ECONOMIC ANALYSIS OF PUNITIVE DAMAGES

Louis T. Visscher*

I. Introduction

A. Economic Goals of Tort Law and the Law of Damages

- Many lawyers regard compensation as the most important goal of tort law. However, in the words of Williams, "this...does not look below the surface of things. Granted that the immediate object of the tort action is to compensate the plaintiff at the expense of the tortfeasor, why do we wish to do this?... An intelligent approach to the study of law must take account of its purpose, and must be prepared to test the law critically in the light of its purpose." Keeton argues that the primary function of tort law is not to compensate the losses, but to determine when compensation is required.² A similar argument is made by Fleming and Rogers.³ Losses of the victim are only shifted to the tortfeasor if there are reasons to do so. These reasons can be found in the goals of tort
- In the economic analysis of tort law, minimization of total accident costs is regarded as the paramount goal. These costs are subdivided into primary accident costs (the costs of precautionary measures and the losses that still occur), secondary accident costs (the costs of having to bear a certain loss) and tertiary accident costs (the administrative costs of the legal system dealing with the accident losses). 4 The reduction of primary costs is achieved by deterrence and the reduction of secondary accident costs by loss spreading. Tertiary costs decrease if the costs of administering the treatment of accidents are reduced. Hence, in the economic analysis of tort law, compensation is not regarded as a goal, but as a means with which the goal of cost reduction is striven for.5
 - Dr. Louis Visscher, LL.M. is Associate Professor in Law and Economics at the Rotterdam Institute of Law and Economics (RILE) of the Erasmus School of Law, the Nether-
 - G. Williams, The Aims of the Law of Tort, Current Legal Problems (CLP) 4 (1951) 137.
 - W.P. Keeton et al. (eds.), Prosser and Keeton on the Law of Torts (5th student edition
 - J.G. Fleming, The Law of Torts (8th edition 1992) 3; W.V.H. Rogers, Winfield and Jolowicz on Tort, (17th edition 2006) 4.
 - G. Calabresi, The Cost of Accidents: A Legal and Economic Analysis (5th printing 1977) 27 ff.
 - For an extensive overview of empirical literature regarding the question whether tort law achieves deterrence, compensation and corrective justice, see D. Dewees, D. Duff and M.

While engaged in their activities, people may create negative externalities, i.e. a probability for others to suffer losses as a result of the activity. Tort law is seen as an instrument that can provide behavioural incentives to the actors, so that they internalize these externalities. The threat of being held liable induces the actors to incorporate the possible losses of others into their decision on how much care to take and how often to engage in the activity. Taking more care and/or reducing the activity level can lower the probability of an accident and thereby the expected accident losses. Optimal care and optimal activity are taken when the marginal costs of taking more care or further reducing the activity level equal the marginal benefits thereof in the sense of a reduction in the expected accident losses.

This economic line of reasoning implies that damages should be high enough for the injurer to internalize the externalities he has caused. Under a rule of strict liability, this in essence means that damages should fully compensate the victim for his losses. Under a rule of negligence, damages should be high enough to make taking due care, which from an economic point of view should equal optimal care, more attractive than applying a lower care level. In situations where both the injurer and the victim can influence the accident probability, no rule is able to provide both parties with the correct activity incentives. Only the residual risk bearer will incorporate all relevant costs in his activity decisions. Under strict liability this is the injurer, under the negligence rule, the victim (because the injurer who takes due care is not liable and hence does not bear the expected accident losses).

Trebilcock, Exploring the Domain of Accident Law. Taking the Facts Seriously (1996). For more recent empirical literature on the prevention goal, see e.g. J.D. Cummins, R.D. Phillips and M.A. Weiss, The Incentive Effects of No-Fault Automobile Insurance, The Journal of Law and Economics (J. L. & Econ.) 44 (2001) 427-464; A. Cohen and R. Dehejia, The Effect of Automobile Insurance and Accident Liability Laws on Traffic Fatalities', J. L. & Econ. 47 (2004) 357-393 and D.P. Kessler and D.L. Rubinfeld, Empirical Study of the Civil Justice System, in: A.M. Polinsky and S. Shavell (Eds.), Handbook of Law and Economics Volume 1 (2007) 343-402.

- ⁶ The remark of *A.P. Scarso* in no. 4 of his report on Italy that "a deterrent purpose of tort law exists when the role of the defendant's conduct is assessed either as a factor affecting the imposition of liability or the amount of damages awarded or in cases where the benefits gained through the damaging event are taken into account in determining the amount of damages" therefore is contrary to the Law and Economics point of view. After all, also a rule of strict liability, where the conduct of the tortfeasor is not assessed, can provide correct care and activity incentives to the tortfeasor.
- See among many others R.A. Posner, A Theory of Negligence, Journal of Legal Studies (JLS) 1 (1972) 29-96; S. Shavell, Strict Liability versus Negligence, JLS 9 (1980) 1-25 S. Shavell, Economic Analysis of Accident Law (1987); W.M. Landes and R.A. Posner, The Economic Structure of Tort Law (1987); S. Shavell, Foundations of Economic Analysis of Law (2004) 178 ff.; H.-B. Schäfer and C. Ott, Lehrbuch der ökonomischen Analyse des Zivilrechts (4th ed. 2005) 129 ff.; R.D. Cooter and T.S. Ulen, Law and Economics (5th ed. 2008) 336 ff.

