dining during that week at another restaurant owned by restaurateur 2. Further assume that the profits gained by restaurateur 2 from restaurateur 1’s patrons are 100, in addition to the regular profits made from his own regular patrons. From a social perspective—so goes the economic argument—no social harm has been done: profits were just transferred from one person to another. With no social harm, injurers should not take any costly precautions, so liability should be nil.

This argument, however, has limits. To start with, if restaurateur 1 is a recurring loser and restaurateur 2 a recurring winner, no liability might inefficiently suppress restaurateur 1’s activity (Bishop 1982).

Second, with no liability, restaurateur 1 might take costly precautions to avoid the harm; if the injurer, rather than restaurateur 1, is the cheapest cost-avoider, liability would save the costs of precautions of restaurateur 1, even if at a cost to the injurer who—given his expected liability—would take precautions and prevent the harm (Dari-Mattiacci & Schäfer 2007).

Third, loss of profits—as in our example—are not necessarily just transfers of value from one person to another: The destruction of input might increase the marginal cost of production, leading to production decrease and higher prices (Rizzo, 1982a & 1982b).

Fourth, with pure economic losses causation is often hard to prove, so even if we think that those losses should be prevented, there is a risk that with liability the injurer will pay more than what he actually caused and be over-deterred. The reason why causation is harder to prove in pure economic loss cases than in physical injury cases is that those losses are of a type which occurs regularly with no wrongdoing, and therefore it is hard to distinguish them from non-wrongful losses. To illustrate, assume that in the

---

\(^8\) This, however, is not always so. Sometimes pure economic losses are a substitute for physical losses which are not compensated for practical reasons. For example, if the defendant polluted a river with chemical effluents and destroyed the wildlife in a certain area, since no one in particular owns the wildlife, recovery for the physical harm is impossible. Fishermen, however, might be able to recover for lost profits even if those are pure economic losses: the lost profits could serve as a proxy for the social value of the lost wildlife. Cf. Pruitt v. Allied Chemical Corp., 523 F. supp. 975 (1981).
nuisance example, restaurant 1 was not shut down, but restaurateur 1 nonetheless lost profits due to low attendance by patrons. Low attendance, however, could be caused by many other causes—most of them non-wrongful—and it could be hard, sometimes even impossible, to isolate the effects of the wrongful from those of the non-wrongful causes. If courts tend to resolve uncertainties in favor of victims—and they are often so inclined—they might impose too heavy a liability burden on the injurer (Abraham 2011, pp. 1781-1783).

Fifth, pure economic losses often have two characteristics which make the victim an especially effective cost-avoider. The first characteristic is, that those losses are accumulated over time and therefore victims have a relatively long period of time to mitigate them. The second characteristic—which has already been mentioned—is that economic losses are of a type which occurs regularly with no wrongdoing; as a result, victims have expertise in handling and reducing them. The nuisance case is not the best example to illustrate these two characteristics, especially if the restaurant is shut down because of the nuisance. So let’s take another example: suppose a person is wrongfully injured in a road accident and has not showed up at work for two months. The injured person’s employer argues that he has suffered pure economic losses in terms of lost profits due to his employee’s absence from work. Obviously, the employer will not be able to recover. A possible justification for this result is that the employer is a very effective cost-avoider: her lost profits accumulate from day to day, so she has time to consider how to mitigate them, e.g., by hiring a substitute employee, reducing her activity level, or postponing the performance of some of the work for a few months. Furthermore, the employer is accustomed to handle such losses on a daily basis—employees are often absent from work for non-wrongful causes—and must have expertise in minimizing them. Indeed, even if the employer had been entitled to compensation for her lost profits she would have been required to mitigate those losses as a precondition for any recovery; however, most of the employer’s failures to mitigate losses are nonverifiable, and therefore the most effective way to encourage her to efficiently mitigate losses is just to let her bear them.
C. Non-Pecuniary Loss

Non-pecuniary losses are losses which have no economic impact on the victim, such as emotional distress, agony, disappointment, and pain and suffering. In contracts, non-pecuniary losses are compensated almost only when the main interest protected by the contract is non-pecuniary in nature (Farnsworth 2004, p. 810). Typical examples are tour package contracts,\(^9\) contracts to perform cosmetic surgeries,\(^10\) and contracts for providing services for weddings or funerals.\(^11\) In torts, non-pecuniary losses are typically compensated when accompanied by physical injury, mainly bodily injury, and only rarely when those are standalone losses (Dobbs 2000, pp. 1050-1053).

One objection to the imposition of liability for non-pecuniary losses is that those losses are subjective, and therefore the risk of plaintiffs’ bringing frivolous claims is high. If those claims were to be allowed—so the argument goes—injurers might pay excessive damages and be over-deterred. This objection, however, is less persuasive when there is some objective evidence for the existence of the non-pecuniary losses and their magnitude, such as when the victim suffers pain and suffering accompanied by bodily injury (Bovbjerg et al. 1989).

Another objection is that inflicting non-pecuniary losses on the victim typically does not decrease her marginal utility of money, so transferring payments from the injurer to the victim typically does not improve social welfare (cf. Shavell 1987, pp. 228-31; Danzon 1984, pp. 517, 521). This objection is attenuated and even disappears when deterrence is considered: since non-pecuniary losses are social losses, injurers and promisors should internalize them (Rea 1982).

The marginal utility of money argument coupled with the countervailing deterrence argument provides a compelling explanation of courts’ willingness to allow compensation for non-pecuniary losses (almost) only when the interest protected by the contract is mainly non-pecuniary. Imagine a construction contract between Builder and Owner, where the parties anticipate when making their contract that with a certain probability performance will be delayed and as a result Owner will be disappointed. Would they agree that in case of a breach Owner will be compensated for his

---

disappointment? If yes, liability should be the contractual default rule; otherwise, no-liability should be the default rule.

Let’s start with the marginal utility of money argument. According to this argument, Owner will never insure against disappointment loss, because suffering this loss will not affect her marginal utility of money. Similarly, Owner will not “insure” against such a loss through the contract (since he will be required to pay a premium through the contract price). The deterrence argument does not change this conclusion, as long as we assume that most of the losses resulting from the breach are pecuniary. Since pecuniary losses are compensated, deterrence is reasonably attained, even if a small fraction of the losses—those which are non-pecuniary—are not. Thus, in our example, the marginal utility of money argument seems to overcome the countervailing deterrence argument and the parties would prefer, ex ante, not to have liability for non-pecuniary losses (cf. Rea 1982).

If instead, most of the losses expected to result from the breach are non-pecuniary, the conclusion will be different. Take, for example, a contract made between a travel agency and a traveler for providing a package tour by the former to the latter. Here, the efficient rule, which most contractual parties would prefer, is liability for non-pecuniary losses. Although the traveler would never insure against such losses with an insurer, he would insure against them with the travel agency, since such “insurance” would provide the agency with efficient incentives to perform the contract. In this example, as opposed to the previous one, most of the expected losses are non-pecuniary, and no liability for those losses would result in severe under-deterrence.12

D. Caps on Consequential Damages

The foreseeability requirement, which has been discussed in Section A, caps damages by allowing recovery for foreseeable losses only. Contractual parties, however, often cap damages even further, sometimes precluding liability for any consequential losses altogether (Farnsworth 2004, p. 799).

One reason for contractual parties to cap consequential losses relates to the fact that consequential losses are often unique to the promisee, as when the breach causes the

---

12 For the argument that victims might be willing to insure against non-pecuniary losses, see Croley & Hanson 1995.
promisee a loss of profits coming from third parties. In such cases the promisee can often take steps to mitigate his consequential losses in nonverifiable ways, and making the promisee bear those losses would incentivize him to do so. At the same time, because of their uniqueness, those losses are often not accurately anticipated by the promisor (even if they are foreseeable enough to pass the foreseeability threshold), and therefore capping them promotes certainty. Indeed, without liability the promisor’s incentives to perform efficiently decreases since he externalizes some of the costs of the breach to the promisee, but given the advantages described above, this flaw might be a price worth paying.\footnote{Leaving some losses uncompensated also mitigates the promisee’s moral hazard problem: see Shavell 1979.}