B. Economic Goals of Punitive Damages

- of punitive damages. Several country reports in this book mention the same, as does Geistfeld regarding the opinion of the U.S. Supreme Court. This opens the question how this viewpoint fits in the economic framework. After all, if the injurer has to pay higher damages than the losses he has caused, a risk of over-deterrence might be created. In such a situation, the injurer takes too much care and/or engages too little in his activity, as compared to the socially optimal care and activity level. Given that the prevention goal aims at *optimal* levels, this goal then is not reached. Furthermore, economic literature regarding 'optimal enforcement' deals with the question of when tort law is the preferable legal instrument, and when criminal law is. These insights are relevant in discussing the punishment goal of punitive damages.
- 6 In the following sections, I will discuss the economic arguments in favour of punitive damages, both from a deterrence point of view as well as from a punishment perspective. Where relevant, I will connect the economic insights to the country reports in this book, to put the economic analysis of punitive damages into perspective. It is important to realize that tort law and criminal law do not operate in a vacuum, but are different legal instruments which both are concerned with undesirable behaviour. In order to keep the analysis focussed on the topic of punitive damages, which is embedded in tort law, my analysis also centres on tort law. Hence, I will not fully discuss the possibilities of criminal law to address the issues that tort law also faces. ¹⁰ I do, however, in Section III.A explain that in Law and Economics, both tort law and criminal law are primarily regarded as instruments which may deter undesirable behaviour. Hence, the view that tort law aims at compensation while criminal law
 - 8 A.J. Sebok, report United States no. 1; V. Wilcox, report England no. 1; J.-S. Borghetti, report France no. 4 and 35; N. Jansen and L. Rademacher, report Germany no. 4; B. Askeland, report Scandinavia no. 4; A. Menyhárd, report Hungary no 6 and 27. B.A. Koch in his report on European Law explains in no. 14 ff. that legislation and ECJ decisions require sanctions to be 'effective, proportionate and dissuasive'. This phrasing suggests that the deterrent function of tort damages is taken seriously and may result in supra compensatory damages. However, Koch makes clear in no. 39 that, even though violations should be prevented to the extent possible, terminology such as 'dissuasive' or 'deterrent' are not intended to promote supra compensatory damages. In no. 41 Koch mentions the possible exception of private enforcement of antitrust rules by way of noncompensatory damages, a topic which I will discuss in no. 15 below.
 - M. Geistfeld, Punitive Damages, Retribution, and Due Process, Southern California Law Review (S. Cal. L. Rev.) 81 (2008) 269. In the recent decision Exxon Shipping Co. et al. v. Baker et al., the Supreme Court again clearly states this, where it considers in Section IV B that 'regardless of the alternative rationales over the years, the consensus today is that punitives are aimed not at compensation but principally at retribution and deterring harmful conduct'.
 - This topic is discussed in the Law and Economics literature on "optimal enforcement", see footnotes 51 ff. below.

aims at punishment and deterrence, is not shared. Given the focus on tort law, I do not discuss the possibilities of the so-called *act-based* sanctions from e.g. administrative law and criminal law, where the sanction is based on a mere wrongful act. Tort damages are so-called *harm-based* sanctions, which can only be applied after harm has occurred. The strengths and weaknesses of act-based and harm-based sanctions, as well as of *preclusionary* measures which make the behaviour impossible in the first place, is extensively discussed in the economic literature. Hence, it would not be correct to assume that the economic analysis of tort law has to conclude that mere risk-creating activities should already be regarded as torts and should open the possibilities of damages. In cases where it is better to attach the sanction to the act than to the harm, tort law is not the suited instrument.¹¹

7 To be sure, I do not consider compensation to be a goal of punitive damages. First, compensation is not regarded as an independent goal in the economic analysis of tort law to begin with.¹² Second, even if compensation were regarded as a goal of tort law, the fact that punitive damages exceed compensatory damages already shows that compensation cannot be the goal of punitive damages.¹³

II. Economic Reasons for Punitive Damages: Deterrence

A. Probability of Being Held Liable is Below 100%

- 8 The above explained economic line of reasoning that actors derive behavioural incentives from the tort system, implicitly assumed that if a tortfeasor causes losses for which he should be liable, he will indeed be held liable. He then faces the full negative externalities he has caused. However, many reasons exist why the probability of being held liable falls below 100%.
 - 11 R.J. Van den Bergh and L.T. Visscher, Optimal Enforcement of Safety Law, in: R.V. de Mulder (Ed.), Mitigating Risk in the Context of Safety and Security. How relevant is a rational approach?, Rotterdam: Erasmus University Rotterdam (2008) 47 ff. A clear example of the different types of measures is the following. Losses due to traffic accidents as a result of speeding can be targeted by preclusionary measures (e.g. Intelligent Speed Adaptation (ISA) which limits the vehicle to the maximum speed at that location), act-based sanction (fines for speeding) or harm-based sanctions (liability in cases where speeding has caused an accident).
 - This sharply contrasts with the Spanish report, where *P. del Olmo* explains in no. 4 that in Spain compensation is regarded as the *only* normative goal and that prevention is at best seen as a by-product of non-contractual liability. *Menyhárd* in the Hungarian report, on the other hand, mentions in no. 27 that prevention is regarded as a main function of civil liability.
 - See e.g. A. Duggan, Exemplary Damages in Equity: A Law and Economics Perspective, Oxford Journal of Legal Studies (OJLS) 26 (2006) 308.

- 9 First, it might be difficult or even impossible for the victim to prove negligence (if required) or causation on the side of the tortfeasor. Polinsky and Shavell provide the example of an individual who develops a form of cancer that could have developed naturally, but also due to exposure to a man-made carcinogen. The same would obviously hold true if not only the tortfeasor, but also the victim himself might be the cause of the losses, for example an employee of an asbestos processing company who smokes and later develops lung cancer. It could also be the case that the victim does not know the identity of the tortfeasor in the first place, for instance if he was involved in a hit-and-run accident in which the identity of the tortfeasor remains unknown.
- 10 Second, the victim might suffer from what is known in the economic analysis of law as 'rational apathy'. This means that the victim might find it too expensive to bring a suit against the tortfeasor, when comparing the costs to the expected outcome of the trial. This problem might occur especially in situations where the losses are scattered over many victims. The total losses however might be substantial so that it would be socially advantageous if the tortfeasor would be held liable after all.
- 11 Third, the injurer might take steps to avoid detection in cases where he intentionally committed the tort.¹⁵ This obviously lowers the probability of being held liable.
- 12 In all these situations, a tort has been committed and the tortfeasor should be liable for the resulting losses. However, if the victim does not bring a suit, the tortfeasor does not face liability. If not all victims bring a suit, or if they sue but fail because they cannot prove all the required elements, the probability of the tortfeasor being held liable falls below 100%. Hence, the tortfeasor no longer correctly weighs the costs of precautionary measures against the decrease they cause in the total losses, but only the decrease they cause in his expected liability. Given that the losses exceed the expected liability, the tortfeasor does not take adequate precautions and/or engages in the activity too often.
- 13 Punitive damages can ameliorate this situation. ¹⁶ After all, if the probability of being held liable lies below 100%, but the damages to be paid *if* held liable exceed the losses, expected liability can again have the correct size. The factor with which compensatory damages should be multiplied, is the reciprocal of

¹⁴ A.M. Polinksy and S. Shavell, Punitive Damages, in: P. Newman (ed.), The New Palgrave Dictionary of Economics and the Law (1998) 193.

¹⁵ W.M. Landes and R.A. Posner, The Economic Structure of Tort Law (1987) 160.

¹⁶ See e.g. D.D. Ellis, Fairness and Efficiency in the Law of Punitive Damages, S. Cal. L. Rev. 56 (1982) 25 and 26; R.D. Cooter, Punitive Damages for Deterrence: When and How Much?, Alabama Law Review (Ala. L. Rev.) 40 (1989) 1148 ff; A.M. Polinksy and S. Shavell, Punitive Damages: An Economic Analysis, Harvard Law Review (Harv. L. Rev.) 111 (1998) 887 ff.

the probability of being held liable. So, if the probability of being held liable is 50%, compensatory damages should be doubled to provide the correct incentives. Polinsky and Shavell term this reciprocal the 'total damages multiplier'. Punitive damages then consist of total damages minus compensatory damages. The remark of Sebok that 'the size of punitive-damages awards... is predictably determined by the size of the compensatory award' is in my view consistent with this idea of punitive damages as a multiple of compensatory damages.

- An early example of the idea that the sanction should be more severe if the probability of 'being caught' is lower, can be found in the law of the Eshnunna (about 2000 B.C.): someone who was caught in the house or in the field of a palace or temple hierarch during daytime had to pay ten shekels of silver, someone who was caught at night was sentenced to death. A contemporary example is that a person, who causes a traffic accident and leaves the scene to escape sanctioning, not only may have committed a tort in causing the accident, but has also committed a crime in leaving the scene. The possible applicable sanctions now become higher. Economically speaking, this makes sense to counter the decreased probability of conviction. The fact that punitive damages, if they are allowed, are often possible in settings of intentional torts²⁰ where the tortfeasor may try to avoid being caught is consistent with this economic line of reasoning. Hence, punitive damages and criminal law both may improve the incentives which tort law provides through compensatory damages.
- 15 The idea that punitive damages also serve to overcome the problem of rational apathy is present in the concept of treble damages in American antitrust law²¹ and in the European debate on the possibility of double damages in EC antitrust law.²² The prospect of being able to collect more than just compensatory damages might induce victims of law infringements to bring a suit, even

 $^{^{17}}$ Polinksy and Shavell (fn. 14) 193.

¹⁸ Sebok, no. 5.

M.H. Fried, The State, the Chicken, and the Egg: or, What Came First?, in R. Cohen and E.R. Service (eds.), Origins of the State: the Anthropology of Political Evolution 44. This reference was found through the Dutch publication of H.O. Kerkmeester, Punitive damages ter compensatie van een lage veroordelingskans ("Punitive damages to offset a low probability of being convicted"), Nederlands Juristenblad (NJB) 73 (1998) 1808.

²⁰ Sebok, no. 6 and 63; Wilcox, no. 63 ff.; Borghetti, no. 46; Askeland, no. 12; Del Olmo, no. 6(d).

²¹ See e.g. W.P.J. Wils, Should Private Antitrust Enforcement Be Encouraged in Europe?, World Competition (W. Comp.) 26 (2003) 476.

See the Commission Staff Working Paper accompanying the White Paper on Damages actions for breach of the EC antitrust rules, SEC(2008) 404. Also see R. Van den Bergh, W. van Boom and M. van der Woude, The EC Green Paper on Damages Actions in Antritrust Cases - An Academic Comment (2006) 14 <ec.europa.eu/comm/competition/antitrust/actionsdamages/files_green_paper_comments /erasmus_university.pdf>.

though the costs outweigh the expected compensatory damages. Sebok in his report on the United States mentions that punitive damages can serve the goal of compensating costs which might not be covered by compensatory damages.²³ This statement fits well into the idea of overcoming the rational apathy problem, by increasing the expected benefits of a lawsuit. It should be noted that, in as far as punitive damages are indeed able to ameliorate this problem, the probability of being held liable increases, and the damage multiplier should be reduced proportionally to avoid over-deterrence.

A different way of putting the argument is this: Victims who claim damages in essence serve the social goal of deterrence. However, starting a lawsuit entails costs, which are privately borne. This might lead to too few lawsuits being brought. Increasing the expected damages of victims by awarding punitive damages may solve this problem. Wilcox explicitly mentions this line of reasoning in her report on England.²⁴ Polinsky and Shavell however warn against the idea that punitive damages may be used to induce parties to bring suits. After all, lawsuits entail litigation costs. From that perspective, it is better that the damage multiplier is high enough to offset the low probability of being held liable, without actually increasing the number of lawsuits.²⁵

B. Underestimation of Harm

17 If there is a risk that compensatory damages fall short of the true losses of the victim, the injurer does not receive adequate behavioural incentives. This risk especially exists in situations where the losses are difficult to assess, for instance in cases of immaterial losses, or if the subjective valuation of the negative externality as experienced by the victim is difficult to determine. In cases where certain types of losses are excluded from compensation, this problem also occurs. 27

²³ Sebok, no. 36. The relevance of this factor is obviously influenced by the applicable rule regarding recovery of legal expenses. Given that under the English rule, as opposed to the American rule, a prevailing plaintiff can recover (part of) those costs from the defendant, punitive damages are not required to cover these costs. On this issue, also see e.g. V. Behr, Punitive Damages in American and German Law – Tendencies Towards Approximation of Apparently Irreconcilable Concepts, Chicago-Kent Law Review (Chi.-Kent L. Rev.) 78 (2003), 122 ff.

²⁴ Wilcox, no. 113. Also see Menyhárd, no. 7 and 30.

²⁵ Polinksy and Shavell (fn. 14) 197.

²⁶ It therefore should come as no surprise that punitive damages frequently occur in cases of defamation, where immaterial losses are important. See Sebok, no. 6.

²⁷ Ellis, S. Cal. L. Rev. 56 (1982) 27. If the legal system has deliberately excluded certain types of losses from compensation, it remains to be seen if they should be included in the punitive damages, as this would boil down to a circumvention of the exclusion. The relevant question from a Law and Economics point of view then is, whether these losses are rightfully excluded from compensation.