\textit{A second reason}, which applies mostly to consumer contracts, is that capping damages is an effective tool for avoiding adverse selection. Take an example used in the literature (Epstein 1989). A shipper ships packages for owners which sometimes are damaged. Some packages are worth more, others worth less. The shipper charges a uniform price (discriminating in prices is either too costly or illegal), and when a package is damaged he is liable for the full amount of the harm done. Under this liability scheme, owners whose packages are worth less than average subsidize owners whose packages are worth more than average, since, by assumption, they all pay a uniform premium through the price. This cross-subsidization is unjust, especially if owners of high-value packages are typically wealthier than owners of low-value packages. Moreover, cross-subsidization results in inefficient consumption and adverse selection. Specifically, because owners with low-value packages pay more than what they should have paid given their low expected harm, their consumption level is too low; in contrast, because owners with high-value packages pay less than what they should have paid given their high expected harm, their consumption level is too high. In addition, in the long run, because fewer owners of low-value packages will consume the shipping services, prices will go up (since the average expected harm will go up). As a result, even fewer owners of low-value packages will consume the shipping services, prices will go up again, and so on and so forth. At the end, only consumers with very high-value packages will consume the shipping services and, in the extreme case, the shipping services will shut down. One
way to solve the problem is to cap damages. Thus, the shipper could offer contracts to consumers under which damages are limited to the value of low-value packages. This would mitigate the inefficiencies, including the adverse selection problem, and would also be more just. Consumers with high-value packages would be able to secure full compensation through first-party insurance, if they choose to do so.

IV. The Measure of Recovery

Generally, in both torts and contracts damages are compensatory. But it may sometimes be unclear what compensatory damages are. The first section deals with this question in a very specific tort context: damages for bodily injury. The second section deals with one contractual exception to the compensatory damages principle, which is liquidated damages.

A. Bodily Injury and Lost Income

A major component in any award of damages for bodily injury is lost income. The result is that a high-income (“rich”) victim receives compensation which is higher—sometimes much higher—than what a low-income (“poor”) victim receives, even if both victims suffer from the same bodily impairment because of the wrongdoing. By contrast, when courts set the standard of care they do not distinguish between rich and poor victims: the injurer is required to take the same level of care toward the victim regardless of his income, even if his type as rich or poor can be anticipated by the injurer.\(^\text{14}\) This leads to an inconsistency in tort law: there is a misalignment between the standard of care and damages (Porat 2011). The following example illustrates this point.

Example 3. Poor and rich neighborhoods. John drives his car at a speed of 30 mph in a rich neighborhood. Unfortunately, he hits a pedestrian as she is crossing the street. Had John driven a bit more slowly, he would have succeeded in stopping his car in time and preventing the accident. A day later, John drives his car again at the same speed, but this time in a poor neighborhood. Once again, he hits a pedestrian. All driving conditions are exactly the same as they were in the rich neighborhood the day before; therefore, in this case as well, the accident would have been avoided had John driven his car a bit more slowly. Is it possible that, under a rule of negligence, the same court would find John liable for the first accident but not for the second?

\(^{14}\) When a victim’s type as rich or poor cannot be anticipated in advance by the injurer, there can be no question that the standard of care should be set according to the average victim.
Assuming that in the rich neighborhood most people have a higher income than the residents of the poor neighborhood, one could argue that different standards of care should be applied in the two neighborhoods. It is quite possible, even reasonable, then, that the same court would find that: (a) John failed to take due care in the rich neighborhood and therefore should be held liable toward his victim; and (b) John took due care in the poor neighborhood and therefore should be exempt from all liability.

Courts, however, do not set different standards of care for driving in rich and poor neighborhoods. Similarly, they also do not set different standards of care for doctors treating rich and poor patients. If a court were required to explain the application of the same standard of care for the rich and the poor, it would reason that the lives and limbs of the rich and poor have identical social value and, therefore, are deserving of the same level of legal protection. But such reasoning, convincing as it may be, is inconsistent with the practice of awarding higher damages to rich victims. This practice suggests that rich people’s lives and limbs are more highly valued by the law relative to poor victims. To be consistent with this practice, so it seems, injurers should take greater care toward the rich than the poor, just as they should be more careful in their interactions with high-value property. Therefore, to restore consistency to the law, courts should have chosen one of two routes: to either apply different standards of care to rich and poor victims (contrary to what they actually do), coupled with different levels of compensation (as they actually do), or, alternatively, to apply the same standard of care to rich and poor victims (as they actually do), coupled with the same level of compensation (contrary to what they actually do).