18 Including such losses in punitive damages, however, might be problematic in itself, because in order to determine the correct amount of punitive damages, the same measurement problems that caused these losses to be excluded from compensatory losses, reappear.²⁸ Polinksy and Shavell argue that *if* certain types of losses should be included in damages, they should be included in compensatory damages. After all, punitive damages are measured less accurately than compensatory damages and they are applied much less often. Hence, the problem of incomplete compensatory damages is no good reason to include these losses in punitive damages instead.²⁹

19 If the risk of underestimating the harm of the victim is realistic, could it then make sense to base the damages of the injurer on his gains, if these are higher, instead of on the harm? In absence of the problem of underestimation of harm, economic theory prefers damages to be based on harm rather than on gains. After all, if the injurer compensates the harm he has caused, he internalizes the externality and receives the correct incentives. Basing damages on the assumedly higher gains would provide excessive behavioural incentives. Only if the gains can be labelled socially illicit, the literature sees a reason to remove the injurer's gains.³⁰ Obviously, in countries where 'unjustified enrichment' constitutes a separate legal action, this may be the indicated instrument. However, in debates regarding the goals of tort law, avoiding unjustified enrichment of the tortfeasor is often mentioned as a separate goal of this body of law as well.31 An additional reason why economic literature prefers to base damages on harm instead of gains is that an underestimation of those gains would still make the tort worthwhile to the tortfeasor. This would lead to a decrease in social welfare in situations where these gains are lower than the harm.³² However, given that the problem of underestimating the harm of the victim provides the injurer with inadequate incentives, removing the profits of the injurer may be a good solution after all, because it induces the injurer not to act in the first place. The English category 'conduct calculated by the defendant to make a profit for himself which may well exceed the compensation payable to the claimant', nicely fits into this idea.³³ However, if the risk of underestimation of the losses is not present, the fact that the defendant yields higher gains is not enough to warrant punitive damages from an economic point of view.34

²⁸ Ellis, S. Cal. L. Rev. 56 (1982) 31; *Polinksy* and *Shavell*, Harv. L. Rev. 111 (1998) 940.

²⁹ Polinksy and Shavell, Harv. L. Rev. 111 (1998) 940, 941; Polinksy and Shavell (fn. 14) 194.

³⁰ *Polinksy* and *Shavell*, Harv. L. Rev. 111 (1998) 919.

³¹ See e.g. *Behr*, Chi.-Kent L. Rev. 78 (2003), 137 ff.

³² A.M. Polinsky and S. Shavell, Should Liability Be Based on the Harm to the Victim or the Gain to the Injurer?, Journal of Law, Economics & Organization (J.L. Econ. & Org.) 10 (1994) 427-437.

³³ Wilcox, no. 18 ff.

³⁴ However, also see no. 50 ff. below.

20 Several country reports discuss the possibility of basing damages on gains rather than losses. Wilcox quotes Lord Diplock, who argues that damages might even have to exceed the actual gain of the defendant, even if these gains already outweigh the losses of the victim. After all, if the damages equal the gains, the defendant has nothing to lose from committing the act, while in situations where the plaintiff does not sue or does not succeed in his claim, the injurer keeps his gains.³⁵ Borghetti in contrast states that in France, damages regarding illegal reproduction of a work protected by intellectual property cannot exceed the amount of the illicit profits made by the tortfeasor.³⁶ Also in cases of unfair competition, French courts sometimes take the profits of the defendant into account.³⁷

C. Socially Unaccepted Costs or Benefits

- A next argument mentioned in economic literature regarding punitive damages is the situation where an injurer derives certain benefits from his tort which are regarded as socially unacceptable or alternatively, taking due care would create certain additional costs for the injurer, which are regarded as socially irrelevant. For instance, if I derive pleasure from causing someone else pain, merely having to compensate his losses might not deter me from my act, because after compensating the losses, I still have experienced my pleasure. Alternatively if, e.g., keeping to the speed limit does not only cost me time and efforts, but in addition I loose an unexceptionally large thrill of speeding in a populated area, the mere threat of liability might not adequately deter me from speeding.
- 22 If the utility the injurer derives from the act is regarded as socially illicit, the act should be deterred completely. Damages then should be so high that they deter even the injurer who yields these unaccepted benefits or bears these unaccepted costs.³⁸ According to Polinksy and Shavell, this reason for punitive damages is limited in scope, because many socially undesirable acts do not seem to be associated with social illicit utility, since they are not aimed at causing harm.³⁹
- 23 In addition to this relativization, the idea of labelling costs or benefits as socially illicit is criticized by, among others, Friedman. Labelling certain activities as socially unacceptable, even if they would yield more benefits to the injurer than losses to the victim, assumes the conclusion that such acts are unde-

³⁵ Wilcox, no. 183.

³⁶ Borghetti, no. 12.

³⁷ Borghetti, no. 29.

³⁸ Ellis, S. Cal. L. Rev. 56 (1982) 32; R.D. Cooter, Economic Analysis of Punitive Damages, S. Cal. L. Rev. 56 (1982) 87 ff.; Polinksy and Shavell (fn. 14) 194.

³⁹ Polinksy and Shavell (fn. 14) 194. Also see Sebok, no. 29.

sirable, instead of *proving* it.⁴⁰ However, given that with such intentional torts, the gains (if they exist at all) are often outweighed by the losses,⁴¹ they will often indeed be socially undesirable because they lower social welfare. It therefore makes economic sense that in cases of intentional torts, punitive damages occur relatively frequently.⁴²