From a social perspective, there is no compelling reason why lost income should be the main criterion for valuing people’s lives and limbs, but this question is beyond the scope of this chapter. Therefore, let us assume first that lost income is the right criterion, and consider the efficiency of the law under this assumption. If lost income is the right criterion, and given that the standard of care is set uniformly according to average income but damages are awarded in the amount of the victim’s lost income, injurers will comply with the standard of care when they expect a rich victim but under-

\[15\] Lost income could be correlated with people’s productivity. It is questionable, however, whether this correlation is strong, and whether productivity is the main value of people’s lives and health.
comply when they expect a poor victim. To see why, assume that the expected harm of rich victims is 15, the expected harm of poor victims is 5, and average expected harm is 10. Further assume that the standard of care is set at 10, namely, injurers are required to take precautions up to 10 to reduce the risk to either rich or poor victims. With these figures, injurers will take precautions up to 10 toward rich victims (and pay no damages if harm occurs) and up to 5 toward poor victims (and pay damages if harm occurs). Since the expected harm for rich victims is 15 and for poor victims 5, injurers will be under-detected toward rich victims and optimally-detected toward poor victims.

Assume now that lost income is not the right criterion for awarding damages for bodily injury, and that rich and poor people’s lives and limbs have the same value. Further assume that people’s lives and limbs are determined by average income. Now, taking precautions by injurers up to 10 toward rich victims is efficient, since their expected harm is assumed to be 10. Conversely, taking precautions up to 5 toward poor victims is inefficient, since 10, rather than 5, is assumed to be their expected harm.

The analysis so far has implicitly assumed no courts’ errors in setting the standard of care and awarding damages and no injurers’ errors in anticipating courts’ decisions. With the risk of errors, the analysis becomes more complex (Porat 2011), but does not change the basic conclusion: efficiency-wise, the value of people’s lives and limbs should be reflected in both the standard of care and damages in a consistent manner, otherwise inefficiencies will result.

B. Liquidated Damages

While the default rule in contract law is that damages are compensatory, the parties are free to opt out of the default rule by incorporating a liquidated damages clause. If, however, damages are set too high courts are authorized to strike down the liquidated damages clause as being a penalty, and to award compensatory damages instead. When deciding whether to uphold a liquidated damages clause, courts are instructed to consider the anticipated loss at the time of contracting, the actual loss, and the difficulties of proving the actual loss (Restatement 2d Contracts, §356, 1981). As liquidated damages

---

16 This is a simplifying assumption: a plausible argument is that average income should also not play a major role in valuing victims’ lives and limbs (Friedman 1982; Porat & Tabbach 2011).

17 Distributive justice considerations could provide at least a partial explanation for the inconsistency in the law (Porat 2011, at pp. 105-7).
are closer to both anticipated and actual losses, and as there are more difficulties in proving actual loss, courts are more willing to uphold the liquidated damages clause. The Uniform Commercial Code ("UCC") adds a fourth consideration to be taken into account, which is “the inconvenience or nonfeasibility of otherwise obtaining an adequate remedy” (UCC, §2-718).

It is puzzling why courts are authorized to scrutinize liquidated damages clauses, while they are generally not authorized to do the same thing with other clauses. (Posner 1979, p. 290). Indeed, at least when the parties are rational and well-informed, there is no reason to assume that the liquidated damages clause they incorporated into their contract sets damages too high (Schwartz 1990). In the next paragraphs both the advantages and disadvantages of liquidated damages are discussed, when the question which arises is whether the disadvantages can justify the wide discretion granted to courts to strike down liquidated damages clauses.

Incorporating a liquidated damages clause into their contracts has several advantages for the parties. First, it saves litigation costs. Indeed, since the actual loss is a consideration for the court whether to uphold the clause, the parties might litigate as to the exact magnitude of the actual loss. But that would not be necessary in most cases, since as long as the gap between the actual loss and liquidated damages is not large, courts will tend to uphold the clause; on many occasions, the defendant will not even try to argue that the gap is large, and it would not be necessary to litigate about actual losses (Goetz and Scott 1977).

Second, a liquidated damages clause promotes certainty since with this clause the parties can accurately anticipate the amount of damages to be paid by the promisor to the promisee in case of a breach. This advantage is prevalent if the liquidated damages clause sets both a floor and a ceiling, as is the case when the parties have not agreed otherwise.

Third, a liquidated damages clause protects interests which otherwise are not adequately protected by contract law. Thus, if the parties anticipate that a breach would result in non-pecuniary losses, or losses which are hard to prove (such as reputational losses), they might incorporate into their contract a liquidated damages clause relating to
such losses (Goetz and Scott 1977). Consequently, not only will compensation be secured, but the promisor's incentives will improve as well.  

*Fourth*, a liquidated damages clause is a solution to the promisee’s over-reliance problem: as has been explained, with liquidated damages the promisee internalizes both the costs and benefits of his reliance and therefore relies efficiently.  

*Fifth* and last, a liquidated damages clause enables the promisor to signal his credibility to the promisee: by undertaking to pay a large enough amount of damages in case of a breach, he is able to signal to the promisee that the probability of a breach is low (Posner 2011, p. 160). Thus, a landlord might hesitate to lease his property to a tenant he hardly knows, fearing the latter may damage the property or fail to evacuate it on time. Given that the level of enforcement is lower than 100%, and given litigation costs, the landlord’s entitlement to compensatory damages might not be a satisfactory guarantee for the tenant’s performance and the landlord will not lease the property to her. If, however, an enforceable liquidated damages clause, stipulating damages in an amount which is much higher than expected losses, is incorporated into the contract, the landlord might be convinced to lease his property to the tenant after all.

This advantage, however, is generally unattainable, since in order to attain it the parties must stipulate damages in an amount which is much higher than both expected and actual losses; courts would not enforce such a stipulation and would strike it down as a penalty.

Liquidating damages clauses come at a cost: sometimes they might result in inefficiencies. *First*, they might lead to inefficient breach or inefficient performance (and also inefficient investment in precautions); the former will occur if liquidated damages are lower than actual losses, the latter if they are higher than actual losses.  

If this were the main disadvantage of liquidated damages clauses, courts’ power to strike them down should have been a default rule. After all, it is for the parties to decide whether the liquidated damages clause should be set aside if it largely deviates from actual losses.

---

18 Although under-compensating some losses could improve the promisee’s incentives. See Shavell 1979.
19 *Supra* Section II.C.
20 On the other hand, the promisor who is generally reluctant to breach, even efficiently, for moral reasons, might be willing to breach when damages are stipulated in the contract. *Cf.* Wilkinson-Ryan 2010. Thus, liquidated damages clauses might sometimes encourage efficient breaches.
Second, if liquidated damages are sets too high, the promissee's expectation for over-compensation might tempt her to inefficiently induce a breach in non-verifiable ways (Clarkson, Miller and Muris 1978). This risk, however, is not necessarily a reason for courts to intervene in liquidated damages clauses, as long as the parties are assumed to be aware of such risk when stipulating damages in their contracts.

Third, because of lack of information, irrationality or bounded rationality, the promisor might not be aware of the harsh consequences of a liquidated damages clause, and agree to set it too high. For example, he might be overly optimistic about his ability to perform the contract, believing he will almost never be subject to the liquidated damages clause. Liquidated damages set too high might encourage the promisor to overinvest in precautions and perform even when a breach is efficient. If this were the main disadvantage of liquidated damages clauses, courts’ power to strike them down should have been limited to cases where asymmetric information or irrationality is a real concern, such as in some consumer contracts.

Forth, sometimes, through a liquidated damages clause, the parties might try to externalize costs to third parties by setting damages much higher than what efficiency requires. Thus, suppose that in our previous example the tenant is under a substantial risk of bankruptcy and the landlord is aware of this. The parties might reach an agreement, setting damages five times higher than anticipated losses, knowing that if the tenant eventually goes bankrupt, part of the costs of the liquidated damages will be borne by his creditors. In such cases, intervention by courts is essential regardless of the parties’ wishes, even if the parties are well informed and fully rational.21

V. Partial Recoveries

In this part, two concepts of partial recoveries are discussed: probabilistic recoveries and offsetting risks. The first concept is well known and has been applied by courts in some specific categories of cases, while the second is less known and has not been applied by courts so far.

---

21 A liquidated damages clause which is above expected harm might have anti-competitive effects since it makes the breach for the promisor very costly (Chung 1992; Spier and Whinston 1995).
A. Probabilistic Recoveries

In civil cases the plaintiff succeeds in trial if he proves his case by a preponderance of the evidence. The question which arises is whether in certain cases tort victims should be entitled to a probabilistic recovery if they have failed to prove causation by a preponderance of the evidence. Typical cases where this question emerges are medical malpractice cases when the doctor’s negligence reduced the patient’s chances of recovery and the patient eventually did not recover (King 1981; Levmore 1990; Porat and Stein 2001, pp. 122-5). For example, suppose that the patient’s chances of recovery were 30%, but because of negligent misdiagnosis by the doctor those chances were reduced to zero. The probability that the doctor’s negligence caused the patient’s non-recovery is 30%. Should the patient recover for 30% of his losses since the doctor deprived him of his chances of recovery?\(^{22}\) In some jurisdictions the answer is yes, while in others the answer is no (Porat and Stein 2001, pp. 74-6).