D. Induce Voluntary Transfers

- 24 In situations where transaction costs are low enough for parties to be able to negotiate with each other about the price to pay to transfer an entitlement, economic theory has a preference for voluntary transfers over involuntary transfers. In economic terms, entitlements in such situations are protected by *property rules*, where the only accepted way of transfer is a voluntary transaction. In situations of high transaction costs, on the other hand, entitlements are protected by *liability rules*, where the entitlement can also be taken without consent of the owner. The taker subsequently has to pay an objectively determined amount to the owner of the entitlement, i.e. damages. ⁴³ Property rule protection in situations of high transaction costs could effectively hinder a desirable reallocation of resources to take place, because it is too expensive for the parties involved to achieve a transaction.
- 25 Property rules are preferred in settings of low transaction costs because the parties involved are assumed to know their own preferences better than a judge does. Hence, the price that results in a voluntary transfer is a better assessment of the valuations of the parties involved than the damages that are set after a tort. In the words of Landes and Posner: 'When the costs of voluntary market transactions are low, the property approach is economically preferable to the liability approach because the market is a more reliable register of values than the legal system.' ⁴⁴ In addition, the assessment of losses, which
 - ⁴⁰ D.D. Friedman, An Economic Explanation of Punitive Damages, Ala. L. Rev. 40 (1989) 1128 ff. and D.D. Friedman, Law's Order: What Economics Has to Do with Law and Why it Matters (2000) 230 ff. Also see D.D. Haddock, F.S. McChesney and M. Spiegel, An Ordinary Economic Rationale for Extraordinary Legal Sanctions, California Law Review (Cal. L. Rev.) 78 (1990) 12; S. Marks, Utility and Community: Musings on the Tort/Crime Distinction, Boston University Law Review (B.U.L.Rev.) 76 (1996) 215 ff; K.N. Hylton, Punitive Damages and the Economic Theory of Penalties, The Georgetown Law Journal (Geo. L.J.) 87 (1998) 464 ff.
 - 41 D.D. Ellis, An Economic Theory of Intentional Torts: A Comment, International Review of Law and Economics (Int. Rev. Law and Econ.) 3 (1983) 50.
 - ⁴² Sebok, no. 6.
 - See, among others, G. Calabresi and A.D. Melamed, Property Rules, Liability Rules, and Inalienability: One View of the Cathedral, Harv. L. Rev 85 (1972) 1089-1128; J.E. Krier and S.J. Schwab, Property Rules and Liability Rules: The Cathedral in Another Light, New York University Law Review (N.Y.U. L. Rev.) 70 (1995) 440-483; L. Kaplow and S. Shavell, Property Rules versus Liability Rules: An Economic Analysis, Harv. L. Rev 109 (1996) 713-790.
 - 44 Landes and Posner (fn. 7) 31.

is necessary under liability rules, can be very costly. Furthermore, even though parties could also negotiate in a setting of low transaction costs if entitlements were protected by liability rules, such a form of protection would cause several problems: (1) if the owner successfully has bought off a potential infringer, another possible injurer might try to take away the entitlement after all; (2) the owner and potential infringers may invest resources to avoid respectively enable the infringement (e.g. a burglar alarm and instruments to circumvent the alarm), which is socially undesirable; and (3) the owner may, after his entitlement is taken from him, try to take it back from the infringer, who then tries again, et cetera. Finally, if owners have to accept infringements of their entitlements, they might have fewer incentives to invest in acquiring property.

- Punitive damages may induce a potential injurer to seek a voluntary transfer rather than to commit the tort, if liability including punitive damages is more costly than seeking the voluntary transfer. Without these punitive damages, the injurer could decide to commit the tort after all, if compensatory damages fall short of his private gain. Given the difficulties in correctly assessing damages, the private loss to the victim, however, might exceed the gains to the injurer, so that the involuntary transfer lowers social welfare. If only a voluntary transfer would have been possible, this problem would not occur, because the transfer then only takes place if the potential buyer values the entitlement higher than the potential seller.
- 27 Polinsky and Shavell mention an additional problem. If injurers know that compensatory damages fall short of true losses, they might spend resources to look for property which they can take without having to compensate its full value and victims invest in avoiding this. Both expenditures lower social welfare. The authors provide the example of copyright infringements.⁴⁷ Here again, the construction of basing damages on the profits of the infringer may be a good solution to avoid this problem. Several country reports show that infringements of intellectual property are indeed situations in which punitive damages can be granted, or at least where damages may be based on the gains of the infringer rather than the losses of the victim.⁴⁸
- 28 The line of reasoning that a higher sanction induces voluntary transfers is also an important argument in the economic theory of criminal law. The idea then is that, even if the probability of being held liable is 100%, mere damages would not be an adequate remedy against certain acts. After all, if an injurer in a setting of low transaction costs would take an entitlement without consent

⁴⁵ Krier and Schwab, N.Y.U. L. Rev. 70 (1995) 440-483; Kaplow and Shavell, Harv. L. Rev 109 (1996) 713-790.

⁴⁶ Kaplow and Shavell, Harv. L. Rev 109 (1996) 766 ff.

⁴⁷ Polinksy and Shavell (fn. 14) 195.

⁴⁸ J. Neethling, report South Africa, no. 26; Wilcox, no. 39 ff.; Borghetti, no. 11 ff.; Jansen and Rademacher, no. 12 ff.; Del Olmo, no. 15 ff.

and pay damages with a 100% certainty, in essence he converts a property rule protection into a liability rule protection. In order to avoid this, the sanction has to outweigh the losses. The prospective injurer will then choose to acquire the entitlement in a voluntary transfer instead of through an involuntary taking.

- 29 For example, suppose that I value the car of my neighbour higher than its market value, but that the subjective valuation of my neighbour for his car is only known to him. If I *buy* the car from my neighbour in a voluntary transaction, it is certain that I value the car higher than he does, because the price I was willing to pay was high enough for him to sell the car. Our transaction hence has improved our situation. However, were I to *steal* his car and subsequently would only have to pay the market value, our joint situation may worsen. After all, even though my personal valuation of the car was higher than the market value so that I was willing to pay the objective price, it is possible that it is lower than the subjective valuation of my neighbour. Hence, the amount I have to pay in the case of theft should be higher than the value of the car.
- 30 This higher sanction than the value of the entitlement is known as a 'kicker'. 49 Punitive damages likewise may serve to avoid the injurer from converting a property rule into a liability rule. 50 Disgorging all gains the injurer yielded by committing the tort fits into this idea.

III. Economic Reasons for Punitive Damages: Punishment

A. Tort Law and Criminal Law

31 In the economic analysis of crimes and criminal law, maximization of social welfare is often regarded as the most important goal.⁵¹ Hence, both tort law and criminal law serve the same goal: deterrence. In the literature on the topic of 'optimal enforcement', several arguments are developed as to why criminal law is needed as a deterrence mechanism alongside tort law.⁵²

⁴⁹ Calabresi and Melamed, Harv. L. Rev 85 (1972) 1126; A.K. Klevorick, On the Economic Theory of Crime, in J.R. Pennock and J.W. Chapman (eds.), Criminal Justice, NOMOS XXVII (1985) 289-309; J.L. Coleman, Crime, Kickers and Transaction Structures, in J.R. Pennock and J.W. Chapman (eds.), Criminal Justice, NOMOS XXVII (1985) 311-328.