Does efficiency require compensation in this case? Not necessarily, if we assume that in the specific category of cases doctors handle not only cases of less than 50% chances of recovery, but also cases of more than 50% chances of recovery, and there is symmetry between the two groups of cases. With such symmetry, a doctor’s liability would be the same under both a probabilistic recovery and a preponderance of the evidence rule. To see why, imagine that a patient has either a 30% chance of recovery or a 70% chance of recovery, with equal probabilities. The doctor is negligent and reduces the patient’s chances of recovery to zero. The patient does not recover and suffers harm of H. Under a probabilistic recovery rule, the negligent doctor’s expected liability is \(50\% \times 30\% \times H + 50\% \times 70\% \times H = 50\% H\). Under a preponderance of the evidence rule, the negligent doctor’s expected liability is the same: \(50\% x0 + 50\% xH = 50\% H\).\(^{23}\) Liability of

\(^{22}\) Suppose that chances of recovery were reduced by the doctor's negligence from 70% to 40% and the patient has not recovered; under a probabilistic recovery rule the patient should recover 50%--not 30%--of the ultimate harm, since the probability that she suffered that harm from the doctor's negligence is 50% (Porat & Stein, at p. 124, 2001; Restatement 3d Torts: Liability for Physical and Emotional Harm, § 26, cmt. n, 2010).

\(^{23}\) There is a probability of 50% that the patient has a 30% chance of recovery, and then liability would be 30%H, and there is a probability of 50% that the patient has a 70% chance of recovery, and then liability would be 70%H.

\(^{24}\) There is a probability of 50% that the patient has a 30% chance of recovery, and then liability would be zero, and there is a probability of 50% that the patient has a 70% chance of recovery, and then liability would be H.
50%H in this example is also efficient, since the expected harm of the doctor’s negligence is 50%H.\textsuperscript{25}

Things are different if there is an asymmetry between the two groups of cases. Consider an extreme example where in a certain hospital department all the patients have a less than 50% chance of recovery. Under a preponderance of the evidence rule, those patients would never be entitled to compensation, since in each and every case when harm occurs the probability that the negligent doctor caused the harm is less than 50%. Therefore, under the latter rule, doctors will not be deterred. In contrast, under the probabilistic recovery rule, patients who suffer harm will always be entitled to compensation from negligent doctors, and the latter’s expected liability will equal expected harm, as required for efficiency. Note that even when there is symmetry between the two groups of cases, probabilistic recovery is essential, as long as the doctor can identify in advance whether her patient has a less than or more than 50% chance of recovery.

It is important to note that probabilistic recovery is more essential if the typical case is less than 50% chance of recovery rather than more than 50% chance of recovery. When chances are less than 50%, with a preponderance of the evidence rule, under-deterrence will result. When chances are more than 50%, with the same rule, over-deterrence might result, but only if we assume courts’ errors in setting the standard of care and awarding damages or injurers’ errors in anticipating courts’ decisions. (Porat 2011, pp. 112-4).

To see why, suppose that in a certain hospital department all the patients have a 70% chance of recovery, and the harm, if it occurs, is H. Bearing these figures in mind, the patient’s expected harm is 70%H, and the doctor should be required—efficiency-wise—to take precautions up to 70%H. Would the doctor take higher precautions, since if harm occurs he bears a liability of H? The answer is no: the doctor will take precautions up to 70%H, thereby satisfying the standard of care and bearing no liability. The doctor might, however, over-comply if there is a risk of courts’ errors in setting the standard of care and awarding damages or injurers’ errors in anticipating courts’ decisions. Since

\textsuperscript{25} There is a probability of 50% that the patient has a 30% chance of recovery, and then expected harm is 30%H, and there is a probability of 50% that the patient has a 70% chance of recovery, and then expected harm is 70%H. Thus, expected harm is 50% \times 30% \times H + 50% \times 70% \times H = 50% H.
such risk is common, the probabilistic recovery rule might be superior to a preponderance of the evidence rule, not only in less-than-50% chance cases, but also in more-than-50% chance cases.

Beyond lost chances of recovery cases, the Market Share Liability doctrine (MSL), which is a form of probabilistic recovery rule, has been applied by courts in some jurisdictions. In one case to which the doctrine was applied, numerous manufacturers produced the same generic drug, which later was found to be defective and harmful. Since in most cases plaintiffs could not tell which manufacturer’s drug caused their harm, a preponderance of the evidence rule allowed all manufacturers in most suits to escape liability. Under MSL, however, each manufacturer was required to bear liability according to his market share at the relevant time with respect to the harmful drug. Thus, if one manufacturer’s market share was 10%, he was required to compensate each and every plaintiff who suffered harm of H, in the amount of 10%H (Porat and Stein 2001, pp. 58-69). Although MSL does not provide fully efficient incentives to injurers, it is more efficient than no liability. Inefficiency might still result because the manufacturers’ behaviors are often nonverifiable, and even with MSL a manufacturer might refrain from taking costly but efficient precautions, since most of the benefits of his precautions would be captured by the other manufacturers (Cooter and Porat 2007).

B. Offsetting Risks

Consider the following example (Porat 2011):

Example 4. Choosing between two medical treatments. A doctor must decide between Treatment A and Treatment B for his patient. Each treatment entails different risks but produces the same utility if the risks do not materialize. This utility is much greater than the respective risks of each treatment. The costs of administering the treatments are the same, and the costs of choosing between them are low. Treatment A entails a risk of 500 to the patient’s left arm (there is a probability of .01 that the treatment will produce harm of 50,000), and Treatment B entails a risk of 400 to the patient’s right arm (there is a probability of .01 that the treatment will produce a harm of 40,000). The risks of Treatments A and B are not correlated: the realization of the risk from one treatment has no bearing on the probability of the realization of the risk from the other treatment. The doctor negligently chooses Treatment A, and a harm of 50,000 materializes. Should the doctor be held liable? If so, in what amount?

26 Note that one of the treatments could be an omission, such as not operating on the patient or not administering a certain medicine.
Under prevailing tort law, the doctor in Example 4 would be found liable because he was negligent: he could have reduced the total risk to the patient by 100 (500-400) at a low cost, but failed to do so. The negligent doctor’s liability under prevailing tort law would amount to the entire harm, which is 50,000, since that is the harm caused by his negligence. Thus, while the net risk created by the doctor’s negligence is 100, his expected liability is five times higher: .01 x 50,000 = 500. The reason for the misalignment between net risk and expected liability is that tort law ignores the fact that the negligent doctor, by choosing treatment A, not only increased the risk to the patient's left arm (of 500), but also decreased the risk to the patient's right arm (of 400). If the doctor bears liability for the increased risk (internalizes the negative externalities), without being credited for the decreased risk (does not internalize the positive externalities), his expected liability will be higher, even much higher, than social costs, which are the net rather than the gross risk.

Liability for far more than social costs is likely to result in over-deterrence, which in the medical context often takes the form of defensive medicine. Over-deterrence and defensive medicine will result if there are courts’ errors in setting the standard of care and awarding damages or doctors’ errors in anticipating courts’ decisions (nonverifiability of some harms but not others is a typical problem leading to defensive medicine). As doctors pay more in damages, over-deterrence and defensive medicine become more severe (Porat 2007).

To restore the alignment between social costs and expected liability, the doctor’s liability in Example 4 should be 10,000 rather than 50,000. Liability of 10,000 would result in this example in expected liability of 100, which is the social cost of the doctor’s negligence. In more general terms, a doctor’s liability should be \( H(R_a-R_b)/R_a \), where \( H \) stands for the materialized harm, \( R_a \) for the risk which was increased by the doctor’s negligence and materialized into harm, and \( R_b \) for the risk which was decreased by the doctor’s negligence.

The same formula can be applied to other cases where an injurer increased the risks to person A and decreased the risks to person B (Porat 2007) (although such application is expected to raise objections mainly from non-economic lawyers). In such cases, as long as a restitution claim against person B is legally or practically impossible,
there is room for the argument that, efficiency-wise, the injurer’s liability toward person A should be decreased in accordance with the risks reduced to person B. A similar argument applies to cases where person B received a certain benefit (as opposed to a probabilistic benefit in terms of decreasing B’s exposure to risks): as long as the benefit is social rather than private, and some other conditions are met, offsetting the benefit from damages awarded to person A would be justified (again, under the assumption that a restitution claim against person B is legally or practically impossible). (Porat and Posner 2014).

References