⁵⁰ Haddock, McChesney and Spiegel, Cal. L. Rev. 78 (1990) 17 ff.

⁵¹ See e.g. G.S. Becker, Crime and Punishment: An Economic Approach', Journal of Political Economy (J.P.E.) 76 (1968) 169-217; R.A. Posner, Economic Analysis of Law (6th ed. 2003) 217 ff; Shavell (fn. 7) 543 ff.; Cooter and Ulen (fn. 7) 510.

See e.g. S. Shavell, The Optimal Structure of Law Enforcement, J. L. & Econ. 36 (1993) 255-287.

- 32 First, if the probability of being held liable in tort law is below 100%, the expected sanction is not high enough to deter potential injurers. Above, this same topic is presented as an argument if favour of punitive damages. However, in countries where punitive damages are not permissible, criminal law may be used to increase the expected sanction.⁵³
- 33 Second, if the injurer is judgment proof, a financial sanction might not adequately deter him. Criminal law can apply non-monetary sanctions, which may be able to provide the necessary incentives. Obviously, the use of punitive damages might suffer from this same problem: if the injurer cannot pay the punitive damages, he might not be deterred by them. Applying criminal law could then be a necessary step.
- 34 Third, and related, criminal law stigmatizes the convict. The fear of this stigma may provide behavioural incentives to the potential wrongdoer, where financial sanctions might not have been enough.⁵⁴
- 35 Fourth, the use of criminal law solves the rational apathy problem. After all, it is no longer the victim who has to initiate the procedure, but the State.⁵⁵ Also in situations where the victim does not have enough information regarding the identity of the wrongdoer or about the existence of an infringement in the first place, the fact that the State initiates the procedure is an advantage of the criminal law system. After all, the State can make use of investigative methods and information systems such as fingerprint and DNA databases, which the victim cannot apply.⁵⁶
- 36 The drawback of criminal law is that the administrative costs generally outweigh those of the tort system. After all, the severity of the sanction and the fact that criminal law does not shift an existing loss but rather adds another loss, calls for procedural safeguards to avoid wrongful convictions.⁵⁷ In addi-

⁵³ See e.g. *Shavell*, J. L. & Econ. 36 (1993) 266, 276ff.

For economically oriented literature on stigmatization, see e.g. E. Rasmusen, Stigma and Self-Fulfilling Expectations of Criminality, J. L. & Econ. 39 (1996) 519-543; D.M. Kahan and E.A. Posner, Shaming White-Collar Criminals: A Proposal for Reform of the Federal Sentencing Guidelines, J. L. & Econ. 42 (1999) 365-391; P. Funk, On the Effective Use of Stigma as a Crime-Deterrent, European Economic Review (E.E.R.) 48 (2004) 715-728.

Due to limited resources, the State will not be able to respond to all contraventions, so that some violations will go undeterred. It is, however, important to realise that the idea of optimal enforcement does not aim at maximum deterrence, but at optimal deterrence, where the costs and benefits of additional enforcement measures are weighed. See e.g. A.M. Polinksy and S. Shavell, Enforcement Costs and the Optimal Magnitude and Probability of Fines, J. L. & Econ. (35) 1992, 133, 138.

⁵⁶ Shavell, J. L. & Econ. 36 (1993) 269, 278.

⁵⁷ From an economic point of view, this is a crucial difference with punitive damages, which are intended to offset the problems which frustrate the preventive potential of tort

tion, non-monetary sanctions (i.e. imprisonment) are much more expensive to execute than monetary sanctions (i.e. fines and damages). Hence, criminal law should be used as an *ultimum remedium*. The fact that criminal law up to a certain degree may be 'self-enforcing' (in the sense that most people do not want to be associated with crimes) strengthens the ultimum remedium character. After all, if too many acts are criminalized, the self-enforcing character is weakened.

B. Punitive Damages and Punishment

- 37 Given the abovementioned economic approach to criminal law, most punishment arguments for punitive damages are actually deterrence arguments in disguise. The punitive character is, at least from an economic point of view, intended to strengthen the preventive functioning of tort law, where compensatory damages are not sufficient. Increasing total damages is then necessary to offset the too low probability of being held liable, to counterbalance socially illicit costs or benefits or to induce voluntary transfers.
- 38 The advantage of using punitive damages rather than criminal law is that the high administrative costs of criminal law are avoided, and that its ultimum remedium character is maintained. However, the judgment proof problem poses limits to the possibility of punitive damages to solve the problem of the too low probability of being held liable. The non-monetary sanctions of criminal law are then needed to provide the correct incentives.⁵⁸
- Analysed like this, punishment is not a goal in itself, but it serves the goal of prevention. Polinksy and Shavell however also analyse the separate goal of punishment, where the punishment objective is derived from the desire of individuals to have blameworthy parties appropriately punished. The correct level of punishment then depends on the reprehensibility of the party's actions. If the defendant is a firm, it is difficult to punish the blameworthy individuals within the firm through punitive damages. First, it might be difficult for the firm to find the culpable employee. Second, it remains to be seen if the internal sanction that the firm applies is influenced by the punitive damages. Third, especially in cases of dispersed responsibility, it is doubtful whether a culpable employee exists. In the end, it is often the shareholders and the customers who get 'punished', not the blameworthy employee.⁵⁹
- 40 The correct amount of punitive damages for the goal of punishment is determined by the reprehensibility of the wrongdoer's actions. The punishment goal implies that the culpable injurer should suffer a utility loss that corre-

law. Punitive damages hence do not add another loss, but compensate for the too low probability of being held liable, the underestimation of losses, et cetera.

⁵⁸ Also see *Haddock*, *McChesney* and *Spiegel*, Cal. L. Rev. 78 (1990) 48 ff.

⁵⁹ *Polinksy* and *Shavell*, Harv. L. Rev. 111 (1998) 948 ff.

sponds to the blameworthiness of his behaviour. This entails that the level of wealth of the defendant is assessed, because the utility decrease that is caused by the duty to pay damages depends *inter alia* on his wealth. Hence, the wealthier the defendant, the higher punitive damages should be to reach the punishment goal.⁶⁰ Wealth of the defendant is a relevant factor in the United States and in England.⁶¹

IV. Tension between the Goals of Deterrence and Punishment

- 41 After having discussed both the deterrence goal and the punishment goal of punitive damages, it is possible to highlight a few tensions that exist between both goals.
 - A. When Should Punitive Damages be Awarded?
- 42 The deterrence objective of punitive damages dictates that they are awarded if the probability that the injurer is held liable is below 100%, if certain costs or benefits from the injurer are regarded as socially illicit, if compensatory damages systematically under-compensate the victim and if the injurer should be induced to seek a voluntary transfer. The punishment goal, in as far as it does not serve the deterrence goal, states that punitive damages should be awarded if the behaviour of the injurer was reprehensible.
- 43 Reprehensibility of the behaviour is not directly relevant for the deterrence goal. In cases where reprehensible behaviour almost certainly leads to liability, punitive damages are not required to adequately deter the tortfeasor. The Exxon Valdez case is regarded in the economic literature as an example of this, because the probability of a tort suit following the accident was close to 100%.⁶² The reprehensibility of the captain regularly being drunk while on duty then is irrelevant. On the other hand, irreprehensible behaviour with a low probability of being held liable might require punitive damages after all. Only if the reprehensibility is connected to the social illicitness of gains or costs, it might be a relevant factor for the deterrence goal.⁶³
- 44 The Supreme Court itself states that the problem of a low probability of being detected and the problem of rational apathy do not play a role in the Exxon Valdez case and hence cannot justify high punitive awards: 'Heavier punitive awards have been thought to be justifiable when wrongdoing is hard to detect (increasing chances of getting away with it)... or when the value of injury and the corresponding compensatory award are small (providing low incentives to

⁶⁰ Polinksy and Shavell, Harv. L. Rev. 111 (1998) 953.

⁶¹ Sebok, no. 78 ff., Wilcox, no. 76.

⁶² Polinksy and Shavell, Harv. L. Rev. 111 (1998) 904. For a critical view, see Hylton (fn 40) 452 ff

⁶³ Polinksy and Shavell (fn. 14) 196.

sue).' (Section IV B). 'We know, for example, that Congress devised the treble damages remedy for private antitrust actions with an eye to supplementing official enforcement by inducing private litigation, which might otherwise have been too rare if nothing but compensatory damages were available at the end of the day...That concern has no traction here, in this case of staggering damage inevitably provoking governmental enforcers to indict and any number of private parties to sue.' (Section IV F2). 'In a well-functioning system, we would expect that awards at the median or lower would roughly express jurors' sense of reasonable penalties in...cases (again like this one) without the modest economic harm or odds of detection that have opened the door to higher awards' (Section IV F3).

B. Wealth of the Defendant

Above it became clear that the wealth level of the defendant is relevant for the punishment goal, because it influences the utility loss experienced by the injurer when having to pay a certain amount of damages. For the deterrence goal, however, wealth is in principle irrelevant, besides the above-discussed topic of judgment proof.⁶⁴ Liability serves the goal to internalize the externalities caused by the injurer. By compensating the losses, multiplied by a factor to offset the too low probability of being held liable, this internalization is reached. Further increasing damages on the basis of the level of wealth would lead to over-deterrence. However, if injurers cannot insure against punitive damages and they are risk-averse, expected liability need not be full in order to provide adequate behavioural incentives. In such a situation, the poorer the injurer is, the lower the punitive damages award can be to still be able to offset the too low probability of being held liable.⁶⁵

C. Insurance Against Punitive Damages

46 If punitive damages are intended to offset the too low probability of being held liable, injurers should be able to insure against punitive damages. It is well established in the economic literature that the availability of liability insurance increases social welfare, provided that the problems of adverse selection and moral hazard can be adequately tackled. Injurers should hence be able to cover liability for the expected losses. Given that punitive damages serve to increase expected liability to the level of expected losses, they should be insurable. It is argues that the problem of moral hazard is a reason not to allow insurance against punitive damages. The same is mentioned by Ebert in her report on Liability Insurance. She states that the moral hazard problem

⁶⁴ See e.g. *Cooter*, Ala. L. Rev. 40 (1989) 1176, 1177.

⁶⁵ *Polinksy* and *Shavell*, Harv. L. Rev. 111 (1998) 913.

⁶⁶ Shavell (fn. 7) 257 ff.

⁶⁷ Cooter, Ala. L. Rev. 40 (1989) 1182 ff.; Polinksy and Shavell, Harv. L. Rev. 111 (1998) 932 ff.; Polinksy and Shavell (fn. 14) 197.

⁶⁸ Ellis, S. Cal. L. Rev. 56 (1982) 74.

cannot be adequately addressed because there are too few verdicts to estimate the individual risk.⁶⁹ However, given that the economic approach favors a more regular use of punitive damages, which is far better predictable than the current practise and which is strongly connected to the compensatory damages, this problem in my view is rather limited.

- 47 The story obviously is completely different from the punishment perspective. In order for punitive damages to be able to punish the injurer, he should experience the negative utility that is caused by the duty to pay. Insurance against punitive damages would frustrate this objective, so that insurance against punishments should not be allowed. However, it remains to be seen if the injurer, especially if it is an individual, would be able to pay the punitive damages without insurance. If the injurer is judgment proof, punitive damages would not be able to punish adequately. As is already explained in no. 33 above, the judgment proof problem forms an argument for criminal law as addition to tort law, because it can make use of non-monetary sanctions.
- Cooter and Ulen argue, based on the judgment proof problem and the need for non-monetary sanctions that arise from it, that insurance against criminal fines can make economic sense. In order to combat the possible problem of moral hazard, the insurance company would want to monitor policy-holders to deter them from committing crimes. Private enforcement by insurance companies would then supplement public enforcement by the police.⁷² The same line of reasoning could be followed with respect to punitive damages. However, as Shavell has pointed out, insurance against the financial consequences of wrongful behaviour will not be bought, because it is too expensive. It is cheaper for the potential wrongdoer not to commit the wrong, than to commit it and collect the insurance benefits. The premium for such insurance outweighs the benefits.⁷³ Be this as it may, Cooter and Ulen's argument is based on the deterrence goal of criminal law, not on the goal of punishment in itself. In my view, this latter goal still cannot be reconciled with insurance against fines or punitive damages, because insurance would bar the negative consequences of the sanction from fully reaching the wrongdoer.

⁶⁹ I. Ebert, report on Liability Insurance, no. 5. However, in no. 9 she mentions the use of deductibles as a possible way to avoid moral hazard.

Also see A.M. Polinsky and S. Shavell, The Optimal Trade off between the Probability and Magnitude of Fines, The American Economic Review (A.E.R.) 69 (1979) 885 ff.

⁷¹ It does not become clear from the reports on the United States and England whether this problem is a reason for considering the wealth of the defendant. Seboks remark in no. 78 that in Maine, "the punitive damage award must take into account defendant's wealth in order to adequately affect the defendant and fulfil the award's punishment and deterrent purposes" rather pertains to the opposite situation, where the defendant is so wealthy that damages may be too low to have a real effect. Wilcox' remark in no. 76 that "the idea is to take the profit out of wrongdoing" suggests the same.

⁷² *Cooter* and *Ulen* (fn. 7) 514.

⁷³ Shavell (fn. 7) 264 ff.

49 Faure and Heine also argue that insurance of fines as such is not to be considered undesirable. It can increase social utility, provided that the insurance company can monitor the insured so that the preventive function of the criminal sanction is shifted to the insurer. Here again, the authors regard the sanction as an instrument of deterrence. The argument that insurance cannot be reconciled with the goal of punishment as such, therefore in my view still holds true.

- D. Should Punitive Damages be Based on Gains or Losses?
- 50 Besides the points already discussed above in no. 19, the differences between the goals of prevention and punishment are also relevant for answering the question whether punitive damages should be based on the gains to the wrongdoer or the losses of the victim(s). Hylton discusses the possible goals of loss internalization (i.e. confronting the injurer with the losses he has caused) and gain elimination (removing the benefits the injurer obtains from his behaviour) and he distinguishes the situation where the gains are smaller than the losses from the opposite situation.⁷⁵
- 51 It has been argued that the internalization approach is the better option, because: (1) if gains exceed losses, basing damages on the losses compensates the victim while still enabling the injurer with high enough gains to perform the behaviour; ⁷⁶ and (2) if gains are smaller than the losses, the behaviour is undesirable and any sanction that exceeds the gains would be able to deter the injurer. However, Hylton challenges this pro-internalization view.
- According to Hylton, if the gains are smaller than the losses, the optimal sanction might differ from the loss-internalization level. If due to excessive discounting some people would not be deterred by the mere elimination of gains, higher sanctions might be required, even higher than the loss-internalization level. If, on the other hand, higher sanctions would lead to more violent behaviour (e.g. to avoid detection), a lower sanction might be optimal. Another reason why the sanction should not always be based on loss-internalization, is that the losses are sometimes more difficult to assess than the gains, e.g. in antitrust cases.
- 53 In situations where the gains exceed the losses, this does not necessarily mean that the loss-internalization level is the correct basis for the sanction. After all, there are more social losses than the mere losses of the victim (e.g. costs of

M. Faure and G. Heine, The Insurance of Fines: the Case of Oil Pollution, The Geneva Papers on Risk and Insurance – Issues and Practice (G.P.R.I.I.P.) 16 (1991), 47.

⁷⁵ *Hylton* (fn 40) 423 ff.

⁷⁶ It is interesting to note that exactly this possibility of the tortfeasor still performing the behaviour because his gains exceed the losses of the victim, is sometimes regarded as an argument for punitive damages. See *Wilcox*, no. 180 ff.; *Borghetti*, no. 36; *Jansen* and *Rademacher*, no. 11.

law enforcement and litigation and the losses of others than the direct victim). Basing the sanction on the gains instead of on the losses might then be a good solution, because it deters the behaviour altogether. Especially in cases where the possible losses due to over-deterrence are small (e.g. in cases of theft, where the potential thief who values the good higher than the owner still has the option of buying it from the owner), this is a good solution.⁷⁷

54 In order to determine the correct amount of punitive damages, it is therefore important to assess the gains to the injurer and the social losses of his behaviour. Reprehensibility of the conduct might serve as an indication that the gains do not exceed the social losses, because the gain only materializes at the expense of a loss of the victim. The choice between loss-internalization damages and gain-removing damages ultimately boils down to the question whether the goal is to induce the potential injurer to choose an optimal level of care and activity, or to completely deter his behaviour. In the first scenario, losses should be internalized, in the second, gains should be eliminated. According to Hylton, in many cases of punitive damages, gain-elimination is preferable and the risk of over deterrence is not relevant, because the behaviour should be deterred completely.

V. Conclusion

- In the economic analysis of law, prevention and spreading of losses are regarded as the most important goals of tort law. Damages are the instrument with which these goals are aimed for. The economic analysis of criminal law focuses at deterrence as well. Punitive damages are therefore primarily regarded as an instrument which can, when necessary, improve the deterrent function of tort law. They may be required in this respect to offset the fact that the probability of being held liable falls short of 100%, to respond to the problem that the harm is sometimes underestimated, to correct for socially unaccepted costs or benefits of the tortfeasor or to induce the latter to seek a voluntary transfer with the victim rather than to commit the tort.
- 56 Given the focus on deterrence, the goal of punishment in itself does not play a major role in the economic analysis of law in general, or in the economic analysis of punitive damages specifically. If both goals are treated separately, nonetheless, some tensions prove to exist. They regard the role of reprehensibility, the wealth level of the defendant, the question whether insurance

⁷⁷ Hylton (fn 40) 438.

⁷⁸ Hylton (fn 40) 441.

⁷⁹ Hylton (fn 40) 456. Also see M.F. Grady, Punitive Damages and Subjective States of Mind: A Positive Economic Theory, Ala. L. Rev. 40 (1989) 1214 ff, who argues that courts rightfully distinguish between inadvertent negligence (which often is efficient) and deliberate negligence (which is not).

 $^{^{80}\,}$ Hylton (fn 40) 467.

against punitive damages is allowed and the questions whether damages should be based on harm to the victim or gains of the tortfeasor.

57 When connecting the economic insights to the country reports, it became clear that some reports contrast with the economic analysis, e.g. the Spanish report which states that compensation is regarded as the only normative goal of noncontractual liability and the emphasis that some reports put on the punishment goal in itself. In many instances, however, the economic analysis nicely fits the legal treatment of punitive damages, e.g. the connection between the size of compensatory damages and punitive damages, the role of punitive damages in overcoming the problem of rational apathy and the fact that punitive damages are often granted in situation of immaterial losses or with intentional torts.